Review Article

Evolution of Black-Box Models Based on Volterra Series

Daniel D. Silveira, Thiago V. N. Coelho, and Alexandre Bessa dos Santos

Department of Electrical Circuits, Federal University of Juiz de Fora, Campus Universitario, Rua Jose Lourenco Kelmer, s/n, 36030-900 Juiz de Fora, MG, Brazil

Correspondence should be addressed to Daniel D. Silveira; danieldsilveira@gmail.com

This paper presents a historical review of the many behavioral models actually used to model radio frequency power amplifiers and a new classification of these behavioral models. It also discusses the evolution of these models, from a single polynomial to multirate Volterra models, presenting equations and estimation methods. New trends in RF power amplifier behavioral modeling are suggested.

1. Introduction

Modeling nonlinear systems has shown to be a challenge in different areas of science. Most natural phenomena and physical devices present a nonlinear behavior. In this sense, it is very useful to classify nonlinear systems, so the right model can be used for each system.

In [1], a complete classification of nonlinear systems is given. If any of the following phenomena occurs, a nonlinear dynamic model has to be used:

(1) asymmetric responses to symmetric input signal changes (ASYM),

(2) generation of higher-order harmonics in response to a sinusoidal input (HARM),

(3) input multiplicity, meaning that one steady-state response corresponds to more than one steady-state input (IM),

(4) output multiplicity, meaning that one steady-state input corresponds to more than one steady-state output (OM),

(5) generation of subharmonics in response to any periodic input (SHAM),

(6) highly irregular responses to simple inputs like impulses, steps, or sinusoids (CHAOS),

(7) input-dependent stability (IDS).

A nonlinear system is classified due to phenomena presence as follows:

(i) mild: ASYM, HARM, and IM,

(ii) intermediate: IDS,

(iii) strong: OM, SHAM, and CHAOS.

In electrical engineering, signal amplifiers are very often used for different purposes. One of the main uses is for signal transmission, where a power amplifier (PA) is needed. A radio frequency (RF) power amplifier is a typical nonlinear system. Even when the transistor is operating in a quasilinear region, driven with small variance input signals, the output signal has nonlinear components, due to the physics of the transistor.

A PA behavioral model (BM) remains in the mildly nonlinear class. The known PAs to be modeled present these characteristics in normal operation conditions, when tested with sinusoid stimuli. None of the other phenomena (OM, SHAM, CHAOS, or IDS), which imply the need of intermediate or strong nonlinear dynamic models, were observed in amplifier measurements.

This paper will present a classification of BMs and discuss the evolution of BMs based on VS used in the modeling of RF PAs, from some of the simplest models to recent ones reported in the literature.
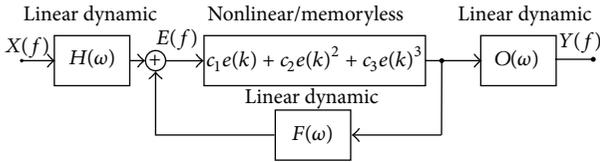
FIGURE 1: A PA representation using a nonlinear feedback structure [2].

2. Classification of Power Amplifier Behavioral Models

The classification of PA BMs used in this work is as in [2]:

(i) memoryless (ML): the output envelope reacting instantaneously to variations in the input envelope,

(ii) linear memory (LM): BMs that account for envelope memory effects attributable to the input and output matching networks' frequency characteristics,

(iii) nonlinear memory (NLM): dynamic interaction of nonlinearities through a dynamic network.

Figure 1 is used by the authors to classify the various BMs. Memoryless models are represented by the block "nonlinear/memoryless." Linear memory models are models that account for the "linear dynamic": $H(\omega)$ and $O(\omega)$ blocks (matching networks). Models that care for nonlinear memory contain all previously mentioned blocks and the feedback path with the block $F(\omega)$, attributed to electrothermal and/or bias circuitry dynamics.

Although this classification was very complete by the time of this paper [2], further developments in the field were noticed, and so an extension of this classification has to be done. Some models were recently reported:

(i) the pruned Volterra series (rVS1) [3],

(ii) the pruned Volterra series (rVS2) [4],

(iii) a parallel cascade model (PCM) composed of a static nonlinearity and a reduced Volterra model (PNLrVS) [5],

(iv) a parallel cascade subsampled reduced Volterra series with the first branch composed of a static nonlinearity and a rVS model and other branches being rVS models, all with the same memory depth (PssVS), as detailed in [5].

The last model (PssVS) is a multirate parallel reduced VS. This model presented the best performance among all BMs reported in the literature, in an extensive comparison presented in [6].

A timeline of publications related to the accuracy of VS models is presented in Figure 2, based on a search in the database of the IEEE Xplore. This search was focused on PA BMs.

A graphical representation of all these models from the initial classification of nonlinear systems to the modern VS models is presented in Figure 3.
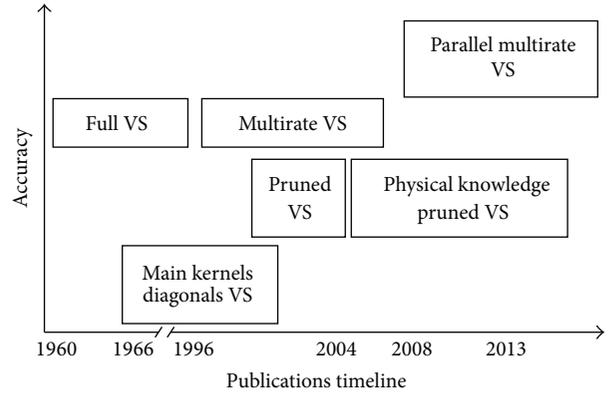

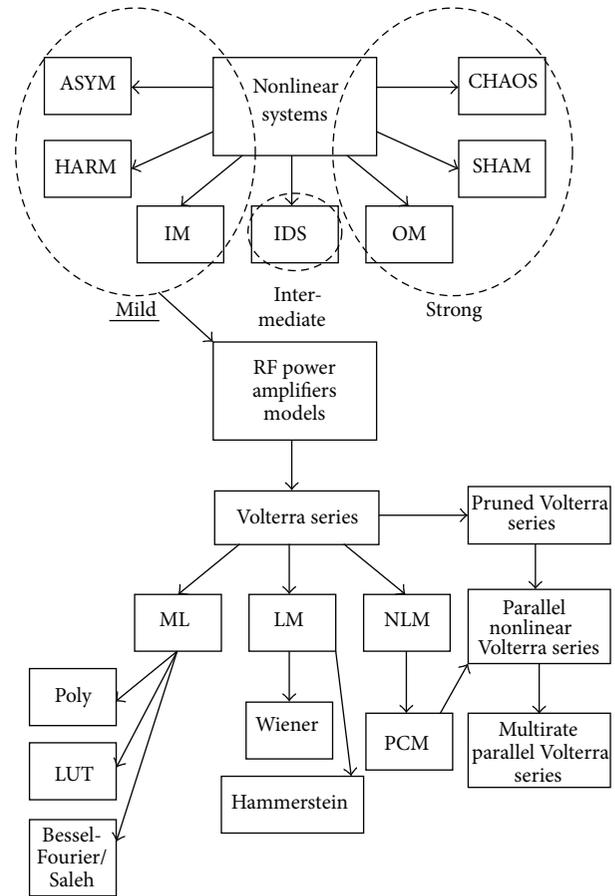FIGURE 2: A timeline of Volterra series models based on accuracy.

FIGURE 3: Representation of various classes of nonlinear models and behavioral models used for RF power amplifiers modeling.

3. Evolution of Power Amplifier Behavioral Models

This section presents the evolution of PA BMs and their equations from nonlinear memoryless models to reduced Volterra Series models.
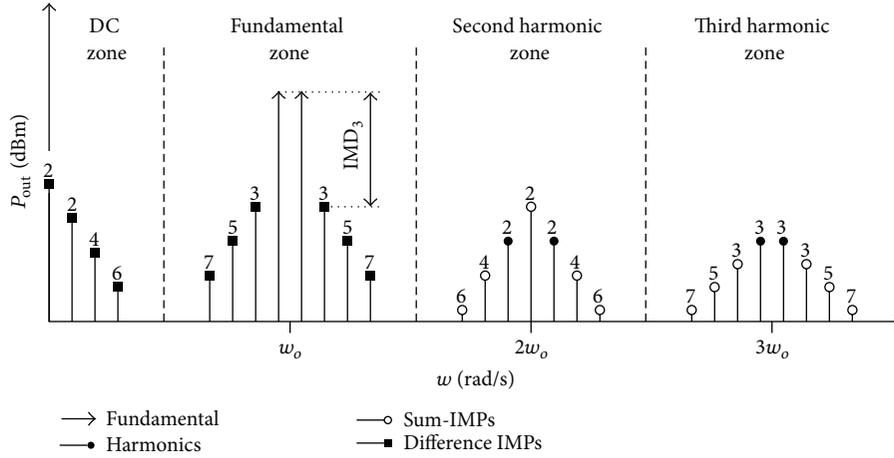
FIGURE 4: Frequency-domain response of a nonlinear amplifier supplied with a two-tone test input signal.

3.1. Nonlinear Memoryless. The nonlinear part of an amplifier model represents the intermodulation distortion (IMD), or the static part, and is usually composed of polynomials or other nonlinear functions (e.g., tangent-sigmoids, look-up tables). These models do not account for dynamics of the system.

In this section, some of the memoryless nonlinear models will be covered.

3.1.1. Power Series. A nonlinear system can be represented by a power series:

$$y(k) = \sum_{p=0}^{P} c_p u^p(k), \tag{1}$$

where c_p are the polynomial coefficients and P is the order.

A simple form to estimate a power series is using linear regression methods, as polynomial coefficients are linear in parameters. The polynomial regression matrix $\mathbf{U}$ for N measurements, a polynomial degree P, and the parameter vector $\widehat{\boldsymbol{\theta}}_P$ is [7]

$$\mathbf{U} = \begin{bmatrix} 1 & u(1) & u^2(1) & \cdots & u^P(1) \\ 1 & u(2) & u^2(2) & \cdots & u^P(2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & u(N) & u^2(N) & \cdots & u^P(N) \end{bmatrix}, \tag{2}$$

$$\widehat{\boldsymbol{\theta}}_P = \begin{bmatrix} c_0 & c_1 & \cdots & c_P \end{bmatrix}^T.$$

Then LS equation can be applied:

$$\widehat{\boldsymbol{\theta}}_P = \left(\mathbf{U}^H \mathbf{U} \right)^{-1} \mathbf{U}^H \mathbf{y}. \tag{3}$$

This regression matrix results in a Hessian with a high condition number (CN), defined as the ratio of the largest to smallest singular value in the singular value decomposition of a matrix [8]. A large CN is not desirable in the estimation

process, as it implies that small errors in the input can cause large errors in the output. For very high CN, orthogonalization of the regression matrix is required to find the solution.

The Hessian CN can be improved if orthogonal polynomials are used. These polynomials are derived based on the input signal used in the system. Thus, the regressors are closer to the ideal situation for a Hessian (regressors mutually orthogonal).

For real valued input signals, Chebyshev (derived for single tones) and Hermite (derived for Gaussian distribution) polynomials are typically applied.

3.1.2. Baseband Power Series. Although polynomial LS estimation is a reasonable possibility to calculate the IMD components, it generates also "out-of-band" harmonics, as shown as an example in Figure 4. The numbers above show the respective IMD products (second, third, and so on) of a Taylor series expansion when a two-tone excitation is applied [9].

These are uninteresting for predistortion purposes, the main objective of behavioral modeling. To solve this problem, the first-zone equivalent (or baseband) polynomial is necessary. It can be derived writing the input signal as [10]

$$u(k) = \operatorname{Re}\left[u(k) e^{jwk} \right] \tag{4}$$

$$= \frac{1}{2} \left[u(k) e^{jwk} + u(k)^* e^{-jwk} \right]. \tag{5}$$

So, a binomial based expression for $u^n(k)$ can be obtained:

$$u^p(k) = \left\{ \frac{1}{2} \left[u(k)e^{jwk} + u(k)^* e^{-jwk} \right] \right\}^p \tag{6}$$

$$= \frac{1}{2^p} \sum_{s=0}^{p} \binom{p}{s} u^p(k) u^{p-s}(k) e^{jw(2s-p)k}. \tag{7}$$

Only the terms where p is odd and $2s - p = \pm 1$ can contribute for the first-zone output or $s = (p+1)/2$ and $s = (p-1)/2$. Then (7) can be written as

$$u_{1z}^p = \frac{1}{2^p} \sum_{s=0}^{p} \begin{pmatrix} p \\ \dfrac{p+1}{2} \end{pmatrix} [u(k)]^{(p+1)/2} [u^*(k)]^{(p-1)/2} e^{jwk}$$

$$+ \frac{1}{2^p} \sum_{s=0}^{p} \begin{pmatrix} p \\ \dfrac{p-1}{2} \end{pmatrix} [u(k)]^{(p-1)/2} [u^*(k)]^{(p+1)/2} e^{-jwk}. \tag{8}$$

Using the binomial property and the relation observed in (5),

$$u_{1z}^p = \frac{1}{2^{p-1}} \begin{pmatrix} p \\ \dfrac{p+1}{2} \end{pmatrix} [u(t)]^{(p+1)/2} [u^*(k)]^{(p-1)/2}. \tag{9}$$

Finally the first-zone filtered input signal can be found as

$$u_{1z}^p = \frac{1}{2^{p-1}} \begin{pmatrix} p \\ \dfrac{p+1}{2} \end{pmatrix} u(k) |u(k)|^{p-1}. \tag{10}$$

The component $(1/2^{p-1}) \binom{p}{(p+1)/2}$ corresponds to the baseband power series coefficients.

If no bias is present in the input/output signals, the regression matrix for the estimation of the coefficients of the baseband polynomial can be written as

$$\mathbf{U} = \begin{bmatrix} u(1) & u(1)|u(1)|^2 & \cdots & u(1)|u(1)|^{p-1} \\ u(2) & u(2)|u(2)|^2 & \cdots & u(2)|u(2)|^{p-1} \\ \vdots & \vdots & \ddots & \vdots \\ u(N) & u(N)|u(N)|^2 & \cdots & u(N)|u(N)|^{p-1} \end{bmatrix}. \tag{11}$$

The baseband polynomial can be written in a compact form:

$$\phi_p[u(k)] = \sum_{p=1}^{P} c_p |u(k)|^{2(p-1)} u(k). \tag{12}$$

For complex Gaussian baseband input signals, a derivation of orthogonal polynomials is found in [11].

3.1.3. Bessel-Fourier Model. The complex Bessel approximation of a memoryless RF power amplifier is obtained by the periodic extension of the instantaneous voltage transfer characteristics by a complex Fourier series expansion. This derivation was extracted from [12]:

$$y(k) = \sum_{p=-\infty}^{\infty} c_p \cdot e^{j\alpha k u(k)}, \tag{13}$$

where $u(k)$ is the input signal and $y(k)$ the output signal, both with finite dynamic range, and c_p are Fourier series

coefficients. The parameter α is determined by the maximum dynamic range of the input, D,

$$\alpha = \frac{2\pi}{D}, \tag{14}$$

which defines the period of the Fourier series periodic extension. Hence, for the general N-carrier input

$$u(k) = \sum_{l=1}^{N} A_l(k) \cos[\omega_l k + \phi_l(k)], \tag{15}$$

the output is

$$y(k) = \sum_{p=-\infty}^{\infty} c_p \prod_{l=1}^{N} e^{j\alpha p A_l(k) \cos(\omega_l k + \phi_l(k))}. \tag{16}$$

And, employing the Bessel function series approximation,

$$e^{jx \cos\theta} = \sum_{n=-\infty}^{\infty} j^n J_n(x) e^{jn\theta} \tag{17}$$

with $x = \alpha p A_l(k)$ and $\theta = \omega_l k + \phi_l(k)$, it is possible to write (16) as

$$y(k) = \sum_{p=-\infty}^{\infty} c_p \prod_{l=1}^{N} \sum_{n=-\infty}^{\infty} j^n J_n[\alpha p A_l(k)] e^{jn_l[\omega_l k + \phi_l(k)]}. \tag{18}$$

Rearranging some terms,

$$y(k) = \sum_{p=-\infty}^{\infty} jc_p \sum_{n=-\infty}^{\infty} \left\{ (-j) \prod_{l=1}^{N} J_n[\alpha p A_l(k)] j^{n_l} \right\}$$
$$\cdot e^{\sum_{l=1}^{m} jn_l[\omega_l k + \phi_l(k)]}. \tag{19}$$

Defining $b_k = j(c_k - c_{-k})$ for $k = 1 \cdots \infty$ and $\sum_{l=1}^{m} n_l = 1$ (as only the first zonal components will be considered),

$$y(k) = \sum_{k=1}^{\infty} b_k \sum_{n=-\infty}^{\infty} \left\{ \prod_{l=1}^{N} J_{n_l}[\alpha k A_l(k)] \right\}$$
$$\cdot e^{j \sum_{l=1}^{m} [\omega_l k + \phi_l(k)]} \tag{20}$$

and J_{n_l} denotes the n_lth-order Bessel function of the first kind. The coefficients b_k may be obtained by using a LS approximation.

3.1.4. Look-Up Tables. Look-up tables (LUTs) are the most common type of nonlinear static models in real-world implementations [7]. An advantage in comparison with other methods is the configuration possibility of the interpolation and extrapolation behavior. LUTs also present good accuracy and very fast evaluation. The drawbacks are poor physical interpretation, high number of parameters, and being not continuously differentiable. Linear interpolation is normally used to determine the points among intervals, but also other methods as cubic interpolation and splines are possible [13].
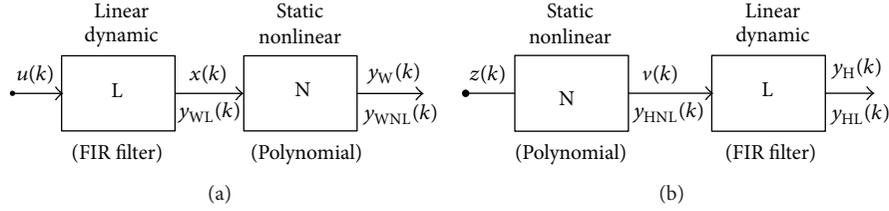
FIGURE 5: Wiener (a) and Hammerstein (b) models.

3.2. Linear Memory.

The two-box modeling techniques are a possibility to represent the linear memory of an amplifier. They are also known as modular approaches [14] or feed-forward block oriented models [1]. They are obtained by combining components from the following two classes: static (or memoryless) nonlinearities and causal, linear time-invariant dynamic subsystems. Parametric and nonparametric modeling methodologies can be used. Flexible arrangements of block structured models in two possibilities are feasible: Wiener model (linear-nonlinear) and Hammerstein model (nonlinear-linear) [15].

The most frequently used configuration for the linear block of this model is a FIR filter. The nonlinear block is commonly represented by a polynomial [1]. Examples of these structures are shown in Figure 5.

If the linear dynamic block is represented by a FIR filter, the output of this block for the Wiener model is

$$y_{\text{WL}}(k) = \sum_{\tau=0}^{M-1} h(\tau) u(k - \tau).$$ (21)

For the Hammerstein model, the FIR filter output is

$$y_{\text{HL}}(k) = \sum_{\tau=0}^{M-1} h(\tau) v(k - \tau).$$ (22)

If the static nonlinearity block is represented by a power series, the output of this block can be formulated for the Wiener Model and for the Hammerstein Model as follows:

$$y_{\text{WNL}}(k) = \sum_{p=0}^{P} c_p x^p(k),$$ (23)

$$y_{\text{HNL}}(k) = \sum_{p=0}^{P} c_p z^p(k).$$ (24)

The overall model output is then the combination of these equations for each model:

$$y_{\text{W}}(k) = \sum_{p=0}^{P} c_p \left(\sum_{\tau=0}^{M-1} h(\tau) u(k - \tau) \right)^p,$$

$$y_{\text{H}}(k) = \sum_{\tau=0}^{M-1} h(\tau) \left(\sum_{p=0}^{P} c_p z^p(k - \tau) \right).$$ (25)

Equations (25) are a simple way to model a nonlinear amplifier with memory.
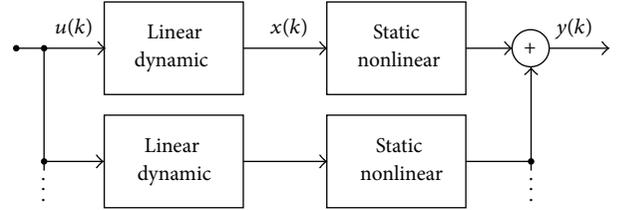

FIGURE 6: Example of a parallel Wiener model.

3.3. Nonlinear Memory.

More complex models are necessary to estimate the nonlinear memory, like parallel models or Volterra series. Examples of these models are parallel cascade models.

Any system that can be represented by a truncated VS (26) can be also modeled exactly using parallel cascaded structures [16].

This technique is the association in branches of various models (Wiener, Hammerstein, Wiener-Hammerstein, etc.). The overall model structure becomes more complicated with each iteration, as each branch is composed of a single model. The value of the cost function decreases or stays constant with each additional branch [17]. An example of this configuration is seen in Figure 6.

This method combines the following favorable properties:

(i) computationally efficient even for high-order models with large memory-bandwidth products,

(ii) allowing the direct extraction of the Volterra kernels,

(iii) offering the convenience to use different methods for the identification of the linear and nonlinear blocks [17].

As a drawback, this method is very sensitive to noise if too many paths are used [14]. Consequently, a proper selection of the paths using parametric FOMs and the system order of the nonlinearity should be made, in order to achieve low noise and good convergence models.

Although the best estimation methods to identify the parallel Wiener model's coefficients would be the nonlinear ones, Korenberg proposed initially linear methods with acceptable results, as described in his paper [17].

3.4. Volterra Series.

Volterra series accounts for a mildly nonlinear class of nonlinear systems and has the property of dynamic interaction of nonlinearities, so it is well suited for the description of PAs.

The finite, discrete VS model is given by [18]

$$
\begin{aligned}
y_V(k) = \sum_{p=0}^{P} \sum_{\tau_1=0}^{M-1} \cdots \sum_{\tau_P=0}^{M-1} h_p\left(\tau_1, \ldots, \tau_p\right) \\
\cdot u\left(k - \tau_1\right) \cdots u\left(k - \tau_p\right),
\end{aligned}
\tag{26}
$$

where h_p is the kernel of order P, k and τ are discrete indices of the sampling interval, and M is the memory length. The sampling interval must be selected to cover the needed input/output signal's bandwidth.

The main disadvantage of a VS based BM is the number of parameters necessary to estimate and consequently to represent the model. A VS model using 5 delay taps needs 5, 125, 625, and 3125 parameters for the 1st, 3rd, 5th, and 7th order kernels, respectively. These values are not practical, since an estimation using so many coefficients is very computational intensive, even for actual computers.

By using the symmetry condition, the complexity of the Volterra kernels as a function of the order of nonlinearity is given by the binomial [19]:

$$
\binom{M + p - 1}{p},
\tag{27}
$$

where M is the number of delay taps used and p is the order of the kernel. Using (27), the above cited model is reduced to 5, 45, 126, and 330 parameters. Unfortunately, this equation is valid only for real valued signals.

3.4.1. Complex Valued Baseband Volterra Series. In order to obtain the best model performance, it is necessary to adapt the BMs under study to the modern PA input/output industry standard signals, once these models are designed for linearization purposes. The excitation signals are complex valued, and as a practical issue only first-zone filtered (baseband) equivalent BMs are frequently used, due to the difficulties to implement bandpass models in hardware.

A closed form for determining the number of independent terms for baseband VS using complex signals is the binomial [20]:

$$
\binom{M + \left\lfloor \dfrac{p}{2} \right\rfloor}{\left\lfloor \dfrac{p}{2} \right\rfloor} \binom{M + \left\lfloor \dfrac{p}{2} \right\rfloor}{\left\lfloor \dfrac{p}{2} \right\rfloor + 1} \frac{1}{M + \lfloor p/2 \rfloor},
\tag{28}
$$

where $\lfloor \cdot \rfloor$ is the floor operation.

As an example, the numbers of parameters of a complex valued baseband VS using 4 delay taps are 4, 40, 200, and 700 for the 1st, 3rd, 5th, and 7th order symmetric kernels. The use of the Volterra kernels symmetry property is necessary in the model extraction process, since it eliminates the linear dependent columns of the kernel to be estimated.

Several techniques are employed to estimate VS. If the system is memoryless, VS are reduced to a Taylor series and can be estimated as described in Section 3.1. If the system has only linear memory, it can be estimated using the techniques
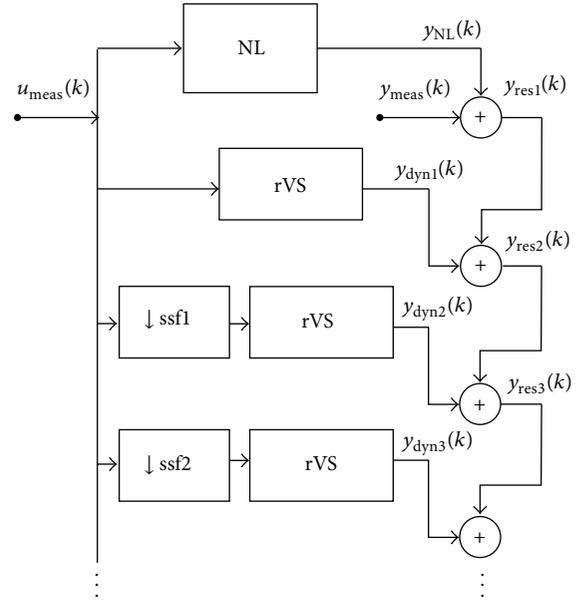

Figure 7: A PA behavioral model representation using multirate and reduced Volterra series in a cascade configuration.

listed in Section 3.2. If the system presents only nonlinear memory or linear and nonlinear memory, some strategies described in Section 3.3 can be employed.

As shown above, VS presents a very high number of coefficients. This can lead to ill-conditioned Hessian matrices, as shown in [20]. The best way to estimate a BM using VS is to prune some terms, losing the minimum accuracy as possible. This simplification includes the use of only main diagonals terms of the VS kernels, the popular parallel Hammerstein models (PH) [21].

3.4.2. Pruned Volterra Series. There are other models that take into account some physical knowledge of the device and include important interactions of the input signal, as in [4, 22, 23]. These are the pruned or reduced VS models, because they include interactions different than the terms in the main diagonal. The accuracy is naturally higher than the PH models, but they also use more coefficients in the model.

3.4.3. Cascade Multirate Pruned Volterra Series. Joining pruning techniques of VS, multirate techniques, and also parallelization of models (cascade), a new model was developed. This model presented different models composing the cascade. The first branch is a nonlinear static block, and the next branches are reduced Volterra series, each one estimated at a different subsampling rate (PssVS). This model and its estimation method are fully explained in [5], and it is shown in Figure 7.

Based on Figure 7 and considering the input signal as $u_{\text{meas}}(k)$, the output signal $y_{\text{NL}}(k)$ is the output of an estimated baseband polynomial (or a look-up table) as in (12). The first residue is

$$
y_{\text{res1}}(k) = y_{\text{meas}}(k) - y_{\text{NL}}(k).
\tag{29}
$$

The second output signal, $y_{\mathrm{dyn1}}(k)$, is the output of an estimated reduced Volterra series BM. Its input signal is $u_{\mathrm{meas}}(k)$, and its output signal is the first residue, $y_{\mathrm{res1}}(k)$. Then the second residue is

$$y_{\mathrm{res2}}(k) = y_{\mathrm{res1}}(k) - y_{\mathrm{dyn1}}(k). \tag{30}$$

So the ith residue for $i \geq 2$ is

$$y_{\mathrm{res}(i)}(k) = y_{\mathrm{res}(i-1)}(k) - y_{\mathrm{dyn}(i-1)}(k). \tag{31}$$

The overall output of the PssVS model is described as

$$y_{\mathrm{tot}}(k) = y_{\mathrm{NL}}(k) + y_{\mathrm{dyn1}}(k) + y_{\mathrm{dyn2}}(k) + y_{\mathrm{dyn3}}(k) + \cdots. \tag{32}$$

This equation is also described as in (33). This equation is linear in parameters and its coefficients can be found using linear regression methods such as least-squares (LS) [6]:

$$
\begin{aligned}
y(k) = & \sum_{p=1}^{P} c_p \left| u(k) \right|^{2(p-1)} u(k) \\
& + \sum_{b=1}^{B} \sum_{\tau_1=0}^{M-1} \cdots \sum_{\tau_p=0}^{M-1} hr_{p(b)} \left[\left(rf_{(b)}\tau_{1(b)} \right), \ldots, \left(rf_{(b)}\tau_{p(b)} \right) \right] \\
& \cdot \prod_{j=1}^{p} u \left(k - rf_{(b)}\tau_{j(b)} \right).
\end{aligned}
\tag{33}
$$

The excitation signals are complex valued, and thus a first-zone filtered odd-order model was designed. This approach is expressed by the polynomial basis function showing only odd terms [24]. The input vector is $u(k)$, B is the number of branches used in the model, P is the NL order of the model, and M is the memory depth. A higher order can be used in the static nonlinear branch, as it can also be replaced by any other static nonlinear block. The advantage of this model is that it can account for different memory effects existing in a PA also at different sampling rates. Applying the input signal at different rates emphasizes the flexibility of the model.

A comparison of this model with other models here cited using simulated and measured data obtained from a LDMOS RF PA is firstly presented in [5], using only one FOM (NMSE). Further on, a comparison was presented in [6] using five different FOMs, namely, normalized mean square error (NMSE) [21], normalized root mean square error (NRMSE) [25], mean absolute error (MAE) [26], maximal absolute error (MaxAE) [27], and coefficient of efficiency (e) [28]. Final results show a superior performance of the PssVS.

These results show a trend in LDMOS RF PA behavioral models that can not be override. Multirate models allow simpler hardware to be used in linearization devices that are the end products of BM. They also present a higher accuracy than any other model reported so far for LDMOS RF PAs. Further efforts in this research direction can reveal even more accurate models that use fewer coefficients with simpler hardware.

4. Conclusion

This paper presented a classification of nonlinear systems and also a modern classification of behavioral models of RF power amplifiers. Then an evolution of behavioral models was also presented, including all equations that characterize these models. This review showed several models, from basic baseband power series to the recent parallel multirate pruned Volterra series models, commenting also on their accuracy. Remarks about future trends in power amplifiers behavioral models were also made.

Conflict of Interests

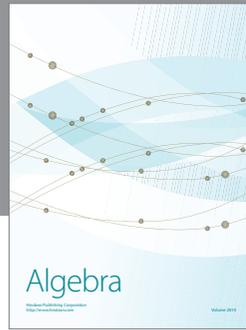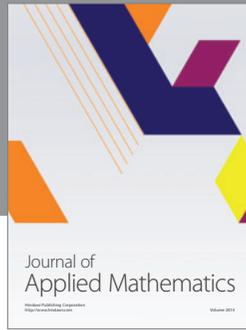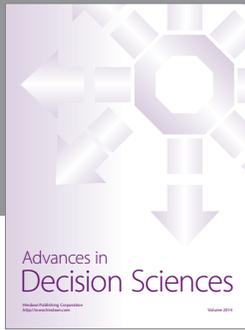The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] R. K. Pearson, "Selecting nonlinear model structures for computer control," *Journal of Process Control*, vol. 13, no. 1, pp. 1–26, 2003.

[2] J. C. Pedro and S. A. Maas, "A comparative overview of microwave and wireless power-amplifier behavioral modeling approaches," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 4, pp. 1150–1163, 2005.

[3] A. Zhu and T. J. Brazil, "RF power amplifier behavioral modeling using volterra expansion with laguerre functions," in *Proceedings of the IEEE MTT-S International Microwave Symposium*, pp. 963–966, June 2005.

[4] J. Reina-Tosina, C. Crespo-Cadenas, and M. J. Madero-Ayora, "A compact volterra model for power amplifiers with memory," in *Proceedings of the IEEE MTT-S International Microwave Symposium (IMS '09)*, pp. 1585–1588, IEEE, Boston, Mass, USA, June 2009.

[5] D. D. Silveira, P. L. Gilabert, P. M. Lavrador et al., "Improvements and analysis of nonlinear parallel behavioral models," *International Journal of RF and Microwave Computer-Aided Engineering*, vol. 19, no. 5, pp. 615–626, 2009.

[6] D. D. Silveira, P. L. Gilabert, A. B. dos Santos, and M. Gadringer, "Analysis of variations of volterra series models for RF power amplifiers," *IEEE Microwave and Wireless Components Letters*, vol. 23, no. 8, pp. 442–444, 2013.

[7] O. Nelles, *Nonlinear System Identification*, Springer, Berlin, Germany, 2001.

[8] D. Lichtblau and E. W. Weisstein, "Condition Number," MathWorld—A Wolfram Web Resource, December 2004, http://mathworld.wolfram.com/ConditionNumber.html.

[9] J. Vuolevi and T. Rahkonen, *Distortion in RF Power Amplifiers*, Artech House, Norwood, Mass, USA, 2003.

[10] M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Simulation of Communication Systems: Modeling, Methodology, and Techniques*, Kluwer Academic Publishers, Norwell, Mass, USA, 2000.

[11] R. Raich and G. T. Zhou, "Orthogonal polynomial for complex Gaussian processes," *IEEE Transactions on Signal Processing*, vol. 52, no. 10, pp. 2788–2797, 2004.

[12] M. S. O'Droma, "Dynamic range and other fundamentals of the complex Bessel function series approximation model for memoryless nonlinear devices.," *IEEE Transactions on Communications*, vol. 37, no. 4, pp. 397–398, 1989.

[13] Mathworks, "Matlab version 7," 2005.

[14] V. Z. Marmarelis, *Nonlinear Dynamic Modeling of Physiological Systems*, John Wiley & Sons, 2004.

[15] W. J. Rugh, *Nonlinear System Theory—The Volterra/Wiener Approach*, Johns Hopkins University Press, Baltimore, Md, USA, 1981.

[16] T. M. Panicker and V. John Mathews, "Parallel-cascade realizations and approximations of truncated Volterra systems," *IEEE Transactions on Signal Processing*, vol. 46, no. 10, pp. 2829–2832, 1998.

[17] M. J. Korenberg, "Parallel cascade identification and Kernel estimation for nonlinear systems," *Annals of Biomedical Engineering*, vol. 19, no. 4, pp. 429–455, 1991.

[18] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*, Krieger Publishing, 1980.

[19] V. J. Mathews and G. L. Sicuranza, *Polynomial Signal Processing*, Wiley Interscience, 2000.

[20] D. D. Silveira and G. Magerl, "Extraction and improvements of a behavioral model based on the wiener-bose structure used for baseband volterra kernels estimation," in *Proceedings of the IEEE MTT-S International Microwave Symposium (IMS '07)*, vol. 1, pp. 2007–2010, Honolulu, Hawaii, USA, June 2007.

[21] A. S. Tehrani, H. Cao, S. Afsardoost, T. Eriksson, M. Isaksson, and C. Fager, "A comparative analysis of the complexity/accuracy tradeoff in power amplifier behavioral models," *IEEE Transactions on Microwave Theory and Techniques*, vol. 58, no. 6, pp. 1510–1520, 2010.

[22] A. Zhu, J. C. Pedro, and T. R. Cunha, "Pruning the volterra series for behavioral modeling of power amplifiers using physical knowledge," *IEEE Transactions on Microwave Theory and Techniques*, vol. 55, no. 5, pp. 813–820, 2007.

[23] C. Crespo-Cadenas, J. Reina-Tosina, M. J. Madero-Ayora, and J. Munoz-Cruzado, "A new approach to pruning Volterra models for power amplifiers," *IEEE Transactions on Signal Processing*, vol. 58, no. 4, pp. 2113–2120, 2010.

[24] S. Benedetto and E. Biglieri, *Principles of Digital Transmission*, Kluwer Academic Publishers, Norwell, Mass, USA, 1999.

[25] A. Zhu, P. J. Draxler, J. J. Yan, T. J. Brazil, D. F. Kimball, and P. M. Asbeck, "Open-loop digital predistorter for RF power amplifiers using dynamic deviation reduction-based volterra series," *IEEE Transactions on Microwave Theory and Techniques*, vol. 56, no. 7, pp. 1524–1534, 2008.

[26] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Climate Research*, vol. 30, no. 1, pp. 79–82, 2005.

[27] P. Karras and N. Mamoulis, "One-pass wavelet synopses for maximum-error metrics," in *Proceedings of the 31st International Conference on Very Large Data Bases (VLDB '05)*, pp. 421–432, September 2005.

[28] M. J. Hall, "How well does your model fit the data?" *Journal of Hydroinformatics*, vol. 3, pp. 49–55, 2001.