Behavioral Modeling of Radio Frequency Power Amplifiers Using a Multiple Depth Memory Volterra Series

Felipe P. Ribeiro
Group of Integrated Circuits and Systems (GICS)
Federal University of Parana
Curitiba, Brazil
felipe.pires@ufpr.br

Eduardo G. Lima
Group of Integrated Circuits and Systems (GICS)
Federal University of Parana
Curitiba, Brazil
eduardo.lima@ufpr.br

***Abstract*—Modeling is an essential step for designing power amplifier linearizers, which enable operation in higher power delivery regions. Black-box modeling had become an attractive method on PA behavioral modeling as it does not require prior knowledge of the physical equations that govern the nonlinear behavior of the device. In this study, the investigation of a black box model based on the Volterra series and modified using multiple-depth memories is presented. For model validation a case study was executed with a class AB PA with GaN based technology, excited by a 900 MHz carrier and stimulated by a WCDMA 3GPP signal with 3.84 MHz of bandwidth, where a total of 280 cases were performed. The proposed model is promising, being able to reduce the normalized mean square error by up to 7.244 dB and computational complexity by up to 81%. The best NMSE result achieved without overfitting was -46.074 dB for the proposed model, against -38.830 dB for the low-pass equivalent Volterra series.**

***Index Terms*—PA, amplifier, behavioral modeling, Volterra, moving average**

I. Introduction

The power amplifier (PA) [1] plays a crucial role in wireless communication systems as it consumes the majority of the energy. Its function is simple: increase the power of the signal to be transmitted without applying any distortion. However, due to physical limitations of the semiconductors used in amplifiers (such as BJTs, FETs, JFETs, MOSFETs, etc.) and the chosen operation mode (class A, B, AB, etc.), the PA can operate in regions where its transistors become saturated, resulting in gain compression. This saturation leads to the generation of harmonic frequencies and spectrum spreading, causing distortions in the transmitted signal. Additionally, the capacitive and inductive effects from the input/output matching networks, as well as thermal and intrinsic effects of semiconductors, can introduce delays in the amplified signal, known as memory effects. Considering modern quality standards, these distortions cannot be neglected.

One widely used method for PA linearization is Digital Pre-Distortion (DPD) [2]. This method involves applying the inverse signal of the PA prior to transmission, resulting in a linear relationship between the DPD input and the PA output. To implement DPD effectively, a behavioral model of the PA with high accuracy and low computational complexity is required.

Among the available models for PA behavioral modeling, two commonly employed approaches are polynomial models, such as the Volterra series and its derivatives and simplifications (e.g., memory polynomials (MP) [2] and generalized memory polynomials (GMP) [2]), and artificial neural networks like multi-layer perceptrons (MLP) [2] and time delay neural networks (TDNN) [2]. In previous investigations [3], [4], the cascades among two and three Volterra series were investigated and better performance was achieved in relation to the traditional model.

The objective of this study was to compare the results of the traditional low-pass equivalent Volterra model and the better result of the cascade model with the results of a modified Volterra model, named Multi-Depth Memories (MDM) Volterra Series model, that takes into account several previous instants, represented by moving averages that have different windows (multi-depth) used as instantaneous memories that are applied on the traditional low-pass equivalent Volterra series. Numerical experiments were conducted in Python to validate the effectiveness of the proposed model.

Moving averages have been already employed in PA behavioral models to separate the PA static and dynamic behaviors for Box-Oriented Models by averaging the dispersion of the signal [5], [6]. However, in this article, moving averages with distinct window sizes are being employed to take into account the usage of the amplifier in multiple time-scales, using a single nonlinear-dynamic model for the PA modeling.

II. Behavioral modeling described by the MDM Volterra series

Behavioral modeling is the process of describing the behavior of a device, either through physical or mathematical representations. Physical models require a comprehensive understanding of the object being modeled, including its internal components, resulting in high computational complexity. Therefore, when it comes to modeling power amplifiers, the use of physical models is avoided due to the preference for

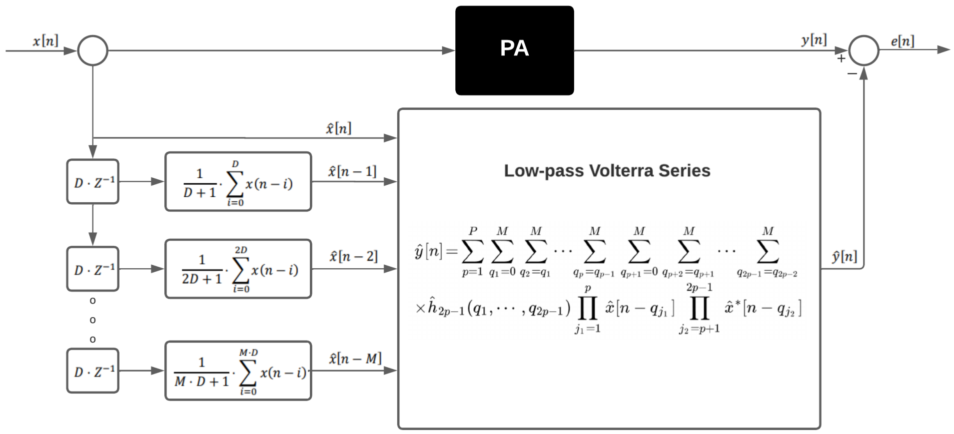

Fig. 1. Proposed Volterra series model with memories based on moving averages of the sampled signal

low-complexity models. Mathematical modeling of the power amplifier involves abstracting the actual process, treating the amplifier as a "black box" with only the input and output information available. With this limited data and a selected model, the coefficients for the behavioral model can be calculated.

In this research, the employed model is based on the Volterra series, which is a polynomial series known for its ability to reproduce memory effects. Unlike models that rely solely on the instantaneous input, the Volterra series takes into account the previous M instants (for discrete-time models). This characteristic enables a non-linear relationship between the output and input, effectively capturing the non-linearity and memory effects of the power amplifier. Furthermore, the series maintains linearity in its parameters without sacrificing generality. For the purpose of modeling radio frequency power amplifiers (RFPA), it is convenient to adopt the discrete-time low-pass equivalent Volterra series model [7], described by:

$$\hat{y}(n) = \sum_{p=1}^{P}\sum_{q_1=0}^{M}\sum_{q_2=q_1}^{M}\cdots\sum_{q_p=q_{p-1}}^{M}\sum_{q_{p+1}=0}^{M}\sum_{q_{p+2}=q_{p+1}}^{M}\cdots\sum_{q_{2p-1}=q_{2p-2}}^{M} \hat{h}_{2p-1}(q_1,\cdots,q_{2p-1})\prod_{j_1=1}^{p}\hat{x}(n-q_{j_1})\prod_{j_2=p+1}^{2p-1}\hat{x}^*(n-q_{j_2}) \quad (1)$$

where $\hat{x}(n)$ and $\hat{y}(n)$ are the complex-valued envelopes, respectively, at the input and output of the PA, $(^*)$ denotes the complex conjugate operator, M is the memory length, $P_0 = 2P - 1$ is the polynomial order truncation and $\hat{h}_{2p-1}(q_1,\cdots,q_{2p-1})$ are the low-pass equivalent Volterra kernels. The number of parameters to be determined in (1) is given by

$$L = \sum_{p=0}^{P-1}\frac{[(M+p)!]^2(M+p+1)}{(M!)^2(p!)^2(p+1)} \quad (2)$$

As the number of coefficients in (2) increases quickly with P and M, we are going to investigate another modeling, looking for complexity reduction. In this particular case, instead of directly using the previous instants of the input signal in the Volterra series, we employ moving averages. A moving average is a mathematical operation that calculates the average of a set of values over a sliding window. By using moving averages as the previous instants, we can effectively capture the temporal characteristics and memory effects of the system.

Instead of using a fixed window size for all the moving averages, we employ varying window sizes. By utilizing different window sizes as the previous instants in the Volterra series, we can capture the system behavior at various time scales. For example, if we have a small window size, it will capture short-term variations in the input signal. This enables us to model and understand the system fast-changing dynamics. On the other hand, a larger window size captures longer-term trends and variations in the input signal, providing insights into the system slower dynamics and memory effects over a more extended period.

By incorporating moving averages with different window sizes into the Volterra series, we obtain a more comprehensive representation of the system behavior. This approach allows us to account for both short-term and long-term memory effects, nonlinearities, and temporal characteristics of the power amplifier accurately.

The proposed model, represented by Fig. 1 in a simplified block diagram, presents the modified Volterra series in parallel with the PA to be modeled, where $D \cdot Z^{-1}$ represents a unit delay block applied D times. In this approach, M is the number of moving averages to be employed, $M \cdot D$ is the memory length of the model, $2P - 1$ is the polynomial order truncation, $\hat{y}[n]$ is the estimated output and $e[n]$ is the error between the measured and the estimated output of the PA. As the parameters of the series, M and P, remain the same, the number of coefficients of the series is not changed.

Since the model is linear in its coefficients, we can extract the coefficients by applying the least squares method. In Python, the coefficient extraction process involves using the output vector Y with dimensions $(N - M \cdot D) \times 1$, the matrix X with dimensions $(N - M \cdot D) \times L$, which corresponds to the products in (1), and the desired coefficient vector H that relates X and Y:

$$Y = X \cdot H \rightarrow H = numpy.linalg.pinv(X)@Y \quad (3)$$

III. Case Study for Model Validation

In this research the proposed model was applied in a case study with a class AB PA with GaN based technology, excited by a 900 MHz carrier and stimulated by a WCDMA 3GPP signal with 3.84 MHz of bandwidth. The signals were measured using a vector signal analyzer (VSA) with a sampling frequency of 61.44 MHz. The data was previously collected and detailed in [8]. For performance analysis, the normalized mean square error (NMSE), was calculated for all computed cases as:

$$NMSE = 10 log \frac{\sum_{n=1}^{N} (\hat{y}[n] - y[n+M-D])^2}{\sum_{n=1}^{N} y[n+M-D]^2} [\text{dB}] \quad (4)$$

where $y[\cdot]$ is the measured output validation data and $\hat{y}[\cdot]$ is the estimated output signal, both complex-valued vectors with dimension $N + M - D$ and N respectively.

For the validation of the method, the model was trained with several combinations of the parameters, that satisfy

$$[1, 0, 1] \leq [P, M, D] \leq [6, 6, 10] \text{ and } L \leq 1000 \quad (5)$$

A total of 280 cases were conducted and subsequently ranked based on decreasing NMSE. Figure 2 illustrates the NMSE curves of the training dataset for both the proposed and traditional methods, along with their respective validation curves. In this graph, if a result exhibits more coefficients and a higher NMSE compared to another result with fewer coefficients on the same curve, it is excluded from the plot. The regions surrounding the curves represent a deviation limit of 3% from the NMSE curve. In this study, a deviation exceeding 3% was considered indicative of overfitting.

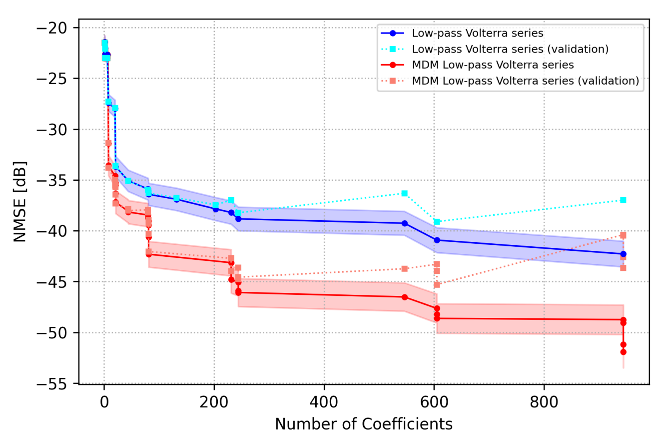

Fig. 2. NMSE curves for the traditional (blue) and proposed (red) models with a 3% trust region

As can be observed from Fig. 2, models with more than 250 coefficients exhibit signs of overfitting. In consequence, only models with fewer than 250 coefficients were considered. Figure 3 presents a 3D visualization of the NMSE curve for the proposed method with the accounted points.

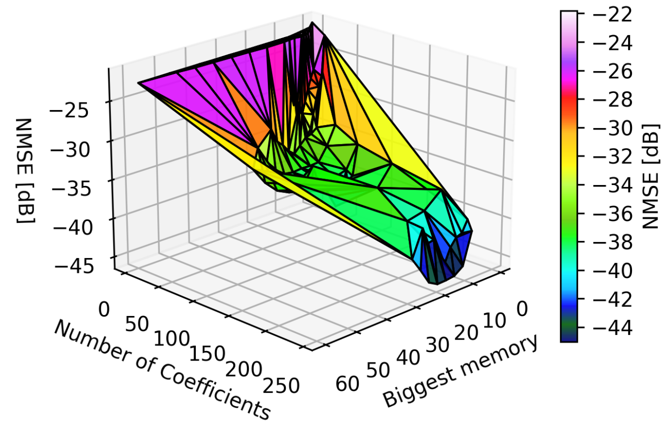

Fig. 3. 3D visualization of the NMSE curve for the proposed method as a function of the number of coefficients and the memory length.

Table I presents a comparative analysis among the best results achieved by three different models: the traditional Low-pass Volterra Series, the Three-blocks Cascade Volterra Series as explored in [4], and the proposed Multi-Depth Memories Volterra Series. Additionally, the initial results that surpassed the traditional model are also included for the last two models.

TABLE I
Comparative Analysis of the Models

M_1	P_1	D	M_2	P_2	M_2	P_3	L	NMSE [dB] Extraction	NMSE [dB] Validation
3	3	-	-	-	-	-	244	-38.830	-38.233
4	3	-	-	-	-	-	944	-42.280	-36.987
0	2	-	1	1	5	2	136	-38.921	-38.571
0	2	-	6	2	1	4	245	-41.359	-40.595
3	2	6	-	-	-	-	81	-42.308	-42.042
3	3	6	-	-	-	-	244	-46.074	-44.566
4	3	6	-	-	-	-	944	-51.921	-43.650

In Fig. 4, samples of the characteristic curves AM-AM and AM-PM are presented. These curves depict the relationship between the amplitude of the input signal and the corresponding output signal, as well as the amplitude of the input signal and the phase shift between the input and output signals. The curves represent the best results obtained from both the traditional model and the proposed model when applied to a validation dataset.

Finally, Fig. 5 presents the spectrum of the input and output signals applied to the PA, as well as the error spectrum for both the traditional and Multi-Depth Memories Volterra series. Additionally, Fig. 6 depicts the error accumulation region using heatmaps for both methodologies.

IV. Conclusions

Through the analysis of the results, we are able to observe the best approaches given by the Multi-Depth Memories Volterra series model. The computational complexity was reduced by up to 81%, when the average between the NMSE of extraction and validation for the proposed model with 44 coefficients had a greater result than the traditional model with 231 coefficients (against 51% of reduction obtained by