

Behavioral Modeling the Power Amplifier with Memory Effect Using the NARMA model

Jiajia Song¹⁺, Yuelin MA² and Lili Guo¹

¹Harbin Engineering University

Harbin, China

²University of Electronic Science and Technology of China

Chengdu, China

Abstract—Behavioral modeling is an important approach to access the characteristic of the power amplifier (PA) and to perform the numerical simulation. Polynomial model is the most popular tool to behaviorally model a PA which is always been considered to be a weak nonlinear system. When wideband signal applied, the frequency-dependence, namely, memory effect is inevitable. The polynomial always becomes very complicated to model the PA with deep memory effect. Nonlinear auto-regressive moving average (NARMA), which can be seen as a recursive polynomial, has few implementations in power amplifier behavioral modeling. As will be shown hereinafter, NARMA model demonstrates low complexity and fairly good accuracy to model an actual PA. To develop a robust and numerically stable identification algorithm, orthogonal polynomial is used to improve the numerical stability of the matrix inverse.

Keywords-Behavior modeling; NARMA; power amplifier model; orthogonal polynomial;

1. Introduction

Wireless communication systems are evolving to provide high data rate transmission and make full use of spectral resource. To enhance the spectral efficiency, much of the importance is attached to development of the compact constellations and multiple access techniques. Such techniques, however, result in tougher requirement to the radio frequency (RF) power amplifier (PA) which plays a key role in wireless communication system, so it has been the subject of intensive studies intended to access its characteristic and to understand its behavior. In recent years, computer aided design (CAD) software is widely used for studying the circuit and system, which significantly reduces the lifespan of system designing and the cost of experiment. As an efficient approach to reproduce the output of the PA in computer simulation, behavior model attracts lot of attentions.

Volterra Series is known to be the most general polynomial to model a weak nonlinear system with memory [1]. It, however, is much too complicated to be used to model a power amplifier. Many researchers focus on how to truncate the Volterra Series while ensuring the fidelity of the model. The most popular behavioral model may be the memory polynomial [2] [3] in which only the diagonal elements of the Volterra Series are considered. To improve the accuracy of modeling a PA with deep memory effect, a so-called generalized memory polynomial is proposed in [4] therein cross terms of time delay are introduced. Some authors also simplify the Volterra Series by regrouping its terms [5] and reduce its dynamics.

⁺ Corresponding author.
E-mail address: sonyajiajia@163.com

In this paper, we consider the NARMA model [6] which is common in nonlinear system theory but has very few practical implementations for PA modeling. To perform the modeling in baseband, we propose a baseband NARMA model which has only standard structures and it is suitable for digital signal processing. To improve the numerical stability of the data processing, orthogonal polynomial is used to effectively reduce the condition number of the matrix. It is shown that NARMA model has low complexity and fairly good performance.

This paper is organized as follows: we begin with the introduction of baseband NARMA model and emphasize the difference from the passband counterpart. In section III, the algorithm used to identify the power amplifier with the NARMA model is presented. The numerical stability problem is also stated in this section and a remedial technique is introduced. It is followed by an experiment to examine the baseband NARMA model. Finally, a conclusion is drawn in the last section.

2. Baseband NARMA Model

Behavior model is to describe a PA by a system operator that maps a function $x(t)$ onto another function $y(t)$. The input-output relation can be expressed by a nonlinear differential function [7]

$$f\left[y(t), \frac{dy(t)}{dt}, \dots, \frac{d^p y(t)}{dt^p}, x(t), \frac{dx(t)}{dt}, \dots, \frac{d^q x(t)}{dt^q}\right] = 0 \quad (1)$$

which states that the output and its derivatives are nonlinearly related to the input and its derivatives. Since the signal is always processed in baseband digital domain, such

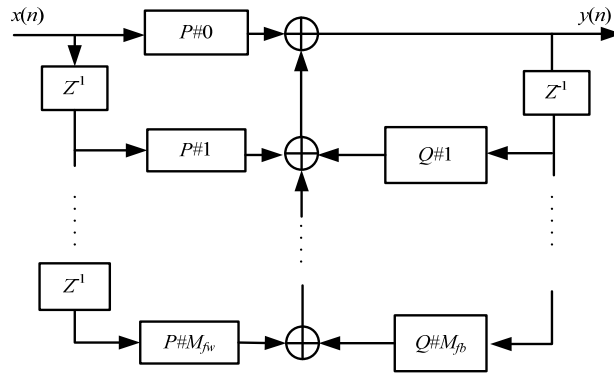


Fig.1. Structure of the NARMA model.

function needs to be modified to baseband discrete version

$$y(n) = f_r[y(n-1), \dots, y(n-p), x(n), x(n-1), \dots, x(n-q)] \quad (2)$$

which indicates that this model is a nonlinear infinite impulse response (IIR) –like system. Under a wide range of conditions, such as causality, stability, fading memory, such system can be approximated by, with any desired error, nonrecursive (finite impulse response (FIR) –like) system [7]

$$y(n) = f_b[x(n), x(n-1), \dots, x(n-M)] \quad (3)$$

where M is the memory length. The reason why recursive model is not so popular as nonrecursive model is that the model in (2) contains feedback paths so that special attention has to be paid to ensure it stable.

The proposed baseband NARMA model can be written as, if we assume the function in (2) to be the power series,

$$\begin{aligned}
y(n) &= \sum_{i=1}^{N_{f_w}} \sum_{j=0}^{M_{f_w}-1} a_{ji} |x(n-j)|^{i-1} x(n-j) \\
&+ \sum_{i=1}^{N_{f_b}} \sum_{j=0}^{M_{f_b}-1} b_{ji} |y(n-j-1)|^{i-1} y(n-j-1)
\end{aligned} \tag{4}$$

where N_{f_w} and N_{f_b} denote the maximum nonlinear order of the input and output, respectively. M_{f_w} and M_{f_b} are the maximum memory length. a_{ji} and b_{ji} are the coefficients. It is worth to mention that even-order terms are introduced in this model though only odd-order terms generate distortion in the frequency band of interest. The description of the advantage of introducing such terms in baseband polynomial is presented in [3] [4].

As mentioned before, (4) is an IIR-like structure, so it contains both forward paths and feedback paths. To clarify this structure we group each term of (4) as $P\#\text{No.}$ or $Q\#\text{No.}$ in (5). Demonstration of such structure is plotted in Figure 1.

The advantage of NARMA model is the introduction of feedback paths which produces infinite order of nonlinearity and infinite memory length. Therefore, the number of delayed samples and the number of nonlinear terms are relaxed so that complexity is significantly reduced.

$$\begin{aligned}
y(n) &= \underbrace{\sum_{i=1}^{N_{f_w}} a_{0i} |x(n)|^{i-1} x(n)}_{P\#0} + \underbrace{\sum_{i=1}^{N_{f_w}} a_{1i} |x(n-1)|^{i-1} x(n-1)}_{P\#1} + \dots + \\
&\underbrace{\sum_{i=1}^{N_{f_w}} a_{M_{f_w}i} |x(n-M_{f_w})|^{i-1} x(n-M_{f_w})}_{P\#M_{f_w}} + \underbrace{\sum_{i=1}^{N_{f_b}} b_{1i} |y(n-1)|^{i-1} y(n-1)}_{Q\#1} + \\
&\dots + \underbrace{\sum_{i=1}^{N_{f_b}} b_{M_{f_b}i} |y(n-M_{f_b})|^{i-1} y(n-M_{f_b})}_{Q\#M_{f_b}}
\end{aligned} \tag{5}$$

3. Power Amplifier Modeling Using NARMA Model

3.1. The Least Square Algorithm

From the point of view of system identification, the coefficients in (4) can be considered generalized impulse response so that a parameter vector W containing all the unknown coefficients a_{ji} and b_{ji} can be formed, and we can define polynomial basis

$$\begin{aligned}
V &= [X_0, |X_0|X_0, \dots, |X_0|^{N_{f_w}-1} X_0, X_1, \dots, |X_1|^{N_{f_w}-1} X_1, \dots \\
&|X_{M_{f_w}}|^{N_{f_w}-1} X_{M_{f_w}}, \dots, Y_1, |Y_1|Y_1, \dots, |Y_1|^{N_{f_b}-1} Y_1, \dots \\
&|Y_{M_{f_b}}|^{N_{f_b}-1} Y_{M_{f_b}}]
\end{aligned} \tag{6}$$

where $X_m = [x(m+1), \dots, x(m+L)]^T$ and $Y_m = [y(m+1), \dots, y(m+L)]^T$ are the observation of the input and output delayed by m samples, respectively. Herein L is the length of available input data and the superscript $(\cdot)^T$ denotes matrix transpose. The model appearing in (4) can be expressed as

$$Y_0 = VW \tag{7}$$

and W is the coefficient vector defined as

$$\begin{aligned}
W &= [a_{00}, a_{01}, \dots, a_{0N_{f_w}}, a_{11}, \dots, a_{M_{f_w}N_{f_w}}, b_{00}, b_{01}, \dots \\
&b_{0N_{f_b}}, b_{11}, \dots, b_{M_{f_b}N_{f_b}}]^T
\end{aligned} \tag{8}$$

We define $E = Y - Y_0 = [e(1), \dots, e(L)]^T$ the error vector, where $Y = [y_{PA}(1), \dots, y_{PA}(L)]^T$ is the desired output (the actual measured output of the PA). The least square (LS) method can estimate W which minimizes the model error criterion shown in (9)

$$J(W) = E^H E \quad (9)$$

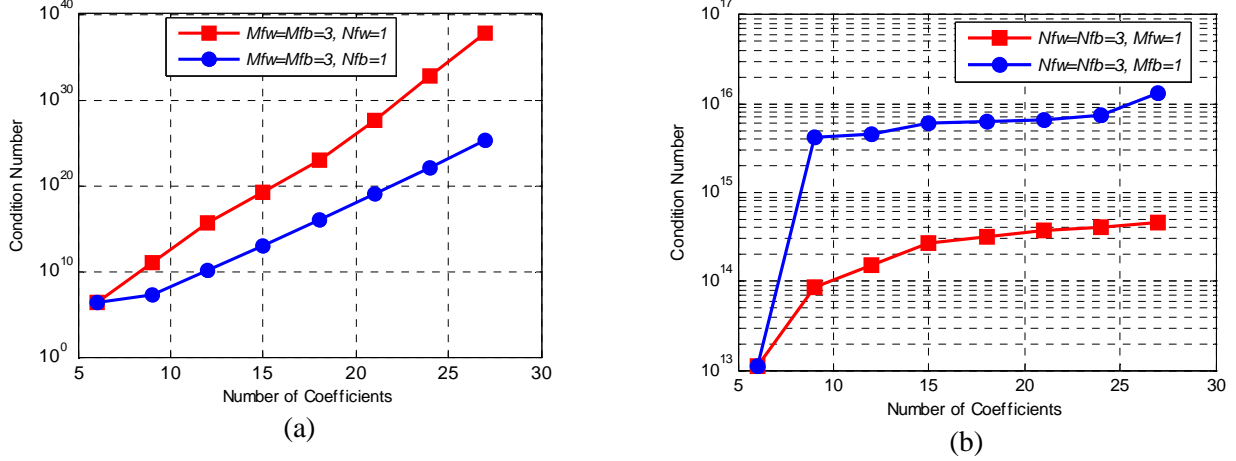


Fig.2. Condition number of $V^H V$: (a) sweeping N_{fw} and N_{fb} , respectively when M_{fw} and M_{fb} are fixed. (b) sweeping M_{fw} and M_{fb} , respectively when N_{fw} and N_{fb} are fixed

where the superscript $(\cdot)^H$ represents the Hermitian transpose. The least square solution is given as

$$W = (V^H V)^{-1} V^H Y \quad (10)$$

The $(V^H V)^{-1} V^H$ in (10) is also known to be the pseudo-inverse of matrix V [8]. To obtain a reliable result, column of V (the data length L) is always set to be from several hundreds to several thousands.

3.2. Numerical Stability of the matrix inverse

As shown in (10), LS algorithm involves matrix inverse, i.e. $(V^H V)^{-1}$, which is to invert a square matrix of size $L \times L$. A metric quantizing the numerical stability of the matrix inverse is the condition number which is defined as the ratio of maximum and minimum eigenvalue amplitudes. Condition number is used to evaluate the ‘‘condition’’, i.e. the data dispersion of a matrix. If the maximum data and minimum data has a large ratio, condition number is big, then the matrix is said to be ill-conditioned. For example, if a matrix has condition number of 10^{10} , the computer is required to be of at least 10^{10} , or 32 bits, to ensure that the computation is error free. On the other hand, having determinant ± 1 and all eigenvalues of magnitude 1 is of great benefit for numerical stability. One implication is that the condition number is 1 (which is the minimum), so that errors are not magnified in the inverse procedure. In practical implementation, it is always impossible to make the condition number to be 1, we can, however, keep it as low as possible to obtain numerical robust computation.

To demonstrate the numerical problem that may be encountered when modeling with the NARMA model, the condition number of $V^H V$ is plotted versus the total number of coefficients in Figure 2, where the signal is assumed to be complex Gaussian distribution, and the length of the signal (L) is set to be 3000. In Figure 2(a), N_{fw} and N_{fb} are varied respectively, while in Figure 2(b), M_{fw} and M_{fb} are swept respectively. Here, the total number of coefficients is defined as $N_{fb} N_{fw} + M_{fb} M_{fw}$. The condition number increases significantly as nonlinearity order, i.e. N_{fw} and N_{fb} , increases. As illustrated in Figure 2(a), the condition number exceeds 10^{20} when the coefficients are more than 20. On the other hand, a 64 bits computer has dynamic range no more than 10^{20} , thus in such case numerical instability may become unavoidable.

To improve the robustness of the matrix inverse, numerical stable algorithms, such as QR decomposition or singular value decomposition, should be applied to $V^H V$.

3.3. Baseband Orthogonal Polynomial

With orthogonal polynomial, the condition number of the target matrix will be reduced to an acceptable extent, thus better performance can be achieved.

We define the orthogonal polynomial basis as

$$\psi_i(x) = \sum_{k=1}^i u_{ik} |x|^{k-1} x \quad (11)$$

and the NARMA model in (4) can be expressed as

$$y(n) = \sum_{i=1}^{N_{fw}} \sum_{j=0}^{M_{fw}-1} a'_{ji} \psi_i(x(n-j)) + \sum_{i=1}^{N_{fb}} \sum_{j=0}^{M_{fb}-1} b'_{ji} \psi_i(y(n-j-1)) \quad (12)$$

Note that (4) and (12) are, actually, equivalent to each other and (12) is constituted by orthogonal polynomial basis presented in (11). The coefficients u_{ik} can be found in [9], where the signal is assumed to be uniformly distributed. The closed form expression is given as

$$u_{ik} = \begin{cases} (-1)^{i+k} \frac{(i+k)!}{(i-1)!(i+1)!(k-i)!} & i \leq k \\ 0 & i > k \end{cases} \quad (13)$$

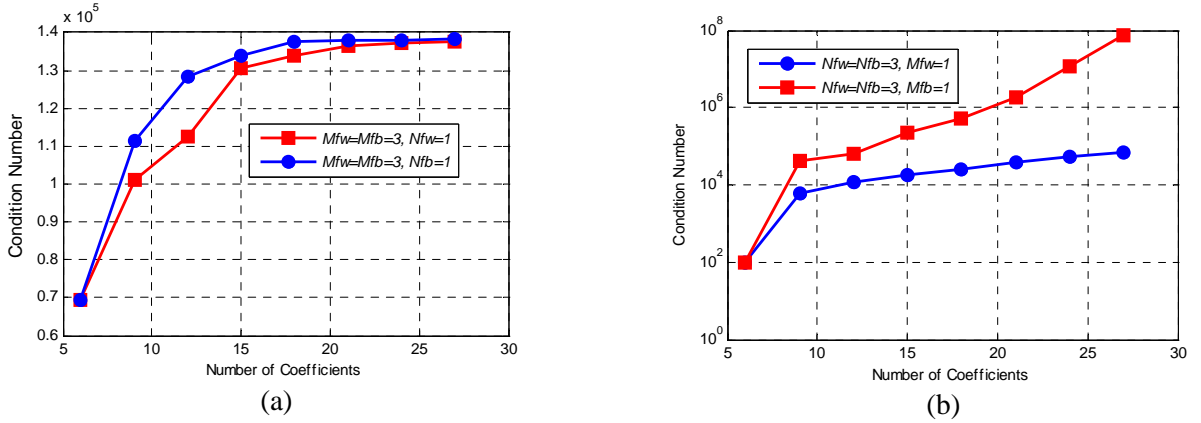


Fig.3. Condition number of $(VU^T)^H VU^T$: (a) sweeping N_{fw} and N_{fb} , respectively when M_{fw} and M_{fb} are fixed. (b) sweeping M_{fw} and M_{fb} , respectively when N_{fw} and N_{fb} are fixed

Assembling u_{ik} in a matrix $\phi[i,k] = u_{ik}$, and define the orthogonal polynomial coefficient matrix

$$U = \underbrace{[\phi \quad \phi \quad \dots \quad \phi]}_{M_{fb} M_{fw} \text{ copies}} \quad (14)$$

The LS solution associated with orthogonal polynomials yields

$$W_{orth} = ((VU^T)^H (VU^T))^{-1} (VU^T)^H Y \quad (15)$$

where, W_{orth} is orthogonal coefficient vector, then the NARMA model can be expressed as

$$Y_0 = VU^T W_{orth} \quad (16)$$

To report the benefit of introduction of orthogonal polynomial, we plot the condition number of $(VU^T)^H VU^T$ in Figure 3. Compared with Figure 2, we can see that with orthogonal polynomial the condition number is effectively reduced.

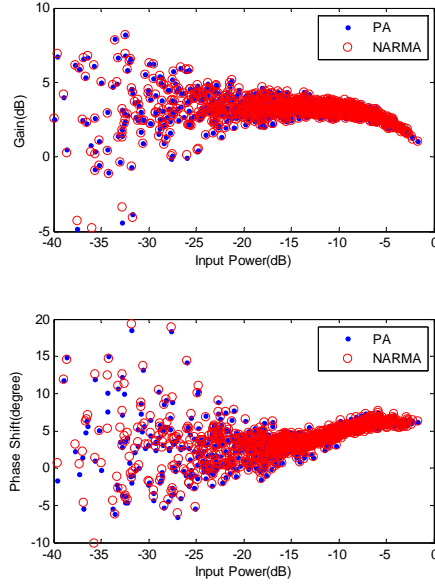


Fig.4. Gain and phase shift versus input power for actual PA and NARMA model

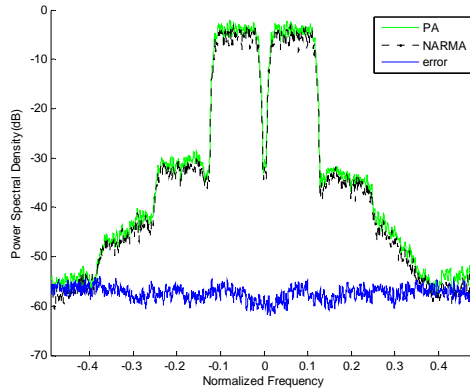


Fig.5. Output power spectral density of the measured data and NARMA model output. The error is also plotted in this figure.

4. Performance Estimation

To estimate the NARMA model’s performance, a MOSFET power amplifier with 10W output power working at 1.7 GHz is fabricated. A two-carrier WCDMA signal with 10MHz bandwidth is used to test the PA.

The signal measured from the input and output of the PA is used to abstract the NARMA model coefficients. The parameters of the model are set to be $N_{fw}=N_{fb}=5$, $M_{fw}=M_{fb}=4$, so the total number of coefficients are $N_{fw}M_{fw}+N_{fb}M_{fb}=40$. The orthogonal polynomial basis presented in the previous section is used to obtain higher numerical robustness.

The gain and phase shift versus input power for the measured data as well as the NARMA model are plotted in Figure 4. The position of modeled samples fit well for the actual power amplifier output. Figure 5 illustrates the power spectral density (PSD) of the PA output and modeled output. It is obvious that the modeled PSD almost coincides with PA output. The PSD of the error is around -58dB and is flat throughout the frequency range.

5. Conclusion

In this paper, we present a baseband NARMA model for behavioral modeling power amplifier with memory effect. The algorithm to abstract the NARMA model coefficients is also presented. In order to improve the numerical robustness of the algorithm, we take the advantage of the orthogonal polynomial to effectively reduce the condition number of the matrix which will be inverted in the least square algorithm. As shown in section VI our model shows good performance for modeling a MOSFET power amplifier.

6. References

- [1] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*, Reprinted, Melbourne, FL:Krieger, 1989.
- [2] G. T. Zhou, H. Qian, L. Ding, and R. Raich, "On the baseband representation of a bandpass nonlinearity," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2953-2957, Aug. 2005
- [3] L. Ding, G. T. Zhou, "Effects of even-order nonlinear terms on power amplifier modeling and predistortion linearization," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 156-162, Jan. 2004.
- [4] D. R. Morgan, Z. MA, J. Kim, M. G. Zierdt, J. Pastalan, "A Generalized Memory Polynomial Model for Digital Predistortion of RF Power Amplifiers," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3852-3860, Oct. 2006.
- [5] A.Zhu, J.C. Pedro, and T.J.Brazil, "Dynamic deviation reduction- based Volterra behavioral modeling of RF power amplifiers," *IEEE Trans. Microw. Theory Tech.*, vol. 54, no.12, pp. 4323-4332, Dec. 2006.
- [6] C. H. Loh, J. Y. Duh, *Analysis of nonlinear system using NARMA models*, Structural Engineering Earthquake, 1996
- [7] J. C. Pedro, S. A. Maas, "A Comparative Overview of Microwave and Wireless Power Amplifier Behavioral Modeling Approaches," *IEEE Trans. Microw. Theory Tech.*, vol. 53, no.4, pp. 1150-1162, Apr. 2005.
- [8] A. Horn, R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.
- [9] R. Raich, H. Qian, and T. Zhou, "Orthogonal polynomials for power amplifier modeling and predistorter design," *IEEE Tras. Veh. Technol.*, vol. 53, no. 5, pp.1468-1479, Sep.2004