A Comparative Analysis of the Complexity/Accuracy Tradeoff in Power Amplifier Behavior Models

Ali Soltani Tehrani, Haiying Cao, Thomas Eriksson, Magnus Isaksson, Christian Fager

***Abstract*—A comparative study of state-of-the-art behavioral models for microwave power amplifiers (PA) is presented in this paper. After establishing a proper definition for accuracy and complexity for power amplifier behavioral models, a short description on various behavioral models is presented. The main focus of this paper is on the modeling accuracy as a function of computational complexity.**

Data is collected from measurements on two PA's - a general purpose amplifier and a Doherty PA designed for WiMAX - and at different output power levels. The models are characterized in terms of accuracy and complexity for both in-band and out-of-band error. The results show that the generalized memory polynomial behavioral model has the best tradeoff for accuracy vs complexity.

***Index Terms*—Behavioral modeling, nonlinear distortion, nonlinear systems, power amplifier, radio communication, Volterra series.**

I. Introduction

THE development of future generation wireless transmission schemes, that operate at higher frequencies and require more bandwidth, has increased the demands on the linearity of power amplifiers. At the same time, with the increase of wireless communication users, the number of power consuming base stations has also increased. This has augmented the importance of developing power-efficient, linear devices in mobile base stations. Power consumption is even more critical in mobile devices where the power is driven from limited battery supplies.

One of the main power consuming devices in a transmitter is the power amplifier. One way to reduce power consumption in power amplifiers is to drive them to high efficiency regions. Unfortunately this has the adverse effect that PAs become more nonlinear [1]. This nonlinearity can introduce spectral regrowth, the degradation of signal quality, and other distortions. Nonlinearities can have even more dramatic effects as the dependency on linearity grows in modern wideband modulation schemes. It is also noticed in such schemes, that the amplifier tends to exhibit strong memory effects on the signal, which can further degrade the signal quality.

Due to the importance of the power amplifier, it is not surprising that the interest in power amplifier modeling has increased in recent years [2]–[4]. The main application of these models is for use in digital pre-distortion (DPD) linearization. DPD has been shown to reduce the size and cost for linearization compared to other linearization methods [5].

In [4] power amplifier behavioral models in literature were presented and classified in terms of memory: models without memory, with linear memory and with nonlinear memory. A similar presentation was done in [6] where a new behavioral model was also proposed. In [7] some important Volterra series-based models were compared and analyzed. The first attempts to compare these behavior models in an experimental setup was done in [8]. This work was extended in [2] with more behavioral models, more input signals, and cross-validation. The effect of signal bandwidth was also analyzed in [2].

While these works provided a necessary basis for power amplifier modeling, the performance was given as a function of model order, and do not consider the complexity of practical implementations. In this paper, we therefore present a comparison based on complexity, and show that the number of model parameters - which has been the dominating approach - is not always a good complexity measure. In order to have a fair comparison between the models, and have a basis where future models may be derived from, an accuracy/complexity comparison for behavioral models is proposed. It will be shown that as the number of parameters grow, most models will exhibit similar performance and only differ on the amount of complexity needed to reach this performance.

The complexity of behavioral models can also be of interest when practical implementations of these models are necessary, where the main factor is computational complexity. Having an accurate and low-complex behavioral model can be a key factor in evaluating communication system performance with computer simulations or when the amount of processing power is limited - such as in practical situations.

When analyzing models in terms of complexity, a proper definition of complexity is needed. The definition of complexity varies in different research fields, so an interpretation for complexity in PA behavioral modeling is established in Section II. In Section III some more-used behavioral models are presented and analyzed. Model evaluation measures are

This research has been carried out in the GigaHertz Centre in a joint research project financed by Swedish Governmental Agency of Innovation Systems (VINNOVA), Chalmers University of Technology, and ComHeat Microwave, Ericsson AB, Infineon Technologies, NXP Semiconductors and Saab AB.

A. Soltani Tehrani and T. Eriksson are with the Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden e-mail: [asoltani,thomase]@chalmers.se.

H. Cao and C. Fager are with the Department of Microtechnology and Nanoscience, Chalmers University of Technology, Gothenburg, Sweden, e-mail: [haiying,christian.fager]@chalmers.se

M. Isaksson is with the Center for RF Measurement Technology, University of Gävle, SE-80176 Gävle, Sweden. e-mail: [min@hig.se].

presented in Section IV. In Section V the measurement setup is introduced and in Section VI results of the comparison on two power amplifiers are provided and discussed.

II. Complexity

In literature, complexity has been notated by different measures [9]. Often it is measured in orders denoted by the Landau symbol $O(\cdot)$. Unfortunately this representation is not precise enough for practical applications [10].

A simple and common representation for behavioral model complexity is to only consider the number of modeling parameters and disregard the computing process complexity. This can determine the memory size needed for each behavioral model, and can be used when there are restrictions on the memory size.

Another approach is to simply record the running time of the different behavioral models in a software package. This is severely dependent on the hardware setup and the algorithm utilized. In order to have a fair comparison in this case, the algorithms must be optimized for the different behavioral models and for the hardware where they are tested.

The number of floating point operations or FLOPs is another widely used measure for complexity. In DSP hardware, complexity mainly stems from additions and multiplications. Since FLOPs is actually a measure for the number of additions, subtractions, and multiplications, it is sufficiently accurate to make a fair comparison between behavioral models. Hence, it will be used as the complexity measure in this paper.

Another important issue in behavioral model complexity is where the complexity originates from. The computational complexity can be classified into *identification complexity*, *running complexity* and *adaptation complexity*.

- *Identification complexity*: The identification procedure differs for the behavioral models. Due to statistical properties of measured signals, most Volterra-based models can be identified with a least squares estimate. Since the identification of the behavioral model is typically done offline, this complexity is normally not a major issue.
- *Adaptation complexity*: In practical systems, due to slight changes in the power amplifier such as temperature change or different mismatching effects, behavioral models might need to be updated at time intervals. These time intervals can normally be much larger than the symbol period. The adaptation of the behavioral model to these changes is considered adaptation complexity. In many instances where the variations are slow, this complexity may be of less importance.
- *Running complexity*: Running complexity is the number of calculations that is done on each sample when the model is utilized. This complexity severely limits the system due to the fact that it is a real-time problem. Depending on the application the maximum acceptable complexity varies. For a base station, there might be room for more complex algorithms and behavioral models while for mobile hand-held devices requirements are stricter. Since one of the main justifications for DPD linearization techniques is to have more power-efficient transmitters, it is essential that the power saved is not all spent on processing the DPD algorithm.

In this work, the focus is on running complexity due to its heavy computational costs on the system.

A final issue for complexity in behavioral models is that of parallelization. Some behavioral model algorithms have the inherent capability to be parallelized easier while others do not. This can be important in some applications where parallelization is possible. All the models that are presented in this study have this capability.

A. Complexity for Volterra-based models

The Volterra series is a widely used mathematical tool for modeling nonlinearities and memory in power amplifiers. The discrete baseband-equivalent form of the Volterra series which consists of a sum of multidimensional convolutions can be written as[1] [12]

$$
\begin{aligned}
y_{\text{Volterra}}[n] = & \sum_{\substack{p=1 \\ p \text{ odd}}}^{P} \sum_{m_1=0}^{M} \sum_{m_2=m_1}^{M} \cdots \sum_{m_{(p+1)/2}=m_{(p-1)/2}}^{M} \\
& \cdots \sum_{m_{(p+3)/2}=0}^{M} \cdots \sum_{m_p=m_{p-1}}^{M} h_{p,m_1,m_2,\cdots,m_p} \\
& \times \prod_{i=1}^{(p+1)/2} x[n-m_i] \prod_{k=(p+3)/2}^{p} x^*[n-m_k]. \quad (1)
\end{aligned}
$$

The Volterra series can be rewritten as:

$$
\mathbf{y}_{\text{Volterra}} = \mathbf{h} * \mathbf{X} \quad (2)
$$

where $\mathbf{h}$ is the vector containing all $h_{p,m_1,m_2,\cdots,m_p}$, and $\mathbf{X}$ is a matrix containing all permutations of $x[n]$:

$$
\mathbf{H}(x[n]) = \prod x[n-m_i] \prod x^*[n-m_k].
$$

Depending on the implementation of the behavioral model algorithm, the complexity for the Volterra-based behavioral models will differ. In [13] a general algorithm for implementing the Volterra series as a behavioral model is proposed. Here it is simplified and given in two steps:

Step i) Construct the basis functions (permutations of $\mathbf{X} = \mathbf{H}(x[n])$).

Step ii) Filter the basis with the kernels ($\mathbf{h} * \mathbf{X}$).

The second step is directly related to the number of kernels, since each kernel will be multiplied by the according basis function and then summed with the remaining results. Thus, it is solely dependent on the number of coefficients. The behavioral models will, however, differ in the construction of the basis functions.

An important issue in efficient algorithm design, is to refrain from generating already available data. For instance, while multiplying two signal values may require a certain number of FLOPs, delaying a signal does not. Therefore it is necessary to fully utilize all available permutations in the behavioral model

[1] In transmitter architectures the effect of filtering the output of the PA results in only odd order nonlinear power terms in the behavioral models [11].

algorithms. For example, $x[n-1]x[n-2]x^*[n-3]$ can easily be constructed from $x[n]x[n-1]x^*[n-2]$ by a simple delay.

Another issue is that terms that will be used in different combinations should be generated beforehand. For example, when constructing $|x[n]|^4$, if $|x[n]|^2$ is already available, using it will result in much lower complexity than constructing from scratch.

Finally, it is important to distinguish between multiplication of two complex numbers, and multiplication of two real-valued numbers. The latter is much less complex and only consists of 1 FLOP, while the former requires 6 FLOPs. Table I shows the operation-FLOP conversion used in this work.

TABLE I
NUMBER OF FLOPS FOR DIFFERENT OPERATIONS

Operation	**Number of FLOPs**
Conjugate	0
Delay	0
Real addition	1
Real multiplication	1
Complex addition	2
Complex-real multiplication	2
$\lvert.\rvert^2$	3
Complex multiplication	6

In the following sections, these considerations were taken into account to generate the behavioral models with minimum complexity.

III. BEHAVIORAL MODELS

PA behavioral models in literature can be classified into a few main groups namely Volterra-based, Artificial Neural Networks, table-lookup methods, and etc. In this section, we derive the computational complexities for Volterra-based models, due to their widespread use.

A. Volterra series

1) Definition: As discussed before, the Volterra series expands the impulse response model of a linear system by representing nonlinearity as a set of higher-order impulse responses named kernels [14]. It has been shown that a wide class of nonlinearities can be represented with good precision with a Volterra filter with nonlinear order P and memory length M [15], [16].

It can be seen that with the increase in memory depth and nonlinear order, the number of coefficients in the Volterra series grows exponentially, rendering the Volterra series useful only for weakly nonlinear systems.

2) Identification method: Due to the nature of the Volterra filters it is possible to identify the coefficients for the kernels with any linear estimation method. In this work after recording the values for the input and output of the amplifier, the least-square (LS) estimator is used with the pseudoinverse (Moore-Penrose pseudoinverse) of the output vs the permutations of the input. In [17] proof for the uniqueness of the results can be obtained.

3) Complexity: Efficient algorithms for implementing Volterra filters has been studied in literature. The complexity of the Volterra series has also been analyzed in [12]. In [13] the complexity for a non-optimized normal implementation of the Volterra series algorithm is given for a real input signal.

An important note to consider for this behavioral model is that all basis that do not contain the $x[n]$ term can be generated freely from previous terms with a simple delay. This follows simply from the previous example, $x[n-1]x[n-2]x^*[n-3]$ can easily be constructed from $x[n]x[n-1]x^*[n-2]$ with little complexity. The number of multiplications for the basis functions of a Volterra series behavioral model can be calculated as:

$$N(M,P) = \sum_{\substack{p=1 \\ p \text{ odd}}}^{P} f_{\text{Volterra}}(M,p) - f_{\text{Volterra}}(M-1,p) \tag{3}$$

where M is the memory depth, P is the nonlinear order, and $f_{\text{Volterra}}(M,p)$ is the number of coefficients in each kernel and is

$$f_{\text{Volterra}}(M,p) = \binom{M+\frac{p+1}{2}}{\frac{p+1}{2}}\binom{M+\frac{p-1}{2}}{\frac{p-1}{2}}. \tag{4}$$

Since all these are complex multiplications except $x[n]x[n]^*$, the total number of FLOPs for the basis construction is:

$$C_{\text{Volterra, basis}}(M,P) = 6N(M,P) - 3. \tag{5}$$

For the filtering each coefficient requires 6 FLOPs for the complex multiplication and 2 for the complex summation. The total number of FLOPs for filtering is thus:

$$C_{\text{Volterra, filter}}(M,P) = 8 \sum_{\substack{p=1 \\ p \text{ odd}}}^{P} f_{\text{Volterra}}(M,p). \tag{6}$$

The total number of FLOPs thus becomes:

$$C_{\text{Volterra}}(M,P) = C_{\text{Volterra, basis}} + C_{\text{Volterra, filter}}. \tag{7}$$

With the rather large number of FLOPs needed for the algorithm, it can be noticed that the Volterra series is useful practically only for relatively low nonlinear orders and memory lengths.

B. Memory Polynomial

1) Definition: The memory polynomial behavioral (MP) model is an extension of the basic polynomial model with linear memory [18]. This model, also known as parallel Hammerstein in literature, is a parallelization of a nonlinear function followed by a linear memory. The baseband equivalent memory polynomial model can be written as:

$$y_{\text{MP}}[n] = \sum_{\substack{p=1 \\ p \text{ odd}}}^{P} \sum_{m=0}^{M} h_{p,m} x[n-m] \left|x[n-m]\right|^{p-1}. \tag{8}$$

The MP is linear in parameters, and the identification is thus similar to the unconstrained Volterra.

2) Complexity: Due to the inherent reusability of the basis functions in this model, the running complexity is much lower than Volterra series model. In general the only term that has to be generated is $x[n]\,|x[n]|^p$ for each p. All other terms, i.e $x[n-m]\,|x[n-m]|^p$ can be generated by delaying existing terms. The basis can be constructed with:

$$C_{\text{MP, basis}}(M,P) = 3 + (P-1) \tag{9}$$

FLOPs.

The number of coefficients in this model is equal to $(M+1)(P+1)/2$ [2] and these will require 8 FLOPs each (similar to the Volterra series):

$$C_{\text{MP, filter}}(M,P) = 8\left(\frac{P+1}{2}\right)(M+1). \tag{10}$$

We notice that the complexity for this model grows linearly with the number of parameters, and the main source of the complexity is in the filtering and summation part.

C. Generalized Memory Polynomial

1) Definition: The generalized memory polynomial (GMP) behavioral model was proposed in [6] and extends the memory polynomial model by including more cross-terms. The formulation is for this model is:

$$\begin{aligned} y_{\text{GMP}}[n] = \sum_{\substack{p=1\\p\text{ odd}}}^{P}\sum_{m=0}^{M}\sum_{g=0}^{G} h_{p,m,g}\, x[n-m] \\ \times |x[n-m-g]|^{p-1}. \end{aligned} \tag{11}$$

This model adds an extra degree of freedom in parameters for the behavioral model that corresponds to the amount of memory in lagging terms which will be called G. When $G=0$ this model becomes equivalent to the MP model. Identification is similar to the MP and Volterra series model.

2) Complexity: The complexity of this behavioral model is similar to the MP model, but with the added terms. The initial basis construction is slightly higher than MP and is equal to:

$$C_{\text{GMP, basis}}(M,P,G) = 3 + (P-1)(G-1) \tag{12}$$

FLOPs.

The number of coefficients for this model is equal to $(G+1)(M+1-G)(P+1)/2$, and they have to be filtered. The complexity for the filtering becomes:

$$\begin{aligned} C_{\text{GMP, filter}}(M,P,G) = {} & 8(G+1) \\ \times & \;(M+1-G)\left(\tfrac{P+1}{2}\right). \end{aligned} \tag{13}$$

The main source of complexity for this model is in the filtering, like the MP model.

D. Volterra with Dynamic Deviation Reduction

1) Definition: In [19] a new mathematical model for power amplifiers is presented based on modeling the static and dynamic parts separately. This work was constructed into the behavioral model format in [20] and [21]. Further work was done in [22] and [23]. The latter is the model that is used in this paper.

In this model the Volterra series is reconstructed with respect to the dynamic deviation in the coefficients, and a parameter which we denote as R is introduced which is the number of dynamic deviations in the model. This gives an extra restriction so the Volterra series can be reduced. The identification is similar to the previous methods. A baseband equivalent of this model here is expanded from [24] and can be written as:

$$\begin{aligned} y_{\text{DDR}}[n] = {} & \sum_{\substack{p=1\\p\text{ odd}}}^{P} h_{p,0}\,\underbrace{x[n]|x[n]|^{p-1}}_{\text{zero order dynamic}} \\ + & \sum_{\substack{p=1\\p\text{ odd}}}^{P}\sum_{m_1=1}^{M} h_{p,m_1}\,\underbrace{x[n-m_1]|x[n]|^{p-1}}_{1^{st}\text{ order dynamics path 1}} \\ + & \sum_{\substack{p=3\\p\text{ odd}}}^{P}\sum_{m_2=1}^{M} h_{p,m_2}\,\underbrace{x^*[n-m_2]x^2[n]|x[n]|^{p-3}}_{1^{st}\text{ order dynamics path 2}} \\ + & \sum_{\substack{p=3\\p\text{ odd}}}^{P}\sum_{m_3=1}^{M}\sum_{m_4=m_3}^{M} h_{p,m_3,m_4} \\ \times & \underbrace{x[n-m_3]x[n-m_4]x^*[n]|x[n]|^{p-3}}_{2^{nd}\text{ order dynamics path 1}} \\ + & \sum_{\substack{p=3\\p\text{ odd}}}^{P}\sum_{m_5=1}^{M}\sum_{m_6=1}^{M} h_{p,m_5,m_6} \\ \times & \underbrace{x[n-m_5]x^*[n-m_6]x[n]|x[n]|^{p-3}}_{2^{nd}\text{ order dynamics path 2}} \\ + & \sum_{\substack{p=5\\p\text{ odd}}}^{P}\sum_{m_7=1}^{M}\sum_{m_8=m_7}^{M} h_{p,m_7,m_8} \\ \times & \underbrace{x^*[n-m_7]x^*[n-m_8]x[n]|x[n]|^{p-5}}_{2^{nd}\text{order dynamics path 3}}. \end{aligned} \tag{14}$$

where up to 2^{nd} order dynamics are shown.

2) Complexity: In this model it is important to note that all basis functions in this model contain the term $x[n]$, and while this is desirable for the accuracy of the modeling, it reduces the reusability of the basis functions and increases complexity. Therefore efficient algorithms for this behavioral model are harder to achieve compared to previous models.

Setting $P>3$ and $R=2$ and using the methods to reduce complexity as previously discussed, the complexity for constructing the basis is calculated as:

$$\begin{aligned} C_{\text{DDR, basis}} = {} & 9 + 6M + (M+1)(P-1) + 6M\left(\tfrac{P-3}{2}\right) \\ + & \;6\left(\tfrac{P-3}{2}\right)\left(M^2 + \tfrac{M(M+1)}{2}\right) \\ + & \;6\left(\tfrac{P-5}{2}\right)\tfrac{M(M+1)}{2}. \end{aligned} \tag{15}$$

The first two terms represent the initial construction of important combinations. The complexity for the zero order and first order dynamic path one is the third term, and the rest are for the different path of the formulation above.

The number of coefficients for this model is given by:

$$f_{\text{DDR}}(P,M) = \frac{P+1}{2}(1+M) + \frac{P-1}{2} \times \left(M + \frac{M(M+1)}{2} + M^2\right) + \frac{P-3}{2}\frac{M(M+1)}{2} \quad (16)$$

and the complexity for filtering is thus:

$$C_{\text{DDR, filter}}(P,M) = 8 f_{\text{DDR}}(P,M). \quad (17)$$

E. Kautz-Volterra and Laguerre-Volterra

1) Definition: The first attempts at constructing orthogonal functions as basis functions for power amplifiers was presented in [25] and [26]. This idea was further expanded in [27] and [28] which resulted in the Laguerre and Kautz-Volterra behavioral models, respectively. The main difference between these two behavioral models is that in the Volterra expansions model with Laguerre functions the orthonormal basis pole λ is chosen to be real, while in the Kautz-Volterra behavioral model these poles are chosen to be complex. In [28] the model's ability to separate linear and nonlinear memory effects was introduced, i.e. the poles for the nonlinear orders could be different from the linear ones.

These models are actually generalizations of the Volterra series model, i.e. the Volterra series is a special case of the Laguerre and Kautz-Volterra model when the poles are at zero.

2) Identification method: Due to the nature of these behavioral models, the identification procedure is not as straightforward as in the previous models. Many identification methods exist, but in this work a full search of poles for per each nonlinear order was done. After finding the optimum poles, the problem becomes a normal least square estimation and can be performed with the same technique as in the Volterra filter. This method becomes attractive when the poles are known beforehand, or when it is possible to have an initial off-line identification of the amplifier to identify the poles. Further extraction methods can be found in [27]. If the orthonormal basis poles are not known before hand the identification is much more complex and can be prone to local minima and maxima.

3) Complexity: While the identification for such models may be problematic, the running complexity is not affected much. Once the poles for the different power levels are calculated, the behavioral model is similar to the Volterra filter with the addition of an extra filter with one pole per nonlinear order. Therefore the construction of basis function requires:

$$C_{\text{KV, basis}}(M,P) = 6N(M,P) - 3 + 8\left(\frac{P-1}{2}\right). \quad (18)$$

where $N(M,P)$ is from (3). The filtering is similar to the Volterra series since they have the same number of parameters:

$$C_{\text{KV, filter}}(M,P) = 8\sum_{\substack{p=1\\ p \text{ odd}}}^{P} f_{\text{Volterra}}(M,p) \quad (19)$$

where $f_{\text{Volterra}}(M,p)$ is given in (4).

IV. Model evaluation measures

In this section, we analyze different model performance measures. Many performance measures have been used in literature to validate power amplifier behavioral models. A study on the different measures used is done in [29]. Some of the more used measures include normalized mean square error (NMSE), adjacent channel power ratio (ACPR) and adjacent channel error power ratio (ACEPR). NMSE is defined as [2]

$$\text{NMSE} = \frac{\sum_n |y_{\text{meas}}[n] - y_{\text{model}}[n]|^2}{\sum_n |y_{\text{meas}}[n]|^2}. \quad (20)$$

The input signal to power amplifiers is normally band-limited, but due to the nonlinearity effect of the PA the output signal has some spectral regrowth. Since most of the power is in-band, NMSE has the inherent characteristic that it mainly measures the in-band error. In instances where the out-of-band performance of the power amplifier is of more importance the adjacent channel power ratio (ACPR) and the adjacent channel error power ratio (ACEPR) are normally used [30].

ACEPR is a measure of the modeling error in the adjacent channels related to the power in the channel and is given by

$$\text{ACEPR} = \max_{m=1,2}\left[\frac{\int_{(\text{adj})_m} |Y_{\text{meas}}(f) - Y_{\text{mod}}(f)|^2}{\int_{\text{ch}} |Y_{\text{meas}}(f)|^2}\right] \quad (21)$$

where $Y_{\text{mod}}(f)$ is the Fourier transform of the model data, $Y_{\text{meas}}(f)$ is the Fourier transform of the measurement data. The integration in the denominator is over the in-band channel signal bandwidth and the integration in the numerator is over the adjacent channels to the signal channel with the same bandwidth. As seen in (21) the ACEPR is defined as the larger of the values evaluated for both the lower ($m = 1$) and upper ($m = 2$) adjacent channels.

Since both NMSE and ACEPR are error measures, lower values show better agreement between the model and the PA measurement. It is important to note that having a low NMSE does not necessarily correspond to having a low ACEPR, i.e, some models have lower NMSE while others can have lower ACEPR.

Further measures also exist in literature like the weighted error-to-signal power ratio (WESPR) proposed in [30], the memory effect ratio (MER) and the memory effect modeling ratio (MEMR) [31], but are not considered in this work.

V. Measurement setup

The block diagram of the measurement setup used to characterize the power amplifier behavior models is shown in Fig. 1. The modulator used is an Agilent E4438C vector signal generator (VSG) and an Agilent 54845A digital storage oscilloscope (DSO) is used as a vector signal analyzer. The baseband I/Q data is generated in the computer and downloaded to VSG. The VSG modulates the data to an RF carrier and in order to have enough input power for the PA under test, fed through a preamplifier. This signal is then fed to the power amplifier which is the device under test (DUT) and both the input of the DUT and the output are captured simultaneously by the DSO. The DSO sends the RF signals

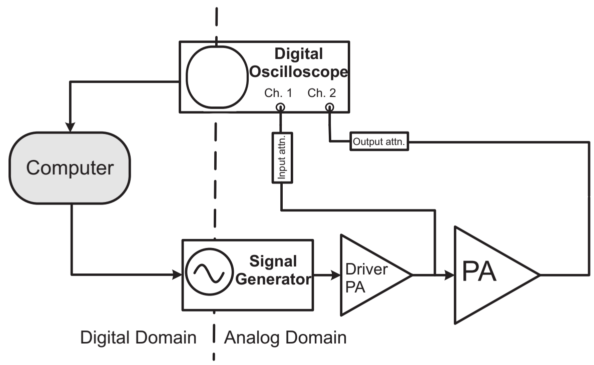


Fig. 1. Outline of the measurement setup used for evaluation of the behavioral models.

back to the PC where they are down-converted to baseband I/Q data. All devices are connected by GPIB and triggered in synch.

To enhance the dynamic range of the signal and decrease the noise variance a statistical averaging technique is used [32]. The experimental results reported here are based on 500 averaged measurements, which resulted in an effective dynamic range of 65 dB. In order to have time alignment the DSO is triggered by the VSG and a 10 MHz reference is connected from the DSO to the VSG. Also to obtain more precise time alignment correlation techniques are utilized.

In order to have a proper open test analysis, the validation of the behavioral models should be done with a different data set than the one used for identification. The procedure that has been analyzed in this paper is as follows:

- Download an I/Q input signal to the VSG to construct the RF signal and record the input and output of the DUT.
- Split the data set to identification data and validation data
- Calculate the behavioral model parameters using the identification data
- Compare the power amplifier output to the behavioral models' prediction using the validation data

The reason for splitting the data set and not re-downloading it to the VSG is to make sure the identification data and validation data are subject to the same temperature and bias conditions.

A. Input Signals

Since the identification process is dependent on the input signal, the experiment should be done with data as similar to a practical case as possible. The WCDMA data used in this work had a bandwidth of 3.84 Mchips/s and was modulated to a carrier frequency of 1 GHz to match the PAs available. The peak to average ratio of this data was 7.6 dB. The WiMAX-like data had 4 MHz bandwidth, peak to average of 7.4 dB and was modulated to a 2.6 GHz carrier.

The in-band channel in this work was defined as the signal bandwidth at the center frequency, and the adjacent channels were defined as the signal bandwidth at ± 5 MHz from the center frequency.

An important issue in the identification process for PA behavioral models is the number of model parameters k vs data set size N. If N is not sufficiently large compared to k, the estimation procedure can be hampered with over-fitting and uncertainties in the model parameters can grow. This effect is seen in the mean-square error for the estimation, which is roughly $(1+2\frac{k}{N})\sigma^2$, where σ^2 is the measurement noise variance [33]. In this work in order to fulfill this requirement, 25000 samples are used for identification, 28000 samples for validation, and the maximum number of model parameters estimated is 350.

B. Power amplifiers

Two power amplifiers were studied in this research: a wideband 3 W class AB commercial solid state PA[2] and a 100 W Doherty power amplifier[3] for WiMAX applications. The class AB amplifier was analyzed at two power levels, one with input power -4 dBm and the other at -12 dBm. For clarity the experiments are classified in three scenarios:

Scenario 1: Class AB power amplifier with WCDMA data and input power -12dBm
Scenario 2: Class AB power amplifier with WCDMA data and input power -4dBm
Scenario 3: Doherty power amplifier with WiMAX-like data

The dynamic AM/AM plot for these power amplifier is shown in Fig. 2(a), and the spectrum for the measured signals in Fig. 2(b).

VI. Results

In this section, the power amplifier models are compared with respect to accuracy vs complexity. The accuracy was evaluated using both NMSE and ACEPR.

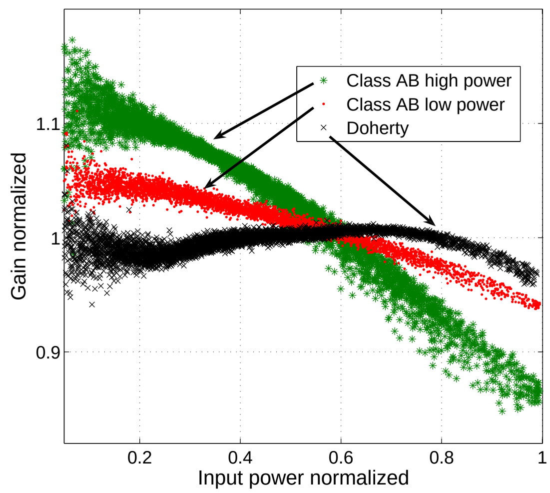
The lowest NMSE that was obtainable regardless of complexity for the different behavioral models is shown in Table II(a). Table II(b) shows the best results obtained regardless of complexity for ACEPR. For the memory polynomial (MP), Volterra, and Kautz-Volterra models, the numbers inside the parenthesis represent (P, M). For Volterra with dynamic deviation reduction (Volterra DDR) they represent (P, M, R) and for generalized memory polynomial (Generalized MP) they represent (P, M, G).

It can be noticed from these tables that the generalized memory polynomial model consistently outperforms the other models in both NMSE and ACEPR. It can also be noted that the second scenario, where the power amplifier is more nonlinear, has lower values for both measures than the other scenarios.

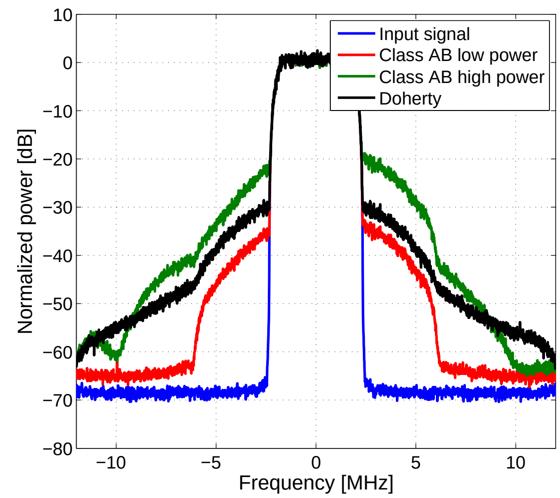
The results here differ from those of [2] in that a much wider range of nonlinear order and memory tap combinations are analyzed, and a more exhaustive search of model parameters has been done.

It is also noticed that the best NMSE and ACEPR results are close for different models. This is because as the nonlinear order and memory depth grows in the models, the uncertainties in modeling parameters increase and dominate the error.

[2]MiniCircuits ZHL-1000-3W
[3]NXP semiconductors



(a)



(b)

Fig. 2. Characteristics for the PAs tested. Left:AM/AM plot for the different scenarios analyzed. Right: Signal spectrum for the three scenarios, from top to bottom scenario 2, scenario 3 and scenario 1 and input signal.

TABLE II
THE BEST RESULTS OBTAINED FOR THE DIFFERENT SCENARIOS REGARDLESS OF COMPLEXITY. THE PARENTHESIS REPRESENT THE CORRESPONDING MODEL ORDER.

(a) Best NMSE results.

Model	Scenario 1	Scenario 2	Scenario 3
MP	-50.6 (9,11)	-36.5 (11,10)	-44.7 (11,8)
Volterra	-51.3 (7,2)	-38.4 (5,3)	-44.9 (9,1)
Kautz-Volterra	-51.34 (5,3)	-38.5 (5,3)	-45.0 (9,1)
Volterra DDR	-51.74 (9,3,2)	-38.5 (9,4,2)	-44.8 (11,3,2)
Generalized MP	-51.8 (9,9,3)	-38.5 (9,7,3)	-45.0 (11,7,3)

(b) Best ACEPR results.

Model	Scenario 1	Scenario 2	Scenario 3
MP	-58.3 (11,12)	-47.5 (11,12)	-52.9 (11,10)
Volterra	-60.1 (7,2)	-49.7 (5,3)	-53.7 (9,1)
Kautz-Volterra	-60.2 (5,2)	-49.8 (5,3)	-53.7 (9,1)
Volterra DDR	-60.6 (9,3,2)	-49.4 (9,4,2)	-53.3 (11,3,2)
Generalized MP	-60.8 (13,9,3)	-49.7 (9,6,3)	-53.3 (11,8,3)

A. Scenario 1

Fig. 3(a) shows a comparison of NMSE between behavioral models with respect to the number of parameters for the class AB power amplifier. A static nonlinear model is also included for reference.

An important issue that should be emphasized in the figures is that an exhaustive search has been done on the parameter space and the optimal curve for each model has been found.

It can be seen in this figure that as the number of parameters grow, the amount of improvement gained with excess parameters is limited. With a large number of parameters, all models have similar performance and higher nonlinear order and memory depths do not yield better results. The rate at which these models achieve this performance however differs between models.

From Fig. 3(a) it can also be seen that the memory polynomial model gives the lowest error compared to other models with a low number of parameters. The Volterra with dynamic deviation reduction model outperforms other models within a range of parameters, and finally the generalized memory polynomial model gives the lowest error when the number of parameters increases further.

While this figure can give certain insight to how model perform compared to one another, as previously discussed this analysis can be unfair. In Fig. 3(b) the comparison is done vs FLOPs.

In Fig. 3(b) we notice that the generalized memory polynomial model outperforms other models consistently. This is due to the fact that this model is less complex to run than the Volterra with dynamic deviation reduction model. This figure supports our hypothesis that complexity is a more appropriate measure than number of parameters.

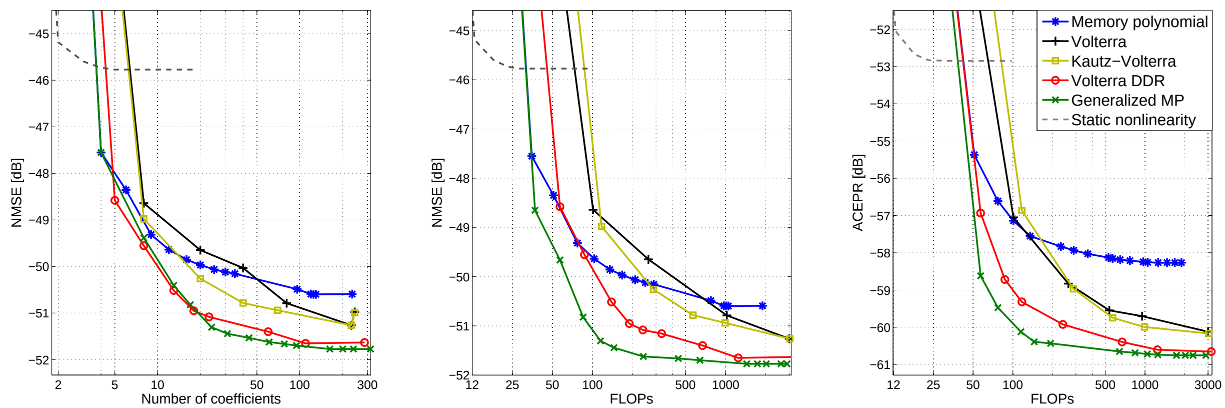
Finally in Fig. 3(c) the out of band performance given by the ACEPR measure for the different models vs FLOPs is presented. It should be noted that in the identification procedure, the minimization criterion was NMSE and not out of band performance.

The ACEPR values yield similar results for this signal input power.

B. Scenario 2

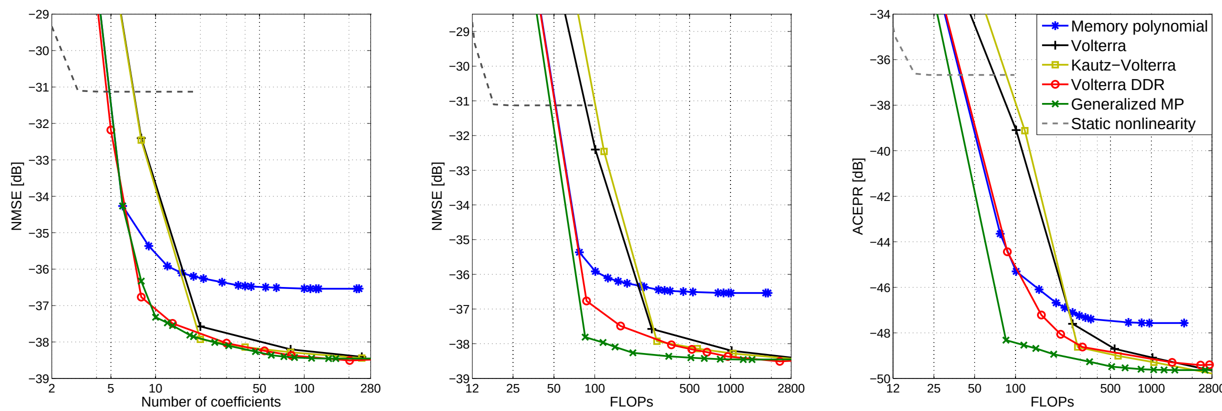
The class AB power amplifier is driven harder and has more nonlinear characteristics in this scenario, therefore the modeling accuracy is degraded. In Fig. 4(a) the comparison is done with respect to the number of parameters, while Fig. 4(b) is with respect to the number of FLOPs.

From these two figures we can notice that while in Fig. 4(a) several models have approximately the same accuracy vs number of parameters (except for memory polynomial), in Fig. 4(b) the generalized memory polynomial model shows the best tradeoff behavior. It can also be noticed that as the number of FLOPs increase, all behavioral models tend to have the same accuracy. The ACEPR can be seen in Fig. 4(c)



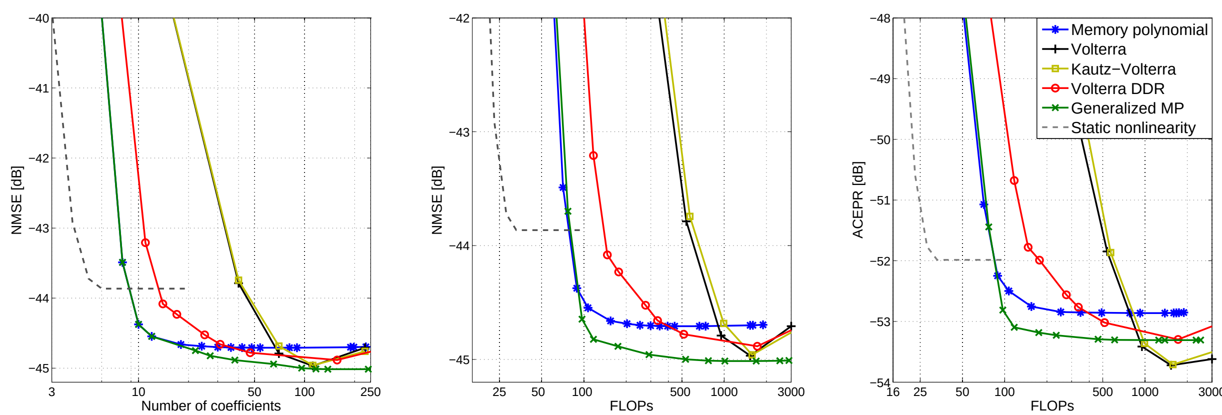
(a) NMSE vs number of parameters (b) NMSE vs FLOPs (c) ACEPR vs FLOPs

Fig. 3. Class AB power amplifier with low input power. Volterra DDR stands for the Volterra with dynamic deviation reduction model. The legend is identical for all figures.



(a) NMSE vs number of parameters (b) NMSE vs FLOPs (c) ACEPR vs FLOPs

Fig. 4. Class AB power amplifier with high input power. Volterra DDR stands for the Volterra with dynamic deviation reduction model. The legend is identical for all figures.



(a) NMSE vs number of parameters (b) NMSE vs FLOPs (c) ACEPR vs FLOPs

Fig. 5. Doherty power amplifier. Volterra DDR stands for the Volterra with dynamic deviation reduction model. The legend is identical for all figures.

C. Scenario 3

In Fig. 5, the results of modeling the Doherty power amplifier is shown. Once again it can be noticed that the GMP model outperforms the other models, but in terms of ACEPR, with a large number of parameters the Volterra and Kautz-Volterra surpasses all other models.

Another important observation is that the models generally do not have significant improvement from the static nonlinear model. In the previous scenarios, a 7-10 dB accuracy gain was achieved with the Volterra based behavioral models compared to the static case. In this scenario only a 1-2 dB gain is achieved compared to the static nonlinear model. One main difference between the behavioral models analyzed and the static model is that the latter does not model memory effects. From these results, we can notice that memory effects are not dominant compared to the static nonlinearity in the Doherty architecture. This can be traced to the internal circuitry for this class of PAs.

VII. Conclusion

In this work, efficient algorithms for some widely used behavioral models were developed, and the the computational complexity of these algorithms were measured in FLOPs. The behavioral models were tested on measurement data from two power amplifiers and it was noticed that the generalized memory polynomial model outperformed all other models in terms of accuracy vs FLOPs consistently. It was also noticed that number of parameters was not necessarily an appropriate measure for behavioral model comparison.

The results indicate that for a Doherty power amplifier, memory effects are not as pronounced as nonlinear distortions, and the existing models are not able to model the memory effects in this PA effectively. This is due to the inherent characteristics of this power amplifier class.

Acknowledgment

The authors would like to thank H. Nemati with the GigaHertz Centre and A. Zhu with University of Dublin for their helpful assistance. The authors would also like to acknowledge Yong Yang at NXP Semiconductors, Cumberland, U.S.A., for providing the Doherty PA used in the measurements.

References

[1] S. Cripps, *RF power amplifiers for wireless communications*, 2nd ed. Boston: Artech House, 2006.

[2] M. Isaksson, D. Wisell, and D. Ronnow, “A comparative analysis of behavioral models for RF power amplifiers,” *IEEE Trans. Microw. Theory Tech.*, vol. 54, no. 1, pp. 348–359, 2006.

[3] A. Zhu, M. Wren, and T. J. Brazil, “An efficient Volterra-based behavioral model for wideband RF power amplifiers,” in *Proc. IEEE MTT-S Int. Micro. Symp. Dig.*, vol. 2, 2003, pp. 787–790 vol.2.

[4] J. C. Pedro and 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. 1150–1163, 2005.

[5] P. B. Kenington, “Linearized transmitters: an enabling technology for software defined radio,” *IEEE Commun. Mag.*, vol. 40, no. 2, pp. 156–162, 2002.

[6] D. R. Morgan, Z. Ma, J. Kim, M. G. Zierdt, and 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, 2006.

[7] A. Zhu and T. J. Brazil, “An overview of Volterra series based behavioral modeling of RF/microwave power amplifiers,” in *Proc. Wirel. Micro. Techn. Conf.*, 2006, pp. 1–5.

[8] I. Santamaria, J. Ibanez, M. Lazaro, C. Pantaleon, and L. Vielvo, “Modeling nonlinear power amplifiers in OFDM systems from subsampled data: a comparative study using real measurements,” *EURASIP J. Appl. Signal Process.*, vol. 2003, no. 1, pp. 1219–1228, 2003, 1283369.

[9] S. LLoyd, *Programming the Universe: A Quantum Computer Scientist Takes on the Cosmos*. Alfred A. Knopf, 2006.

[10] D. de Ridder, E. Pekalska, and R. P. W. Duin, “The economics of classification: Error vs. complexity,” in *Intern. Conf. Patt. Recog.*, vol. p.20244, 2002.

[11] L. Ding and 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, 2004.

[12] V. Mathews and G. L. Sicuranza, *Polynomial Signal Processing*, ser. Wiley Series in Telecommunications and Signal Processing. New York: Wiley, 2000.

[13] G. M. Raz and B. D. van Veen, “Baseband Volterra filters for implementing carrier based nonlinearities,” *IEEE Trans. Signal Process.*, vol. 46, no. 1, pp. 103–114, 1998.

[14] M. J. Reed and M. O. J. Hawksford, “Identification of discrete Volterra series using maximum length sequences,” *IEE Proc.-Circuit Device Syst.*, vol. 143, no. 5, pp. 241–248, 1996.

[15] I. W. Sanderg, “Uniform approximation with doubly finite Volterra series,” *IEEE Trans. Signal Process.*, vol. 40, no. 6, 1992.

[16] R. D. Nowak and B. D. Van Veen, “Tensor product basis approximations for Volterra filters,” *IEEE Trans. Signal Process.*, vol. 44, no. 1, pp. 36–50, 1996.

[17] G. M. Raz and B. D. Van Veen, “Blind equalization and identification of nonlinear and IIR systems-a least squares approach,” *IEEE Trans. Signal Process.*, vol. 48, no. 1, pp. 192–200, 2000.

[18] J. Kim and K. Konstantinou, “Digital predistortion of wideband signals based on power amplifier model with memory,” *Electron. Lett.*, vol. 37, no. 23, pp. 1417–1418, 2001.

[19] F. Filicori and G. Vannini, “Mathematical approach to large-signal modelling of electron devices,” *Electron. Lett.*, vol. 27, no. 4, pp. 357–359, 1991.

[20] D. Mirri, G. Iuculano, F. Filicori, G. A. V. G. Vannini, G. A. P. G. Pasini, and G. A. P. G. Pellegrini, “A modified Volterra series approach for the characterization of non-linear dynamic systems,” in *Proc. IEEE Instr. Meas. Tech. Conf.*, vol. 1, 1996, pp. 710–715 vol.1.

[21] D. Mirri, F. Filicori, G. Iuculano, and G. Pasini, “A nonlinear dynamic model for performance analysis of large-signal amplifiers in communication systems,” *IEEE Trans. Instrum. Meas.*, vol. 53, no. 2, pp. 341–350, 2004.

[22] A. Zhu, J. Dooley, and T. J. Brazil, “Simplified Volterra series based behavioral modeling of rf power amplifiers using deviation-reduction,” in *Proc. IEEE MTT-S Int. Micro. Symp. Dig.*, 2006, pp. 1113–1116.

[23] 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, 2006.

[24] 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 Trans. Microw. Theory Tech.*, vol. 56, no. 7, pp. 1524–1534, 2008, 0018-9480.

[25] P. M. Lavrador, J. C. Pedro, and N. B. Carvalho, “A new Volterra series based orthogonal behavioral model for power amplifiers,” in *Proc. Asia-Pacific Micro. Conf.*, vol. 1, 2005, p. 4 pp.

[26] J. Dooley, B. O’Brien, and T. J. Brazil, “Behavioral modeling of RF power amplifiers using adaptive recursive polynomial functions,” in *Proc. IEEE MTT-S Int. Micro. Symp. Dig.*, 2006, pp. 852–855.

[27] A. Zhu and T. J. Brazil, “RF power amplifier behavioral modeling using Volterra expansion with Laguerre functions,” in *Proc. IEEE MTT-S Int. Micro. Symp. Dig.*, 2005, p. 4 pp.

[28] M. Isaksson and D. Ronnow, “A Kautz-Volterra behavioral model for RF power amplifiers,” in *Proc. IEEE MTT-S Int. Micro. Symp. Dig.*, 2006, pp. 485–488.

[29] P. Landin, M. Isaksson, and P. Handel, “Comparison of evaluation criteria for power amplifier behavioral modeling,” in *Proc. IEEE MTT-S Int. Micro. Symp. Dig.*, 2008, pp. 1441–1444.

[30] D. Wisell, M. Isaksson, and N. Keskitalo, “A general evaluation criteria for behavioral PA models,” in *Proc. 69th ARFTG Conf. Dig.*, 2007, pp. 251–255.

[31] H. Ku and J. S. Kenney, “Behavioral modeling of nonlinear RF power amplifiers considering memory effects,” *IEEE Trans. Microw. Theory Tech.*, vol. 51, no. 12, pp. 2495–2504, 2003.

[32] C. Fager and K. Andersson, "Improvement of oscilloscope based RF measurements by statistical averaging techniques," in *Proc. IEEE MTT-S Int. Micro. Symp. Dig.*, 2006, pp. 1460–1463.
[33] L. Ljung, *System identification: theory for the user, 2nd ed.* Prentice-Hall, Inc., 1999.