

Low-Pass Equivalent Behavioral Modeling of RF Power Amplifiers Using Two Independent Real-Valued Feed-Forward Neural Networks

Luiza B. C. Freire, Caroline de França, and Eduardo G. de Lima*

Abstract—Feed-forward artificial neural networks (ANNs) can provide the adequate model required for the linearization of power amplifiers (PAs) used in wireless communication systems. A common characteristic of previously available ANN-based models for linearization purposes is the use of a single real-valued ANN having two outputs. The contribution of this work is to report the benefits of performing such behavioral modeling based on two independent real-valued ANNs, where each network has a unique output. The proposed ANN-based model is applied to the behavioral modeling of a GaN HEMT class AB PA, and its accuracy is compared to previous approaches in two different scenarios. First, in case of similar number of network parameters, it is observed that the proposed ANN-based model can reduce the normalized mean-square error (NMSE) by up to 1.3 dB. Second, in a situation of comparable modeling accuracy (NMSE = -40 dB), it is observed that the proposed ANN-based model can reduce the number of network parameters by up to 40% (from 62 to 38 real-valued parameters).

1. INTRODUCTION

Transmitters for wireless communication systems have received a lot of attention in the microwave community [1–5]. The requirements of high efficiency and high linearity put a lot of challenges for the design of wireless transmitters, specially the power amplifier (PA). Indeed, PAs based on solid-state transistors can only provide acceptable efficiencies when driven at strong nonlinear regimes [1].

In one hand, rigorous specifications on linearity are imposed by regulatory agencies to avoid interferences among neighbor users and to guarantee the quality of service [2]. On the other hand, high efficiency is crucial to assuring longer time of battery autonomy in mobile devices and also to reducing the costs of heat dissipation and energy consumption in base-stations [2]. The impact of the trade-off between linearity and efficiency is further accentuated because recent wireless standards require the transmission of envelope signals having a high peak-to-average power ratio (PAPR) [3, 4].

The simultaneous demands for linearity and efficiency can only be attained if a linearization scheme is integrated into the transmitter chain. Among the available strategies, digital baseband predistortion (DPD) is a cost-effective linearization solution [5]. The successful design of a DPD scheme relies strongly on the availability of a low complexity and high accurate low-pass equivalent behavioral model for the PA [6]. In literature, there are various techniques that can provide an adequate mathematical description for the PA behavior. Volterra series [7–9] and artificial neural networks (ANNs) [10–15] are the most widely reported techniques that can simultaneously describe nonlinear and dynamic behaviors. ANNs have the advantage of requiring a lower number of parameters than the Volterra series and have more general validity than polynomial approximations.

ANNs having only real-valued inputs, outputs, weights and biases, are widely reported in literature for the low-pass equivalent behavioral modeling of PAs [11–15]. Since low-pass equivalent behavioral models relate complex-valued envelope signals at the PA input and output, complex-valued signals must

Received 2 July 2014, Accepted 29 July 2014, Scheduled 4 August 2014

* Corresponding author: Eduardo Goncalves de Lima (elima@eletrica.ufpr.br).

The authors are with the Department of Electrical Engineering, Federal University of Paraná, Centro Politécnico, CP. 19011, Curitiba, PR 81531-980, Brazil.

be converted into real-valued signals. Indeed, different approaches have been proposed to perform these conversions. Specifically, in [11–13] the complex-valued input and output envelopes are decomposed into rectangular components, in [14] into polar coordinates and in [15] into amplitude and phase difference components. However, the ANN-based behavioral models in [11–15] use a single ANN having two outputs that estimate real-valued components of the PA output signal.

This work first proposes a novel ANN-based PA behavioral model that, contrarily to the previous approaches in [11–15], requires two independent real-valued ANNs, each network having just one output. Second, based on input-output data measured on a GaN HEMT class AB PA, it reports a better trade-off between modeling accuracy and computational complexity of the proposed ANN-based model than previous approaches.

This work is organized as follows. Section 2 reviews PA behavioral modeling and describes basic guidelines to be followed in order to get accurate and low complexity models. Section 3 presents two widely adopted architectures for feed-forward ANNs, namely the three-layer perceptron and radial basis function. Section 4 first addresses how to apply ANNs to the behavioral modeling of bandpass systems in a way that minimizes the computational cost, and then describes the proposed ANN-based model. Using data measured on a GaN HEMT class AB PA, Section 5 assesses the accuracy of the proposed ANN-based model in comparison with previous approaches. Section 6 includes the conclusions.

2. POWER AMPLIFIER BEHAVIORAL MODELING

The power amplifier (PA) is present at the wireless transmitter chain, immediately before the antenna. The purpose of the PA is to amplify the power of the signal to be irradiated by the antenna. In wireless communications, the PA is excited by carrier signals (whose frequency is in the order of GHz) modulated by complex-valued envelope signals having a bandwidth in the order of several MHz [1, 2]. In other words, the signals that must be handled by the PA have non-null energy only at frequencies near the center frequency and also bandwidth much lower than the center frequency. Such signals are commonly designated as bandpass signals [16], and therefore, PAs can be seen as systems having one input and one output, which relate bandpass signals. Indeed, the input and output frequency-selective matching network circuits inside the PA architectures are responsible for the PA bandpass behavior, e.g., only bandpass inputs can produce any measurable output, and the only observable outputs are also bandpass signals. Denoting the PA input and output signals by x and y , respectively, and exploiting their bandpass nature, it is obtained:

$$x = \Re[\tilde{x} \exp(j\omega_c t)] = a \cos(\omega_c t + \theta), \quad (1)$$

and

$$y = \Re[\tilde{y} \exp(j\omega_c t)] = b \cos(\omega_c t + \varphi + \theta), \quad (2)$$

where ω_c is the carrier frequency, while $\tilde{x}$ and $\tilde{y}$ represent the complex-valued input and output envelope signals, respectively.

Ideally, the PA output signal must be a linear replica (apart from constant gain and delay) of the applied input signal. However, in order to improve the power efficiency, especially when a linearization scheme is integrated into the transmitter chain, non-negligible nonlinear behaviors are observed in the PA [2, 5]. The most evident nonlinear behavior is the gain compression that occurs when the PA is driven at higher power levels. Besides, different kinds of dynamic effects are also observable, including low-frequency memory effects (on the order of the envelope bandwidth) associated to active-device self-heating, biasing circuitry and trapping effects [17, 18], as well as high-frequency memory effects (on the order of the carrier frequency) attributed to non-uniform frequency response of the matching networks [6].

For linearization purposes, an accurate and low complexity nonlinear dynamic model for the PA is required [6]. PA models with different levels of internal description and computational complexity can be found. The physics-based models are the most detailed descriptions. For a complete characterization of a PA at a physical level, the partial-differential equations of diffusion and charge transport for the active device modeling, and partial-differential wave equations for the passive device modeling must be included, which results in a model of very high computational complexity. An intermediate description at the circuit-equivalent level is available specifically for the solid-state PAs. It includes the Kirchhoff

laws and the constitutive equations of the elementary devices. Finally, a less detailed description at system level is possible, leading to the so-called behavioral or black-block models. They describe the PA based on mathematical equations with adjustable parameters based on terminal measurements performed on the PA, using no or little *a priori* information about its internal structure.

The most suitable PA model for linearization purposes is the discrete-time behavioral model. Volterra series [7–9] and artificial neural networks (ANNs) [10–15] are the most widely reported techniques that can provide an adequate mathematical description for the PA behavior. The selection of a particular technique targets the improvement of the trade-off between increasing modeling accuracy and reducing computational cost. However, independent of the particular chosen technique (for instance, Volterra or ANN), designers are strongly encouraged to follow two common recommendations, detailed in the sequence. In fact, their adoption can hugely reduce the computational complexity of a behavioral model, without negligible effects on the modeling accuracy.

Behavioral models are classified based on the nature of the signals involved [6]. If they relate the bandpass signals at the PA input and output, they are called bandpass behavioral models. Instead, if it is the relationship between the complex-valued envelopes at the PA input and output being modeled, they are called low-pass equivalent behavioral models. In one hand, accurate bandpass behavioral models require that the instantaneous PA output signal is formulated as a function of the instantaneous, as well as a huge set (on the order of thousands) of previous, samples of the PA input signal. Indeed, an accurate bandpass behavioral model must take into account long-term memory effects (on the order of MHz), at the same time having a sampling frequency of at least twice the carrier frequency (on the order of GHz) to comply with the Nyquist criterion. On the other hand, accurate low-pass equivalent behavioral models can be achieved even when the instantaneous PA output is assumed to be a function of just the instantaneous, plus a few set of previous, samples of the PA input signal. In fact, since only complex-valued envelope signals are available to the low-pass equivalent behavioral model, the maximum frequency that must be handled by the model is reduced to a few times of the envelope bandwidth, and therefore, the sampling frequency can be set to a value on the MHz range. In summary, the choice between bandpass and low-pass equivalent behavioral models has a significant impact on the computational cost of a behavioral model. Therefore, the first common recommendation when choosing a PA model for linearization purposes is the adoption of a low-pass equivalent behavioral model.

Low-pass equivalent behavioral models are insensitive to the carrier frequency (ω_c) information. As a consequence, if $\tilde{x}$ is subject to an arbitrary operator f , out-of-band contributions, having non-null energy only in the vicinity of harmonic frequencies of ω_c , can be easily generated by the low-pass equivalent behavioral model [19]. Out-of-band contributions can not be associated to any measurement performed on the PA and, therefore, do not improve the modeling accuracy. However, the generation of contributions having no physical counterpart significantly increases the computational complexity of a low-pass equivalent behavioral model. Therefore, the second common recommendation when choosing a PA model for linearization purposes is the adoption of a low-pass equivalent behavioral model that avoids the generation of out-of-band contributions. In other words, some constraints associated to the bandpass behavior of PA under narrowband excitations must be imposed to guarantee that only physical in-band contributions are estimated by the low-pass equivalent behavioral models [15, 19]. Indeed, a necessary (but not sufficient) condition is that the PA model must estimate the output envelope signal by an operator f that has odd-parity with respect to the complex-valued envelope signal applied to the PA input. Besides, the operator f must be chosen in a way that the unitary scalar value multiplying the carrier frequency ω_c must be copied from the PA input to the estimated output, e.g., the expression $1\omega_c t$ that is intrinsically part of the signal applied at the PA input must also be part of the estimated output signal with exactly the same coefficient equal to one.

3. ARTIFICIAL NEURAL NETWORKS

In feed-forward ANNs, the network outputs are restricted to be functions of only the signals applied as network inputs [10]. No feed-back mechanism is allowed. Three-layer perceptron (TLP) and radial basis function (RBF) are the most common architectures for feed-forward ANNs. Figure 1(a) shows the block diagram of a discrete-time TLP having E inputs, R neurons in the hidden layer and S outputs, while Figure 1(b) shows the block diagram of a discrete-time RBF also having E inputs, R neurons

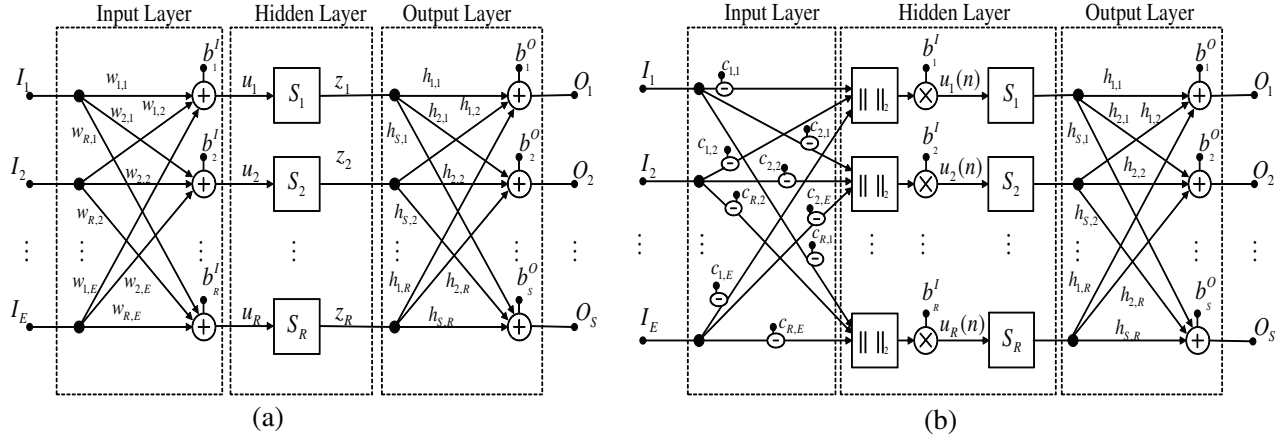


Figure 1. Block diagram of neural networks: (a) three-layer perceptron and (b) radial basis function.

and S outputs. A widely employed application for TLP and RBF is the modeling of dynamic nonlinear systems. Indeed, nonlinear activation functions, designated by S in Figure 1, are required when the TLP and RBF are used to model nonlinear systems. Moreover, to account for dynamic behaviors, TLP and RBF must estimate the instantaneous (n) sample of an output signal based on information of the input signals at instantaneous (n) as well as past ($n-M$), up to the memory length M , samples.

In a TLP, the signals u applied to the activation functions (S) are obtained from the linear combination of the real-valued input signals I , weighted by input parameters (w) and added to input biases (b^I). In an RBF, the signals u applied to the activation functions (S) are obtained by taking the Euclidean distance between the real-valued input vector and the neuron center vector (given by the coefficients c), multiplied by input biases (b^I).

A common choice for S in a TLP is the hyperbolic tangent sigmoid function given by

$$S(u) = \frac{2}{1 + \exp(-2u)} - 1, \quad (3)$$

while the Gaussian function, given by

$$S(u) = \exp(-u^2), \quad (4)$$

is typically used as activation function in an RBF.

The O real-valued output signals of both TLP and RBF are given by the linear combination of the signals z at the output of the activation functions, weighted by output parameters (h) and added to output biases (b^O).

To train the TLP in order to obtain its real-valued network parameters (w, b^I, h, b^O), the back-propagation algorithm is used [10]. To train the RBF in order to obtain its real-valued network parameters (c, b^I, h, b^O), the orthogonal least-squares algorithm is employed [20].

4. ANN-BASED LOW-PASS EQUIVALENT PA BEHAVIORAL MODELING

When applying the ANN architectures shown in Figure 1 to mimic the behavior of a physical PA, the obvious choice is to use an ANN having one input ($E = 1$) and one output ($O = 1$). Indeed, if the same input is applied to both physical PA and ANN model, it is desired that the ANN output provides a very close approximation to the measured physical PA output. However, the ANN architectures shown in Figure 1 represent discrete-time models. By this it means that ANN models can only provide estimations for the PA output at discrete-time instants. Using an ANN having just one input and one output, the best that can be done is to estimate the instantaneous sample of the PA output as a function of just the instantaneous sample of the PA input. Such memoryless model is not able to take into account dynamic behaviors observed in physical PAs. Actually, when applied to the modeling of dynamic systems, ANNs shown in Figure 1 must have more than one input. Therefore, a much more

accurate model is achieved if the instantaneous sample of the PA output is assumed to be a function of the instantaneous sample plus a few past samples (up to the memory length M) of the PA input. In other words, accurate ANN models must have at least $(M + 1)$ inputs and one output.

In order to significantly reduce the computational complexity of an ANN model, as described in Section 2, the use of a low-pass equivalent model is mandatory. In this case, only the complex-valued envelopes at the PA input and output are available to the behavioral model. Therefore, accurate low-pass equivalent models must have at least $(M + 1)$ complex-valued inputs (representing instantaneous as well as past, up to the memory length M , samples) and one complex-valued output. Nevertheless, in the ANN architectures shown in Figure 1, the network inputs and outputs are restricted to be real-valued signals. As a consequence, complex-valued envelopes must be somehow mapped into real-valued signals. Indeed, at least two real-valued signals are necessary to keep all the information provided by a single complex-valued envelope. Therefore, accurate ANN models must have at least $2(M + 1)$ real-valued inputs and two real-valued outputs.

Moreover, as detailed in Section 2, a further reduction in computational cost can be obtained by paying attention to avoid the generation of non-physical contributions that for sure will not improve the modeling accuracy. To that purpose, the strategy adopted in [15] was to modify the ANN real-valued inputs. For instance, in [15] the complex-valued envelopes were mapped into real-valued signals through a non conventional decomposition technique, according to Figure 2(a). The reasoning followed in [15] was to guarantee that all of the ANN real-valued inputs and outputs kept no information about the carrier frequency. Observe that the polar angle of any complex-valued input (denoted by θ and omitting the time sample to simplify the notation) is intrinsically attached to the carrier frequency according to $(\omega_c t + \theta)$. So, the polar angle of complex-valued input envelopes can not be directly applied as ANN inputs if it is intended to guarantee that only physical in-band contributions are generated by the ANN model. In fact, the polar angles must be modified in a way that their connection with ω_c is completely destroyed. For doing that, it suffices to calculate the difference between any two polar angles taken at distinct time samples, once $(\omega_c t + \theta_1) - (\omega_c t + \theta_2) = (\theta_1 - \theta_2)$. Furthermore, observe that the absolute value of a complex-valued number, by definition, has no connection with the polar angle and, as a consequence, no relationship with ω_c . Therefore, the absolute values of the complex-valued input envelopes can be directly applied as ANN inputs, without any modification. Considering that all of ANN inputs shown in Figure 2(a) are insensitive to ω_c , the same also applies to the outputs estimated by the ANN. The operations performed on the ANN outputs shown in Figure 2(a) are essential to re-establish the intrinsically connection between a polar angle of a complex-valued envelope and the carrier frequency. Moreover, it is also necessary to keep the unitary scalar value multiplying the carrier frequency ω_c , e.g., $1\omega_c t$, once the physical PA system relates bandpass real-valued signals given by a carrier signal at ω_c modulated by the complex-valued envelopes available to the low-pass equivalent models.

The previous ANN, proposed in [15] and whose block diagram is shown in Figure 2(a), estimates two separate real-valued signals (b and φ) through a single ANN. Indeed, it is a common practice in PA behavioral modeling literature (see for instance [11–14]) the use of a single ANN having two real-valued outputs. In either the TLP or RBF architectures shown in Figure 1, the signals (z) at the output of the activation functions (S) can be interpreted as elements of a subspace Z having dimension R , according to $Z = z_1, z_2, \dots, z_R$. Therefore, each ANN output can be seen as the best projection (apart from a possible constant shift) of the desired signal into the Z subspace. In fact, in both TLP and RBF shown in Figure 1, each ANN output is obtained by first performing a linear combination of the signals at the output of the activation functions and then shifting the result by a constant bias. In the common practice of employing a single ANN having two outputs, the two desired signals that must be estimated at the ANN outputs are very different from each other. As a consequence, it can be a very hard task to find a unique subspace Z , whose best projections of the two desired outputs into Z can provide simultaneously two accurate estimations. The difficulty is accentuated when the number of neurons in the hidden layer is low, because the number of elements composing the subspace Z is directly given by the number of neurons in the hidden layer.

However, if it relaxes the obligation of a single subspace Z , a much better modeling accuracy is expected, especially in case of a low number of neurons in the hidden layer. Based on this reasoning, Figure 2(b) shows the block diagram of the proposed ANN-based low-pass equivalent PA behavioral

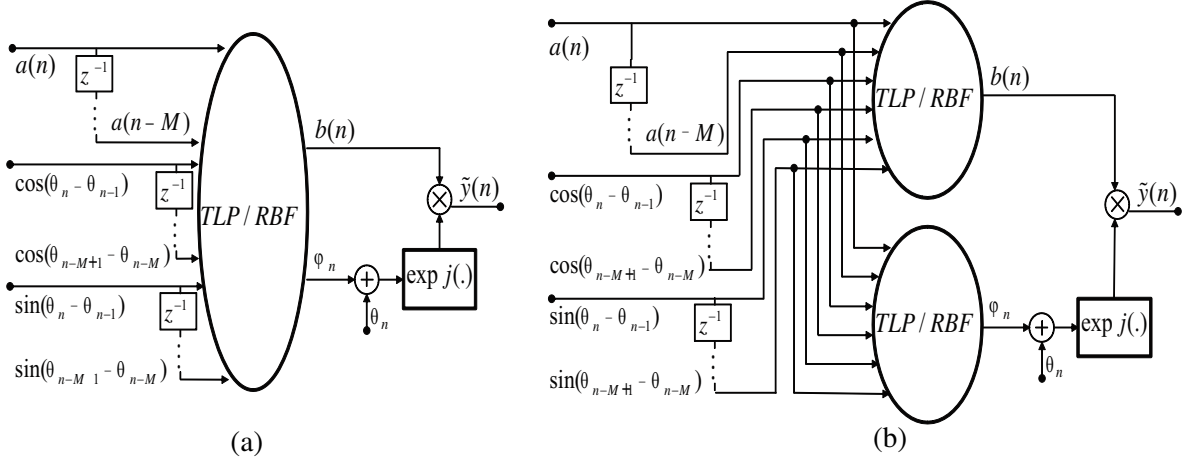


Figure 2. Block diagram of ANN-based low-pass equivalent PA behavioral models: (a) previous [15] and (b) proposed.

model. Observe that, now, two distinct ANNs are employed and, as a consequence, two different subspaces (Z_1 and Z_2) are available. Considering that, now, each ANN has a single output, each ANN output has its own (and exclusive) subspace. Therefore, it is reasonable to expect that accurate projections of desired signals into their exclusive subspaces are possible even when the numbers of elements in the subspaces are low. Additionally, a better trade-off between modeling accuracy and number of network parameters can be achieved if the numbers of neurons in the hidden layers are set separately for each ANN. For instance, the ANN that estimates (b) can have R_1 neurons, while the ANN that estimates (φ) can have R_2 neurons.

5. EXPERIMENTAL VALIDATION

In this section, the accuracy (as a function of the number of network parameters) of the proposed ANN-based model, shown in Figure 2(b), is compared to the previous ANN-based model of [15], shown in Figure 2(a). The number of network parameters is varied by changing the number of neurons in the hidden layer. The difference between measured and estimated complex-valued output envelopes gives the error signal. The accuracies of PA behavioral models are assessed by the normalized mean-square error (NMSE), as defined in [21].

The device under test (DUT) is a GaN HEMT class AB PA. The input signal is a carrier signal at 900 MHz modulated by a 3GPP WCDMA signal having a bandwidth of 8.84 MHz. A Rohde & Schwarz FSQ vector signal analyzer (VSA) is used to measure the discrete-time data sampled at 61.44 MHz. The PA average output power is set to 26 dBm. The measured data is divided into two subsets: one for network training and one for network validation. Only validation results are reported here.

For the behavioral modeling of the DUT, the TLP architecture shown in Figure 1(a) is selected. Hyperbolic tangent sigmoid functions given by (3) are employed as TLP activation functions. The TLP real-valued parameters (w , b^I , h , b^O) are trained by the back-propagation algorithm [10]. The memory length is kept constant and equal to $M = 2$.

Figure 3 shows the NMSE results as a function of the number of network parameters (e.g., for a varying number of neurons in the hidden layer). Observe that a better trade-off between modeling error and number of network parameters is achieved by the proposed model in comparison with the previous model of [15]. Specifically, the proposed model presents lower errors when the number of network parameters are similar (up to 1.3 dB reduction in NMSE) and can significantly reduce the minimum number of network parameters required for fulfilling specific threshold levels on modeling error (up to 40%, from 62 to 38 real-valued network parameters, in order to reduce the NMSE to -40 dB).

Figures 4, 5 and 6 further illustrate that the proposed ANN-based model having just 38 network parameters provides an excellent estimation for the measured PA output signal, in this way validating the proposed ANN-based behavioral model. Specifically, Figure 4 shows the power spectral densities

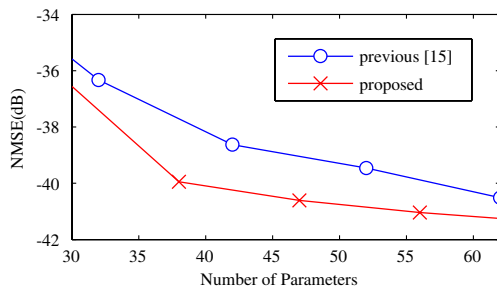


Figure 3. NMSE results as a function of the number of network parameters. The GaN HEMT class AB PA is modeled by TLPs having $M = 2$ and a variable number of neurons in the hidden layer.

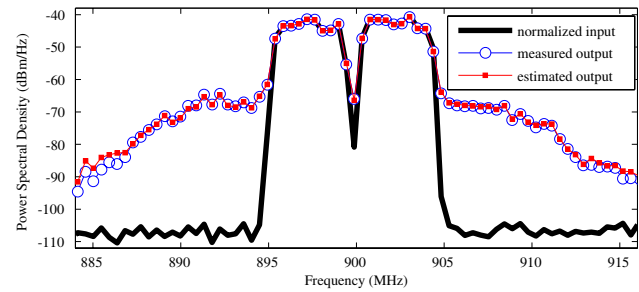


Figure 4. Power spectral densities of the measured and estimated PA output signals. The GaN HEMT class AB PA is modeled by the proposed TLP having $M = 2$ and 38 network parameters.

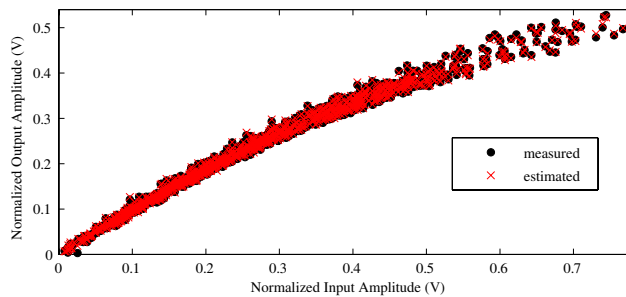


Figure 5. Measured and estimated normalized instantaneous AM-AM conversions. The GaN HEMT class AB PA is modeled by the proposed TLP having $M = 2$ and 38 network parameters.

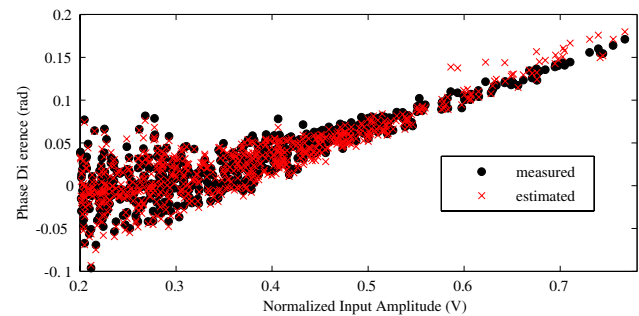


Figure 6. Measured and estimated instantaneous AM-PM conversions. The GaN HEMT class AB PA is modeled by the proposed TLP having $M = 2$ and 38 network parameters.

(PSDs) of the measured and estimated PA output signals, Figure 5 shows the (measured and estimated) normalized instantaneous amplitude of the PA output signal as a function of the instantaneous amplitude of the PA input signal, the so-called instantaneous AM-AM conversion and Figure 6 shows the (measured and estimated) instantaneous phase difference (between PA output and input signals) as a function of the instantaneous amplitude of the PA input signal, the so-called instantaneous AM-PM conversion. Observe that the proposed ANN-based model having just 38 network parameters can accurately estimate the measured AM-AM and AM-PM characteristics.

6. CONCLUSIONS

This work has addressed the benefits of using two independent real-valued feed-forward artificial neural networks, each network having a single output, for the low-pass equivalent behavioral modeling of power amplifiers for wireless communication systems. A better trade-off between modeling error and computational cost achieved by the proposed model was confirmed based on input-output data measured on a GaN HEMT class AB PA. It was reported that the proposed ANN-model, in comparison with previous approaches, requires a much lower number of network parameters in order to comply with threshold levels on modeling error. Alternatively, in a situation of similar number of network parameters, the proposed model can significantly reduce the modeling error in comparison with previous approaches.

ACKNOWLEDGMENT

The authors would like to thank the financial support provided by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

REFERENCES

1. Cripps, S., *RF Power Amplifiers for Wireless Communications*, Artech House, Norwood, 2006.
2. Raab, H., P. Asbeck, S. Cripps, P. B. Kenington, Z. B. Popovic, N. Potheary, J. F. Sevic, and N. O. Sokal, "Power amplifiers and transmitters for RF and microwave," *IEEE Trans. Microw. Theory Tech.*, Vol. 50, No. 3, 814–826, 2002.
3. Raychaudhuri D. and N. B. Mandayam, "Frontiers of wireless and mobile communications," *Proc. IEEE*, Vol. 100, No. 4, 824–840, 2012.
4. Piazzon, L., R. Giofre, P. Colantonio, and F. Giannini, "A method for designing broadband Doherty power amplifiers," *Progress In Electromagnetics Research*, Vol. 145, 319–331, 2014.
5. Kenington, P. B., *High Linearity RF Amplifier Design*, Artech House, Norwood, 2000.
6. Pedro, J. C. 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, 1150–1163, 2005.
7. Schetzen, M., "Nonlinear system modeling based on the Wiener theory," *Proc. IEEE*, Vol. 69, No. 12, 1557–1573, 1981.
8. Wang, H., H. Ma, and J. Chen, "A multi-status behavioral model for the elimination of electrothermal memory effect in DPD system," *Progress In Electromagnetics Research C*, Vol. 47, 103–109, 2014.
9. Sun, G., C. Yu, Y. Liu, S. Li, and J. Li, "An accurate complexity-reduced simplified volterra series for RF power amplifiers," *Progress In Electromagnetics Research C*, Vol. 47, 157–166, 2014.
10. Haykin, S., *Neural Networks: a Comprehensive Foundation*, Prentice Hall, New Jersey, 1999.
11. Liu, T., S. Boumaiza, and F. M. Ghannouchi, "Dynamic behavioral modeling of 3G power amplifiers using real-valued time-delay neural networks," *IEEE Trans. Microw. Theory Tech.*, Vol. 52, No. 3, 1025–1033, 2004.
12. Yuan, X.-H. and Q. Feng, "Behavioral modeling of RF power amplifiers with memory effects using orthonormal Hermite polynomial basis neural network," *Progress In Electromagnetics Research C*, Vol. 34, 239–251, 2013.
13. Fu, K., C. L. Law, and T. T. Thein, "Novel neural network model of power amplifier plus IQ imbalances," *Progress In Electromagnetics Research B*, Vol. 46, 177–192, 2013.
14. Isaksson, M., D. Wisell, and D. Ronnow, "Wide-band dynamic modeling of power amplifiers using radial-basis function neural networks," *IEEE Trans. Microw. Theory Tech.*, Vol. 53, No. 11, 3422–3428, 2005.
15. Lima, E. G., T. R. Cunha, and J. C. Pedro, "A physically meaningful neural network behavioral model for wireless transmitters exhibiting PM-AM/PM-PM distortions," *IEEE Trans. Microw. Theory Tech.*, Vol. 59, No. 12, 3512–3521, 2011.
16. Jeruchim, M. C., P. Balaban, and K. S. Shanmugan, *Simulation of Communication Systems — Modeling, Methodology, and Techniques*, Kluwer Academic/Plenum Publishers, New York, 2000.
17. Vuolevi, J. H. K., T. Rahkonen, and J. P. A. Manninen, "Measurement technique for characterizing memory effects in RF power amplifiers," *IEEE Trans. Microw. Theory Tech.*, Vol. 49, No. 8, 1383–1389, 2001.
18. Bosch, W. and G. Gatti, "Measurement and simulation of memory effects in predistortion linearizers," *IEEE Trans. Microw. Theory Tech.*, Vol. 37, No. 12, 1885–1890, 1989.
19. Benedetto, S., E. Biglieri, and R. Daffara, "Modeling and performance evaluation of nonlinear satellite links — A Volterra series approach," *IEEE Trans. Aerosp. Electron. Syst.*, Vol. 15, No. 4, 494–507, 1979.

20. Chen, S., C. F. N. Cowan, and P. M. Grant, "Orthogonal least squares learning algorithm for radial basis function networks," *IEEE Trans. Neural Netw.*, Vol. 2, No. 2, 302–309, 1991.
21. Isaksson, M., D. Wisell, and D. Ronnow, "A comparative analysis of behavioral models for RF power amplifiers," *IEEE Trans. Microw. Theory Tech.*, Vol. 54, No. 1, 348–359, 2006.