

# Reconfigurable DPD Based on ANNs for Wideband Load Modulated Balanced Amplifiers Under Dynamic Operation From 1.8 to 2.4 GHz

Estefanía Guillena, *Graduate Student Member, IEEE*, Wantao Li<sup>ID</sup>, *Graduate Student Member, IEEE*, Gabriel Montoro<sup>ID</sup>, *Member, IEEE*, Roberto Quaglia<sup>ID</sup>, *Member, IEEE*, and Pere L. Gilabert<sup>ID</sup>, *Senior Member, IEEE*

**Abstract**—This article proposes a methodology to ensure linear amplification of a load modulated balanced amplifier (LMBA) while keeping the power efficiency as high as possible over a frequency band ranging from 1.8 to 2.4 GHz and where the transmitted signals can present different bandwidth (BW) configurations. The proposed reconfigurable linearization methodology consists of, in a first step, tuning some free parameters (with dependence on the signal BW and frequency of operation) of the LMBA to trade-off linearity and power efficiency. In a second step, two multipurpose adaptive digital predistortion (DPD) linearizers are considered, properly combined with crest factor reduction (CFR) techniques, to meet the required linearity specifications. Either a DPD based on artificial neural networks or a DPD based on polynomials can be selected taking into account the compromise between computational complexity and linearization performance. Experimental results will validate the proposed methodology to guarantee the linearity levels (ACPR < -45 dBc and EVM < 1%) with high power efficiency in an LMBA under dynamic transmission, where both the signal BW (from 20 and up to 200-MHz instantaneous BW) and frequency of operation (in the range of 1.8–2.4 GHz) change.

**Index Terms**—Artificial neural network (ANN), digital predistortion (DPD), load-modulated balanced amplifiers (LMBAs), power efficiency.

## I. INTRODUCTION

POWER-EFFICIENT amplification has been a hot research topic since the introduction of nonconstant envelope

Manuscript received February 2, 2021; revised April 19, 2021 and May 25, 2021; accepted May 28, 2021. This work was supported in part by the Spanish Government (Ministerio de Ciencia, Innovación y Universidades) and the Fondo Europeo de Desarrollo Regional (FEDER) under Grant TEC2017-83343-C4-2-R and in part by the Generalitat de Catalunya under Grant 2017 SGR 813. This article is an expanded version from the International Workshop on Integrated Nonlinear Microwave and Millimetre-Wave Circuits (INMMiC), Cardiff, U.K., July 2020. (Corresponding author: Pere L. Gilabert.)

This work did not involve human subjects or animals in its research.

Estefanía Guillena, Wantao Li, Gabriel Montoro, and Pere L. Gilabert are with the Department of Signal Theory and Communications, Universitat Politècnica de Catalunya (UPC)—Barcelona Tech, 08034 Barcelona, Spain (e-mail: estefania.guillena@estudiantat.upc.edu; plgilabert@tsc.upc.edu).

Roberto Quaglia is with the Centre for High Frequency Engineering, Cardiff University, Cardiff CF24 3AA, U.K. (e-mail: quagliar@cardiff.ac.uk).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TMTT.2021.3091672>.

Digital Object Identifier 10.1109/TMTT.2021.3091672

digital modulations. Starting from W-CDMA in 3G, the peak-to-average power ratio (PAPR) of signals have kept increasing through the use of orthogonal frequency division multiplexing (OFDM) in 4G (LTE, LTE-Advanced), and now cyclic prefix OFDM in new radio (NR) 5G. In addition, to satisfy the requirements for higher transmission rates, the signal bandwidths (BWs) that the amplification architectures have to accommodate are always increasing (e.g., from a few megahertz in 3G, to tenths of megahertz in 4G, to hundreds of megahertz in NR 5G). In addition, efficient amplification is required in several frequency bands in NR 5G with multiple numerology (i.e., subcarrier spacing) and channel BWs. This demands providing the baseband processing with some degree of reconfigurability to adapt to the changing transmission requirements.

When dealing with signals presenting high PAPR, the power amplifier (PA) needs to operate at large power back-off leading to a serious degradation of average efficiency. To avoid wasting excessive power resources, highly efficient amplification architectures based on dynamic load or dynamic supply modulation have been proposed in the literature. Some of the most popular solutions are envelope tracking PAs [1], Doherty PAs [2], [3], load-modulated balanced amplifiers (LMBAs) [4], [5], and LINC or outphasing PAs [6], [7]. In each case, these highly efficient topologies require the use of digital predistortion (DPD) linearization to guarantee the stringent linearity requirements of today's systems, especially with the increasing signal BW.

All PA architectures based on active load modulation, such as Doherty, LMBA, and outphasing, rely on the nonlinear interaction between multiple transistors to enhance the average efficiency in presence of modulated signals with large dynamic range. Although these architectures can be designed with a single RF input to simplify their use in a transmitter, there are benefits in maintaining separate inputs controlled by different upconverter chains. Therefore, the additional degrees of freedom offered by the separate inputs can be used to optimize the performance on the same or larger BW, or to improve other performance metrics such as linearity and average efficiency. Ma *et al.* [8], for example, propose the use of machine-learning techniques to optimize the configuration parameters of a

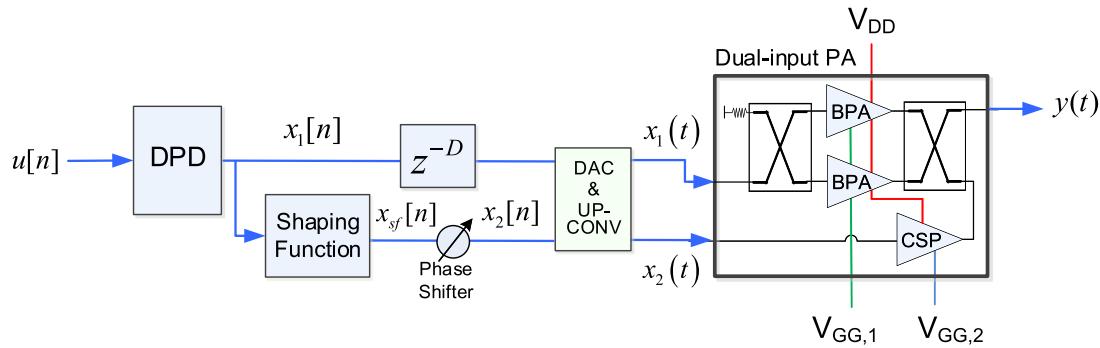


Fig. 1. Block diagram of the LMBA architecture with DPD linearization.

dual-input Doherty PA. Particularizing for dual-input LMBAs, the configuration of certain key free parameters influencing the inherent linearity versus power efficiency trade-off is explored in [9] and [10]. Among the different degrees of freedom that can be considered to optimize the LMBA performance, Guillena *et al.* [9] proposed a method for predicting the optimum relative phase shift between the two input-modulated signals to the LMBA that maximize linearity levels.

Some previous works published in literature addressing the linearization of LMBAs tested the PA with OFDM-based modulated signals but with moderate BWs (i.e., several tenths of megahertz), where the use of memory polynomial (MP), e.g., in [11] and [12], or generalized memory polynomial (GMP), e.g., in [13], behavioral models was enough to meet the out-of-band linearity specifications. To the best of the authors' knowledge, the linearization of LMBAs taking into account signal BWs of hundreds of megahertz (thus, taking advantage of the broadband nature of LMBAs), up to now has been only addressed in [14] and [15]. On the one hand, Cappello *et al.* [14] address the linearization of a supply-modulated LMBA when considering, among others, a NR-5G signal of 100-MHz instantaneous BW. By considering a GMP-based DPD, the out-of-band linearization specifications cannot be met for the 100-MHz BW signal. By using an ideal indirect learning control (sic), however, they show that the linearity specs could somehow be met. On the other hand, Pang *et al.* [15] presented the design and linearization of a LMBA that was tested with several OFDM-based signals taking into account BWs up to 200 MHz. With this last challenging BW configuration of ten-carrier 200-MHz OFDM signal with 10 dB of PAPR, however, the reported adjacent channel leakage ratio (ACLR) after DPD linearization could not reach the threshold of  $-45$  dBc using the magnitude-selective affine (MSA) function model for DPD. Moreover, none of the aforementioned papers addressing the linearization of LMBAs provide information of the error vector magnitude (EVM) to quantify the in-band distortion, which may be of concern when operating PAs with broadband signals.

In this article, a significant step forward is taken with respect to our previous work in [9], by proposing a methodology to ensure power efficient amplification from 1.8 up to 2.4 GHz allowing reconfigurability to meet the linearity

specifications in a dynamic environment, where the center frequency and BW of the transmitted signal change. Consequently, in this dynamic environment, the adaptive DPD linearizer can be reconfigured by selecting different behavioral models according to the BW of the signal. For example, when modeling strong nonlinearities with significant memory effects, the artificial neural network (ANN) DPD [16]–[18] can provide robust global estimation capabilities in contrast to the more local estimations provided by the polynomial-based DPD. For signal BWs of hundreds of megahertz, we show in this article that, given the difficulty of meeting the out-of-band linearity specifications when considering polynomial-based behavioral models, the use of ANNs for DPD linearization is justified. Therefore, unlike previously reported solutions, the ANN-based DPD proposed in this article is capable of meeting the ACPR specs (i.e.,  $<-45$  dBc) with EVM figures lower than 1% when considering four noncontiguous LTE-20 signals of 200-MHz total BW.

Accordingly, the remainder of this article is organized as follows. Section II presents a brief description of the LMBA used in this article. Section III describes the proposed methodology to properly configure the free parameters involved in the LMBA configuration to maximize linearity and power efficiency. Section IV describes the DPD linearization strategy followed, where both GMP- and ANN-based DPDs are used. Details on the ANN configuration and the low-complexity adaptation strategy following a direct learning approach are also discussed. Section V describes the experimental test bench and shows experimental results including DPD linearization when considering a dynamic transmission environment (i.e., considering different center frequencies and signal BWs of the transmitted signal over a frequency band ranging from 1.8 up to 2.4 GHz). Finally, the conclusion is given in Section VI.

## II. LMBA

In this article, the LMBA presented in [5] is used as device under test (DUT). A simplified block diagram of the DUT and the DPD linearizer is shown in Fig. 1. There are two separate RF inputs;  $x_1$  controls the balanced power amplifier (BPA) pair, based on two CGH40025F transistors from Wolfspeed, biased in class AB with  $V_{GG,1}$  at  $-2.8$  V corresponding to  $80$  mA of quiescent drain current;  $x_2$  controls the control signal power (CSP) amplifier, also based on a

CGH40025F, and biased in class C, with  $V_{GG,2}$  left as a free parameter within the range of DC voltages  $-3.5$  to  $-5.5$  V. The matching networks and the output hybrid couplers are based on soft-board microstrip networks, with SMD capacitors and resistors for the bypass and stabilization networks. An off-the-shelf hybrid is used on the input. The circuit is mounted on an aluminum fixture, and SMA coaxial launchers are used for the RF ports.

The CW measurements reported in [5] showed, over the 1.7–2.5-GHz frequency range, a maximum power larger than 63 W, and an 8-dB back-off efficiency exceeding 39%. Modulated signal measurements were also performed with 5- and 20-MHz channel LTE signals, showing the linearizability of the LMBA under these conditions. Both sets of measurements were performed with a manual search for the optimum amplitude, phase, and bias settings. In particular, the relative phase was maintained at a constant offset that led to a good compromise between output power and back-off efficiency, whereas the relative amplitude was following a square relation between the BPA and CSP inputs [5].

It is important to stress out that, when considering signals with BWs of several tenths or even hundreds of megahertz with PAPRs exceeding 8 dB (e.g., typical PAPRs of carrier aggregated signals around 10 dB or higher), the average power efficiency shown by this particular LMBA decreases quite abruptly. However, the tuning and linearization methodology proposed in this article is valid for addressing the inherent linearity versus power efficiency trade-off, independently on the specific power efficiency profile shown by the specific LMBA DUT.

The methodology proposed in this article to cope with the LMBA linearity versus efficiency trade-off, when considering a dynamic scenario in which the transmitted signal configuration can change, in terms of BW and frequency of operation in the range of 1.8–2.4 GHz, is depicted in the flowchart of Fig. 2. The first part of this flowchart corresponds to the tuning of the specific free parameters of the LMBA (e.g., the phase shift or the amplitude relationship between the main and the auxiliary signal) that have an impact on its linearity and power efficiency. As will be discussed in Section III, the optimal value of these parameters depends on the signal BW and frequency of operation. Once these parameters are properly tuned, in the second part of the flowchart (see Fig. 2), crest factor reduction (CFR) and adaptive DPD linearization techniques are applied to meet the required linearity specifications (i.e., ACPR <  $-45$  dB) with the best possible power efficiency. Details on the architecture and adaptation process of the proposed DPD linearizers and the criteria to choose one over the other will be discussed in Section IV.

### III. CONFIGURATION OF THE LMBA FREE PARAMETERS

Some of the free parameters of the dual-input LMBA that can be tuned to trade-off linearity and power efficiency are described in the following. Taking into account the notation in the block diagram of Fig. 1, the complex BPA signal is defined as  $x_1[n] = x[n]$ , where  $x[n]$  is the signal at the output of the DPD. The complex CSP signal  $x_2[n]$  is generated using

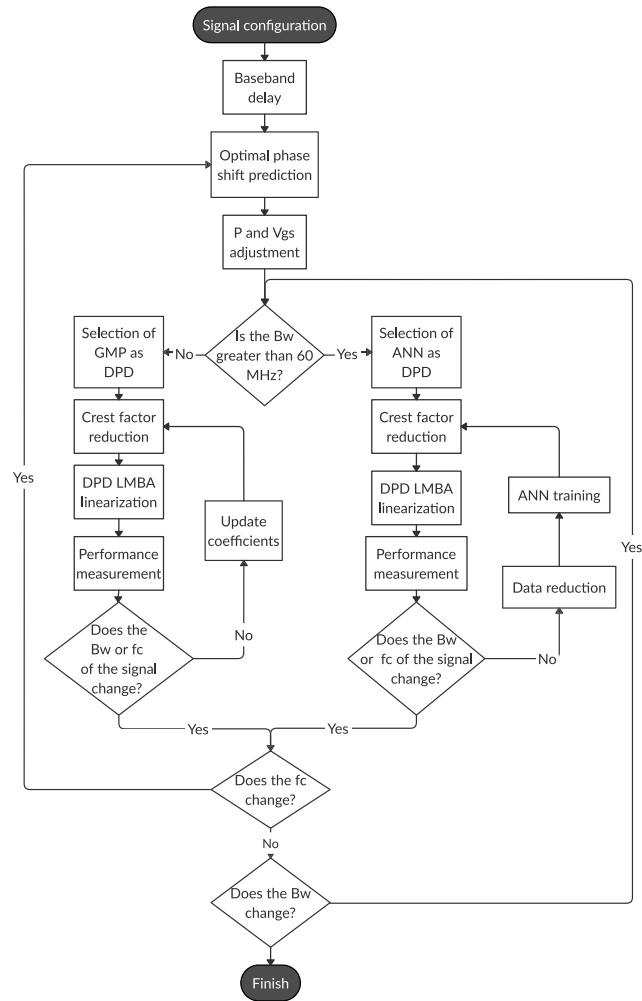


Fig. 2. Flowchart for the LMBA linearization and power efficiency enhancement.

a shaping function that controls the relative amplitude between the BPA and CSP inputs. More specifically, the CSP signal  $x_2[n]$  is defined as

$$x_2[n] = x_{sf}[n]e^{i\Psi_{\text{rel}}} \quad (1)$$

where  $\Psi_{\text{rel}}$  is the relative phase in radians between the BPA and CSP signals; and where the signal after the shaping function  $x_{sf}[n]$  is defined as

$$x_{sf}[n] = A_s[n]K_0e^{i\phi_x} \quad (2)$$

where  $K_0 = (\max\{|x[n]|}/\max\{|A_s[n]|})$ ,  $\phi_x = \text{phase}\{|x[n]|}$  and the amplitude relationship between the two signals is given by the following expression, which is a simplified version of the shaping function used in [9], where  $x_{\min} = 0$ , and thus

$$A_s[n] = (|x[n]|^6)^{\frac{1}{p}} \quad (3)$$

with  $p$  being the degree of the root.

To determine the optimal value of some free parameters such as the relative phase ( $\Psi_{\text{rel}}$ ) or delay between the BPA and CSP signals, or the degree of the root  $p$ , a 20-MHz BW LTE signal (LTE-20) with a PAPR of 10.2 dB was used to

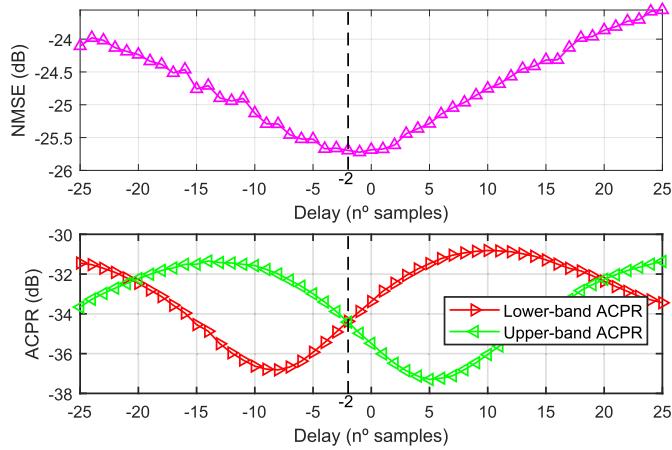


Fig. 3. NMSE, lower-band, and upper-band ACPR for different delay samples between LMBA input signals.

evaluate the linearity and power efficiency when tuning these parameters. In addition, another free parameter to be tuned is the gate voltage of the CSP amplifier,  $V_{GG,2}$ , that can be set to operate between a deep-class C condition that should favor efficiency, and a near-class B bias where linearity should improve.

#### A. Selection of the Baseband Delay

In a first step toward the optimal tuning, the objective is to find a baseband time delay between the LMBA inputs to balance the lower-band and upper-band ACPR of the LTE-20 test signal at the output of the LMBA. Fig. 3 shows the NMSE and ACPR of the amplified output for different delay samples between the LMBA inputs when considering the LTE-20 signal at 2-GHz center frequency, the optimum phase shift in terms of linearity and  $p = 3$ . As observed, the best NMSE values are for delays close to 0. However, without any delay, there is an ACPR difference of about 3 dB between the lower and upper bands. Therefore, for the sake of linearizability, a baseband delay of  $-2$  samples is chosen to balance the ACPR in the lower and upper bands.

#### B. Estimation of the Polynomial Fitting to Predict the Optimal Phase Shift

As introduced in [9], the phase shift between the LMBA's main and auxiliary signals has a strong impact on its linearity. Consequently, selecting an optimal phase shift is of crucial importance, because with certain phase-shift configurations it is not possible to meet the linearity requirements even by applying DPD linearization. In addition, the optimal phase shift depends on the specific frequency of operation.

Considering an LTE-20 test signal and fixing the degree of the root in (3) to  $p = 3$ , the effect on linearity of the phase shift between the LMBA's inputs for different center frequencies is shown in Figs. 4 and 5 in terms of NMSE and ACPR, respectively. However, as shown in Fig. 6, the power efficiency, despite having a strong dependence with the center frequency of operation (with its maximum value around 2 GHz), it is quite invariant with the phase shift between the LMBA inputs.

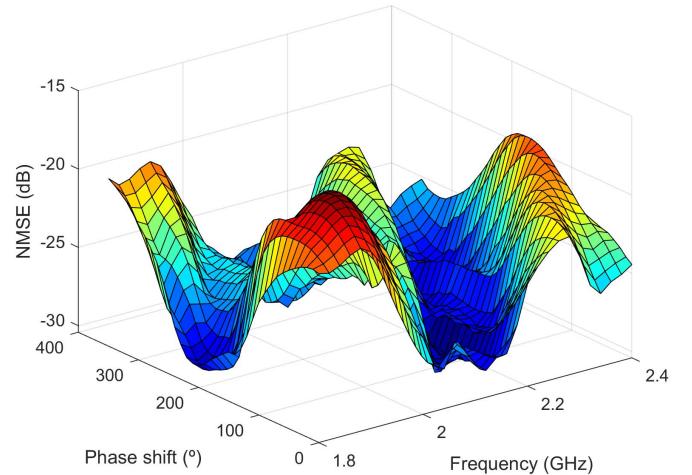


Fig. 4. NMSE value for different center frequencies and different phase shifts between the LMBA's input signals.

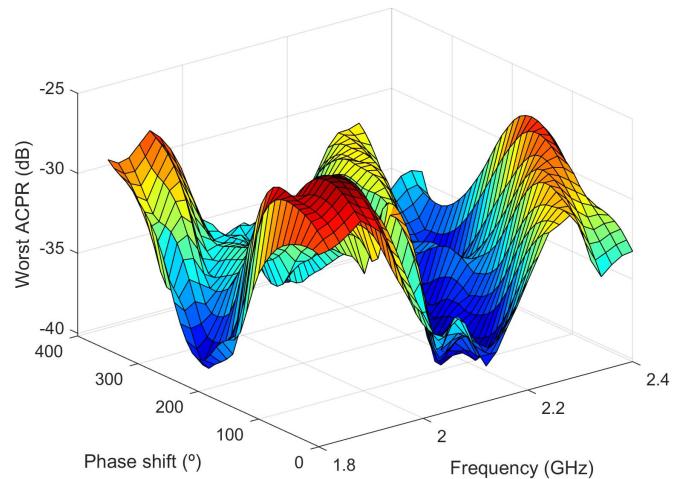


Fig. 5. Worst ACPR for different center frequencies and different phase shifts between the LMBA's input signals.

In terms of ACPR and NMSE, this optimal phase shift has a trend that can be predicted using a polynomial regression, as reported in [9]. In this article, we have considered polynomial regressions of degrees 1, 3, and 5 and a piecewise regression of degree 1. Fig. 7 shows the measured upper and lower limits of the NMSE and ACPR over frequency when considering for each center frequency the worst and best phase shift, respectively. In addition, NMSE and ACPR values obtained with the predicted optimal phase shifts using the different regression strategies are also shown. With all of them, the predicted phase shift resulted in NMSE and ACPR values close to the lower bound for all the frequency range. However, the best prediction was obtained with a simple linear regression or the piecewise functions. From now on, these two regression approaches will be used to predict the optimal phase shift in terms of linearity.

The polynomial regressions were extracted from data measurements of an LTE-20 signal. To validate that the optimal phase-shift predictions are valid when considering other

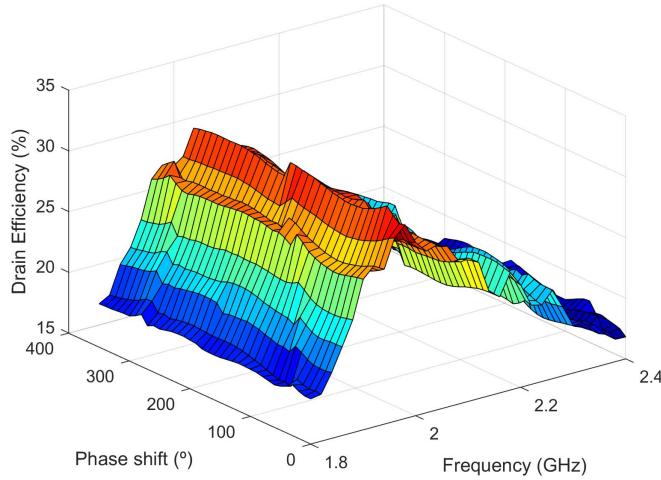


Fig. 6. Drain efficiency for different center frequencies and different phase shifts between the LMBA's input signals when considering an LTE-20 test signal with 10.2 dB of PAPR.

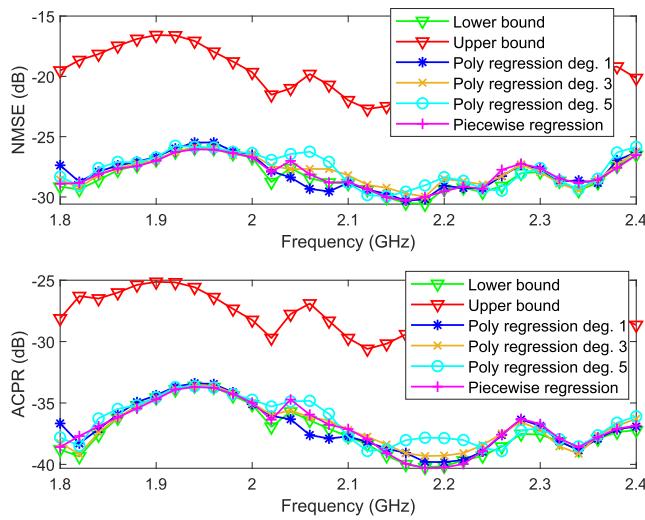


Fig. 7. NMSE and ACPR values obtained considering the predicted phase shifts taking into account different fitting approaches.

signals with different BWs, Figs. 8 and 9 show the results for LTE signals with total BWs of 60 and 200 MHz, respectively (further details on these test signals are given in Section V). As observed, the NMSE and ACPR values obtained with the predicted optimal phase shifts are very close to the optimal values (i.e., the lower bound) found from measurements into the entire frequency range.

### C. Selection of the Optimal Ranges of $p$ and $V_{GG,2}$

The last step proposed to configure the LMBA consists of a fine tuning of the parameter  $p$  in (3) and the auxiliary amplifier supply voltage,  $V_{GG,2}$  (see Fig. 1).

As reported in [9], for values higher than  $p = 4$  the power efficiency starts dropping significantly, while there is a sweet spot for linearity, in terms of NMSE and ACPR, around the value of  $p = 5$ . With these results we can make

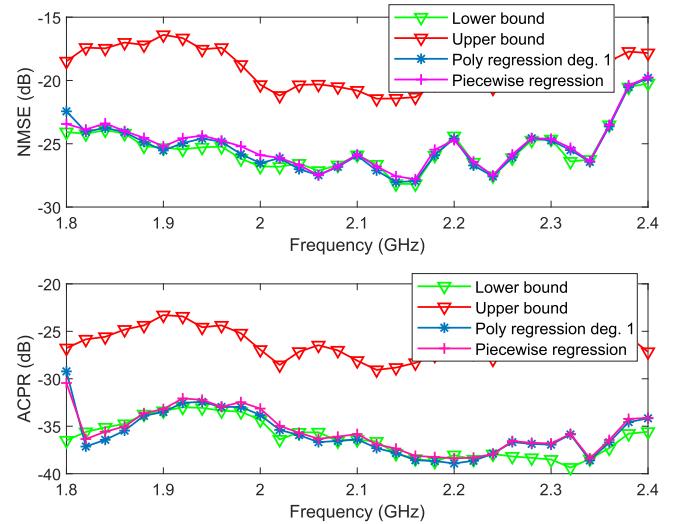


Fig. 8. NMSE and ACPR values for a NC 2×LTE-20 test signal of 60-MHz total BW.

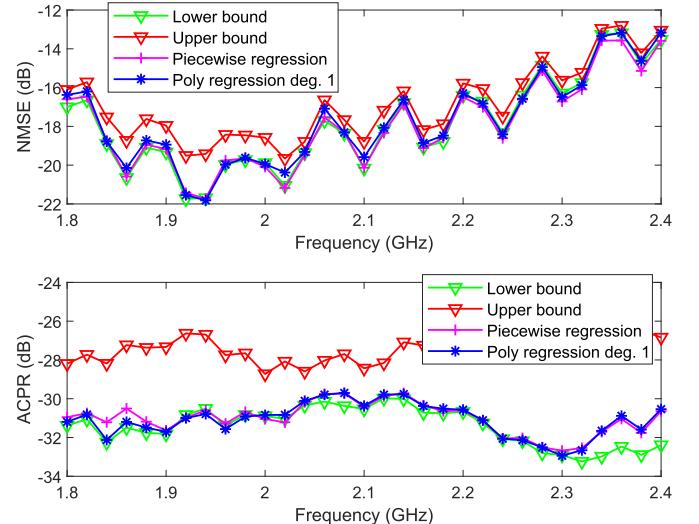


Fig. 9. NMSE and ACPR values for a NC 4×LTE-20 test signal of 200-MHz total BW.

a fine adjustment of this parameter setting the value of  $p$  between 3 and 5.

To analyze the effect of the auxiliary PA gate voltage, we carried out a sweep of  $V_{GG,2}$  exciting the LMBA with an LTE-20 signal at 2 GHz, with an optimal phase shift of  $\Psi_{rel} = 260^\circ$  and  $p = 3$ . Fig. 10 shows how the NMSE and ACPR values get better for  $V_{GG,2}$  between  $-3.5$  and  $-4$  V, while slightly higher power efficiency (the maximum variation of power efficiency is only 0.5 percentage points) is observed for  $V_{GG,2}$  between  $-3.9$  and  $-4.9$  V.

## IV. DPD LINEARIZATION

Once the LMBA is properly tuned, the most suitable (in terms of linearization performance and computational complexity) linearization method is selected, as depicted in the flowchart of Fig. 2. In Sections IV-A and IV-B, we will

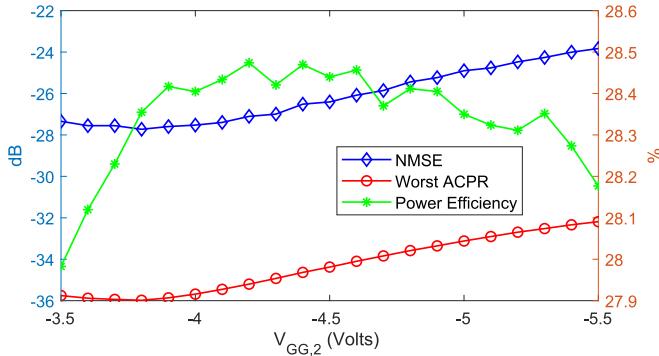


Fig. 10. LMBA linearity and power efficiency for different values of  $V_{GG,2}$ , considering an LTE-20 signal at 2 GHz.

describe both the polynomial- and the ANN-based DPD, providing specific details on the adaptation required to ensure reconfigurability to cope with dynamic environments.

#### A. Polynomial-Based Adaptive DPD

A common approach in literature for DPD is the use of behavioral models in the form of simplified versions of the full Volterra series. The GMP is a popular candidate because it introduces bidimensional kernels (considering cross-term products between the complex signal and the lagging and leading envelope terms), which increase the capability to compensate for the PA memory effects. Following the notation of the block diagram in Fig. 11, the input–output relationship at the DPD block is defined as

$$x[n] = u[n] - d[n] \quad (4)$$

where  $x[n]$  is the signal at the output of the DPD block,  $u[n]$  is the input signal, and  $d[n]$  is an error signal that can be described following a generic GMP structure as

$$\begin{aligned} d(n) = & \sum_{l=0}^{L_a-1} u(n-l)\varphi_l^a(|u(n-l)|) \\ & + \sum_{l=0}^{L_b-1} \sum_{m=1}^{M_b} u(n-l)\varphi_{l,m}^b(|u(n-l-m)|) \\ & + \sum_{l=0}^{L_c-1} \sum_{m=1}^{M_c} u(n-l)\varphi_{l,m}^c(|u(n-l+m)|) \end{aligned} \quad (5)$$

where  $\varphi(\cdot)$  are generic nonlinear functions that depend on envelope terms. These nonlinear functions can be described by polynomials as in the case of the original GMP [19] or by B-splines as presented in [9]. In any case, the GMP is a parametric model that can be expressed as the linear combination of nonlinear basis functions weighted by some parameters. In general, (4) can be rewritten in a matrix notation as

$$\mathbf{x} = \mathbf{u} - \mathbf{U}\mathbf{w} \quad (6)$$

where  $\mathbf{U}$  is the data matrix containing the DPD basis functions and  $\mathbf{w}$  is the vector of parameters. By following a direct

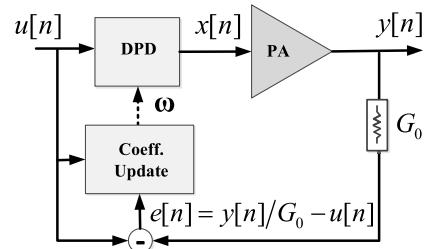


Fig. 11. Adaptive DPD architecture following a direct learning approach.

learning approach (see Fig. 11), the vector of parameters can be estimated iteratively as follows:

$$\mathbf{w}^{j+1} = \mathbf{w}^j + \mu^j (\mathbf{U}^H \mathbf{U})^{-1} \mathbf{U}^H \mathbf{e} \quad (7)$$

where  $\mu^j$  is the learning ratio at iteration  $j$  and  $\mathbf{e}$  is the DPD error vector defined as

$$\mathbf{e} = \frac{\mathbf{y}}{G_0} - \mathbf{u} \quad (8)$$

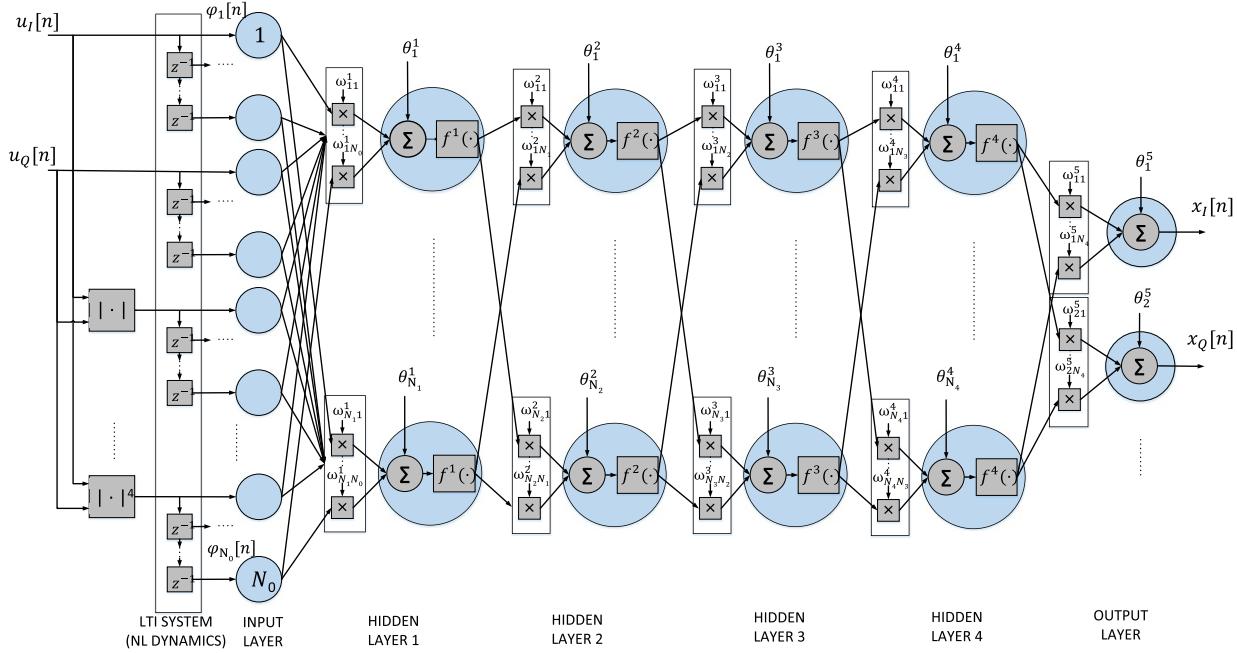
where  $G_0$  determines the desired linear gain of the PA and  $\mathbf{y}$  and  $\mathbf{u}$  are the PA output and transmitted input data vectors, respectively.

#### B. Adaptive DPD Based on ANNs

The most commonly used ANN architecture for DPD linearization is the feedforward time-delayed neural network (FTDNN). To handle complex data, real-valued (RV) FTDNN are used taking as inputs the in-phase ( $I$ ) and quadrature ( $Q$ ) components of the complex signal. Moreover, additional terms, such as envelope or phase terms, are included as inputs to the RV-FTDNN [16], [18], [20], to improve its linearization performance.

To select the different hyperparameters of our ANN (such as the architecture of the network, the input variables, the number of hidden layers, and the number of neurons per layer), we ran some preliminary tests to model the LMBA nonlinear behavior when excited with a challenging signal composed by four noncontiguous LTE-20 channels over a total BW of 200 MHz. As a result, Fig. 12 shows the proposed RV-FTDNN network, composed of four hidden layers (i.e.,  $N_{HL} = 4$ ) and a distribution of 20, 20, 10, and 10 neurons per hidden layer, respectively. The input values correspond to the  $I$  and  $Q$  components of the signal (i.e.,  $u_I[n]$  and  $u_Q[n]$  in Fig. 12), including time-delayed components up to a certain memory depth (e.g., seven delay taps) and envelope dependent terms (i.e.,  $|u[n]|^p$  with  $p = 1, \dots, 4$ ) including also their time-delayed values. A total of  $N_0 = 48$  inputs to the ANN were considered. After running several tests evaluating the performance obtained with different activation functions in the hidden layers of the ANN, the sigmoid tangent function was selected. It is mathematically equivalent to the hyperbolic tangent function but its MATLAB implementation runs faster. Finally, a linear function was selected in the output layer.

Taking into account the RV-FTDNN depicted in Fig. 12 and considering an input layer with  $N_0$  inputs, the input–output

Fig. 12. RV-FTDNN network composed of four hidden layers and a distribution of  $[N_1, N_2, N_3, N_4] = [20, 20, 10, 10]$  neurons per hidden layer.

relationship for both  $I$  and  $Q$  components is defined in the following equations:

$$x_I[n] = \sum_{s=1}^{N_4} \omega_{1,s}^5 f^4 \left( \sum_{k=1}^{N_3} \omega_{s,k}^4 f^3 \left( \sum_{j=1}^{N_2} \omega_{k,j}^3 f^2 \left( \sum_{l=1}^{N_1} \omega_{j,l}^2 f^1 \left( \sum_{i=1}^{N_0} \omega_{l,i}^1 \varphi_i[n] + \sigma_l^1 \right) + \sigma_j^2 \right) + \sigma_k^3 \right) + \sigma_s^4 \right) + \sigma_1^5 \quad (9)$$

$$x_Q[n] = \sum_{s=1}^{N_4} \omega_{2,s}^5 f^4 \left( \sum_{k=1}^{N_3} \omega_{s,k}^4 f^3 \left( \sum_{j=1}^{N_2} \omega_{k,j}^3 f^2 \left( \sum_{l=1}^{N_1} \omega_{j,l}^2 f^1 \left( \sum_{i=1}^{N_0} \omega_{l,i}^1 \varphi_i[n] + \sigma_l^1 \right) + \sigma_j^2 \right) + \sigma_k^3 \right) + \sigma_s^4 \right) + \sigma_2^5 \quad (10)$$

where  $\varphi_i[n]$  with  $(i = 1, \dots, N_0)$  are the inputs to the ANN (e.g.,  $\varphi_1[n] = u_I[n]$ ,  $\varphi_2[n] = u_I[n - 1]$ , etc.),  $\omega_{a,b}^m$  (with  $m = 1, \dots, N_{\text{HL}} + 1$ ) are synaptic weights of the network, the values of  $\sigma_a^m$  are the bias,  $f^r(\cdot)$  (with  $r = 1, \dots, N_{\text{HL}}$ ) are the activation functions, and  $N_r$  are the number of neurons per hidden layer. As explained before, in our particular ANN for DPD purposes  $N_{\text{HL}} = 4$  and  $[N_1, N_2, N_3, N_4] = [20, 20, 10, 10]$ . The number of coefficients of the proposed RV-FTDNN results from the sum of the number of weights and the number of biases, and thus  $O = (N_0 N_1 + N_1 N_2 + N_2 N_3 + N_3 N_4 + 2 N_4) + (N_1 + N_2 + N_3 + N_4 + 2)$ .

ANN-based DPD can be applied adaptively or nonadaptively (i.e., in open-loop). If a nonadaptive scheme is chosen, the ANN coefficients are configured through prior training and remain fixed over time. However, if an adaptive implementation is considered, the coefficients of the ANN are updated by

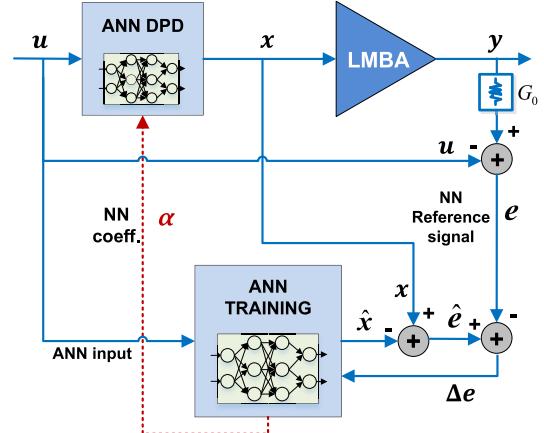


Fig. 13. Adaptive ANN-based DPD following a direct-learning approach.

retraining the network over time. This implementation allows us to adjust the ANN DPD to time-varying LMBA behavior or to a dynamic environment where the transmitted signal characteristics change over time. For example, in systems where a general-purpose DPD has to cope with different signal BWs at different center frequencies or operating the PA with different output back-off levels, some reconfigurability of the DPD parameters is required. Therefore, although we start from an initial condition of the DPD parameters obtained in an off-line training, if we want the best DPD linearization performance for specific operating conditions (i.e., for a given signal input power, BW, center frequency, etc.), coefficients adaptation is required.

The block diagram in Fig. 13 shows the proposed adaptive ANN DPD implementation following a direct learning approach. The Levenberg–Marquardt (LM) backpropagation

algorithm is used to calculate the RV-FTDNN coefficients by minimizing the mean square error (MSE) cost function  $C$  in (11) for each training data batch of length  $K$  samples. This forward–backward process is repeated until the desired modeling performance is met or the ANN fails in the validation procedure [21]

$$\begin{aligned} C &= \frac{1}{2K} \sum_{n=1}^K (\hat{e}_I[n] - e_I[n])^2 + (\hat{e}_Q[n] - e_Q[n])^2 \\ &= \frac{1}{2K} \sum_{n=1}^K |\Delta e[n]|^2 \end{aligned} \quad (11)$$

where  $K$  is the data batch length and  $\Delta e$  is the data batch error vector (see Fig. 13) of  $K$  samples defined as

$$\Delta e = \hat{e} - e \quad (12)$$

with  $e$  being the residual linearization error vector defined in (8) and  $\hat{e}$  being the estimated residual linearization error defined as follows:

$$\hat{e} = x - \hat{x} \quad (13)$$

where  $x[n] = x_I[n] + jx_Q[n]$  is the predistorted output signal and  $\hat{x}[n] = \hat{x}_I[n] + j\hat{x}_Q[n]$  is the output signal ( $I/Q$  pairs) produced at the output layer of the training ANN, as depicted in Fig. 13. The cost function  $C$  is minimized according to the LM algorithm and with respect to the vector of coefficients  $\alpha = [w_{11}^1 \dots w_{N_1 N_0}^1 \theta_1^1 \dots \theta_{N_1}^1 \dots w_{11}^5 \dots w_{2N_4}^5 \theta_1^5 \theta_2^5]^T$  containing the weights and biases of the RV-FTDNN. When going backwards,  $\alpha$  is updated at every epoch  $j$  as

$$\alpha^{j+1} = \alpha^j - (\mathbf{J}^T \mathbf{J} + \lambda \mathbf{I})^{-1} \mathbf{J}^T \Delta e \quad (14)$$

where  $\mathbf{I}$  is the identity matrix,  $\lambda$  is the damping factor, and  $\mathbf{J}$  is the Jacobian matrix being calculated over the error vector  $\Delta e$  with respect to  $\alpha$  as

$$\mathbf{J} = \begin{pmatrix} \frac{\partial \Delta e[1]}{\partial w_{11}^1} & \frac{\partial \Delta e[1]}{\partial w_{12}^1} & \dots & \frac{\partial \Delta e[1]}{\partial \theta_1^5} & \frac{\partial \Delta e[1]}{\partial \theta_2^5} \\ \frac{\partial \Delta e[2]}{\partial w_{11}^1} & \frac{\partial \Delta e[2]}{\partial w_{12}^1} & \dots & \frac{\partial \Delta e[2]}{\partial \theta_1^5} & \frac{\partial \Delta e[2]}{\partial \theta_2^5} \\ \vdots & & & & \\ \frac{\partial \Delta e[K]}{\partial w_{11}^1} & \frac{\partial \Delta e[K]}{\partial w_{12}^1} & \dots & \frac{\partial \Delta e[K]}{\partial \theta_1^5} & \frac{\partial \Delta e[K]}{\partial \theta_2^5} \end{pmatrix}. \quad (15)$$

Fig. 14 presents a comparison of the linearization performance (in terms of ACPR) versus number of averages of the measured PA output signal, when considering adaptation and open-loop ANN DPD. The ACPR results shown with the adaptive DPD were obtained after three update iterations (independently of the number of averages considered) from the initial configuration of parameters. This initial parameter estimation was carried out in a preliminary off-line training and corresponds to the parameter configuration that is permanently used in the open-loop (nonadaptive) DPD. As it will be further described in Section V, the test signal used consisted of four noncontiguous LTE-20 channels over a total BW of 200 MHz. It is well-known that, considering the test bench described in Section V, by averaging several captures

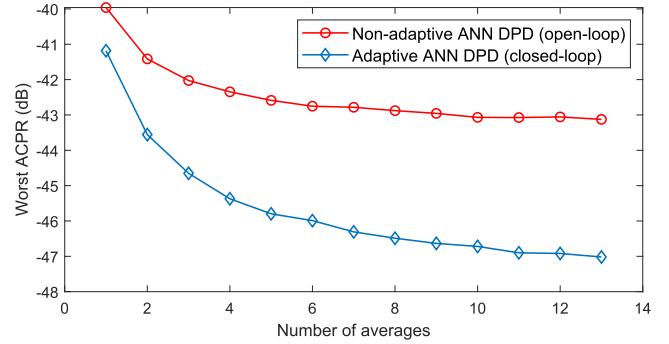


Fig. 14. ACPR values for an adaptive and nonadaptive implementation of the ANN DPD for different number of averages.

of the PA output signal we can reduce the noise floor and thus enhance the DPD linearization capabilities. Consequently, a straightforward conclusion derived from the results depicted in Fig. 14 is that adaptive DPD, with proper averaging, can outperform the spectral regrowth compensation obtained by the open-loop ANN DPD by more than 4 dB.

One of the main problems of the adaptive ANN DPD, besides the computational complexity, is the required training time for each update. To reduce the retraining time between adaptations we have followed two approaches: limiting the number of ANN-training epochs and reducing the number of training samples. Several techniques have been proposed in literature to select the most relevant training data allowing a reliable estimation without significant loss of performance [18], [22]. In this article we have used the mesh-selecting method proposed in [22].

## V. EXPERIMENTAL TEST BENCH AND RESULTS

### A. Experimental Test Bench

The dual-input PA system was experimentally evaluated using a MATLAB-controlled digital linearization test bench, as shown in Fig. 15, interfacing waveform generation and acquisition instruments. To compensate for the out-of-band distortion, a 614.4-MSa/s DPD signal was digitally upconverted to the 2-GHz RF frequency, and digital to analog converted (through the AWG M8190A from Keysight, with a clock rate of 7.9872 GHz and 14 bits) to feed the dual-input PA. The PA output signal was attenuated, RF sampled with the digital storage oscilloscope (DSO) Keysight 90404A at 20 GSa/s with 8-bit resolution (applying averages to reduce the noise floor), digital downconverted, and resampled for time alignment and DPD processing. A Keysight N9020A MXA signal analyzer was used to characterize the spectrum at the output of the PA.

The dataset used for offline training the ANN consisted of 307 200 complex-valued data samples. The estimated parameters of the ANN were later validated and eventually adapted (in a different time-scale than real-time) in closed-loop DPD using different batches of 307 200 complex-valued data samples at each iteration. In the case of the GMP-based DPD, no offline training was applied *a priori*. Instead, the coefficients were directly adapted from scratch in the

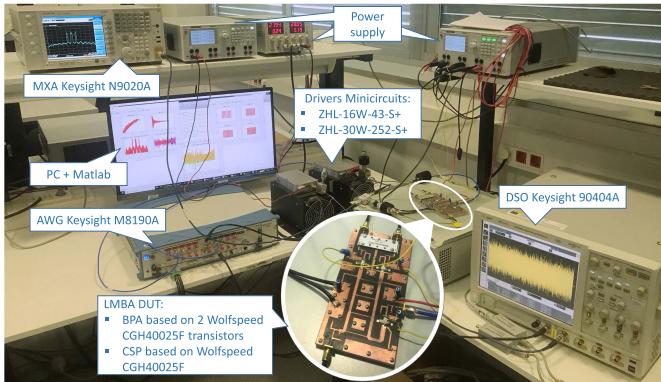


Fig. 15. Laboratory test bench including the LMBA used as DUT.

DPD observation loop considering also different batches of 307 200 complex-valued data samples at each iteration.

### B. GMP Versus ANN DPD for Different Signal BWs at 2-GHz RF Center Frequency

In a first approach, we wanted to evaluate when (or if) it is worth using ANNs instead of simplified Volterra series behavioral models, such as the GMP. Even though ANNs can outperform the linearization capabilities of polynomial-based behavioral models, the price to pay is an increase of computational complexity, for example, stated in terms of number of parameters. Therefore, with this objective in mind, the LMBA was excited with different test signals in order to evaluate linearization performance of both ANN-based and GMP-based DPD linearizers. The test signals' characteristics as well as the LMBA-delivered output power and power efficiency achieved with these signals are described in the following.

- 1) A 64-QAM LTE-20 signal of 20-MHz total BW, centered at 2 GHz and with 10.2 dB of PAPR. With this test signal the LMBA was operated to deliver around 35.5-dBm mean output power with a power efficiency around 28%.
- 2) Two noncontiguous 64-QAM LTE-20 (NC 2× LTE-20) signals of 60-MHz total BW, centered at 2 GHz and with 10.4 dB of PAPR. With this test signal the LMBA was operated to deliver around 35.5-dBm mean output power with a power efficiency around 28%.
- 3) Four noncontiguous 64-QAM LTE-20 (NC 4× LTE-20) signal of 200-MHz total BW centered at 2 GHz and with 10.6 dB of PAPR. With this test signal the LMBA was operated to deliver around 33.3-dBm mean output power with a power efficiency around 21.4%.

The initial configuration of parameters for both GMP and the ANN DPD behavioral models was rich enough to be able to cope with the worst-case scenario, i.e., the linearization of the NC 4×LTE-20 signal. Therefore, as listed in Table I, the same amount of coefficients were used to address the linearization of all three type of signals. Having an initial condition for the GMP parameters is not critical, because the parameters identification is relatively fast. However, in the

case of ANNs, the training time can be significantly long. Consequently, it is of crucial importance to speed up the adaptation process to have an accurate initial condition of the ANN parameters. For that reason, targeting an agile, versatile operation, the ANN was initially trained considering the NC 4× LTE-20 signal of 200-MHz total BW. Once the ANN is trained, it can be used as a DPD for different signals with different BW.

The training time for ANNs depends on several factors such as the hyperparameters of the ANN (number of hidden layers, neurons, etc.), the available hardware to carry out the computation (e.g., hardware accelerators such as GPUs or FPGAs), the software/library used (e.g., MATLAB, Pytorch), the type of solver used, and the batch size or the number of epochs. For the offline training of the ANN parameters (i.e., the initial condition), we used the LM algorithm provided by MATLAB without hardware accelerators, considering a batch size of 307 200 samples and without imposing limitations on the number of epochs. Consequently, the training time was around 6 h. However, once the initial condition for the ANN coefficients was estimated, in order to reduce the retraining time between adaptations, we followed two approaches: limiting the number of ANN training epochs and reducing the number of training samples using the mesh selecting method proposed in [22]. Then, the retraining time was reduced to a few minutes. In particular, between 2 and 5 min depending on the number of selected data samples to carry out the adaptation (there is a tradeoff between computational complexity and modeling accuracy).

Table I summarizes the linearization performance (in terms of ACPR, NMSE, and EVM) of both the ANN- and GMP-based DPD linearizers when considering the previously described test signals. The nonlinear functions in the general GMP behavioral model in (5) are particularized with polynomials. Therefore, the metaparameter configuration of the GMP DPD used in this article is nonlinear order of  $Pa = 7$ ,  $Pb = 5$ ,  $Pc = 5$  and memory depth of  $Ma = 9$ ,  $Mb = 7$ ,  $Mc = 7$ ,  $Qb = 2$ ,  $Qc = 2$ ; corresponding to a total of 248 coefficients. This configuration was selected to have two general-purpose DPD linearizers (i.e., the GMP-based and the ANN-based) with the same memory depth. Even if the GMP and ANN are not specifically dimensioned (i.e., use the minimum required number of coefficients to meet the linearity specs) to linearize the 20- or 60-MHz total BW signals, it is clear that GMP DPD can perfectly cope with the compensation of the nonlinear distortion and memory effects by meeting the targeted ACPR < -45 dB, with significantly less computational complexity than ANNs. Taking into account the aforementioned GMP configuration with 248 coefficients, the time for running the coefficients estimation (one iteration) was less than 10 s considering MATLAB's backslash operation.

However, when considering the NC 4× LTE-20 signal of 200-MHz total BW, the GMP DPD is not capable to meet the required ACPR levels. By increasing the number of coefficients of the GMP model, not only there is no improvement in the spectral regrowth compensation, but it also leads to an ill-conditioned and unreliable estimation

TABLE I  
COMPARISON OF ANN- AND GMP-BASED DPD FOR DIFFERENT TEST SIGNALS AT 2-GHz RF CENTER FREQUENCY

DPD type	Signal type	Worst ACPR	NMSE	EVM	Num. Coff.
No DPD	64-QAM LTE-20, BW=20 MHz	-35.0 dB	-26.4 dB	2.4 %	—
GMP		-51.0 dB	-39.3 dB	0.7 %	248
ANN (mem. depth =7)		-52.9 dB	-40.4 dB	0.6 %	1742
No DPD	NC 64-QAM 2×LTE-20, BW=60 MHz	-32.1 dB	-24.1 dB	2.7 %	—
GMP		-51.0 dB	-39.6 dB	0.6 %	248
ANN (mem. depth =7)		-49.9 dB	-39.2 dB	0.6 %	1742
No DPD	NC 64-QAM 4×LTE-20, BW=200 MHz	-28.4 dB	-19.8 dB	4.7 %	—
GMP		-35.4 dB	-28.0 dB	2.1 %	248
ANN (mem. depth =7)		-44.0 dB	-35.9 dB	0.7 %	1742
ANN (mem. depth =9)		-46.2 dB	-36.3 dB	0.9 %	1982

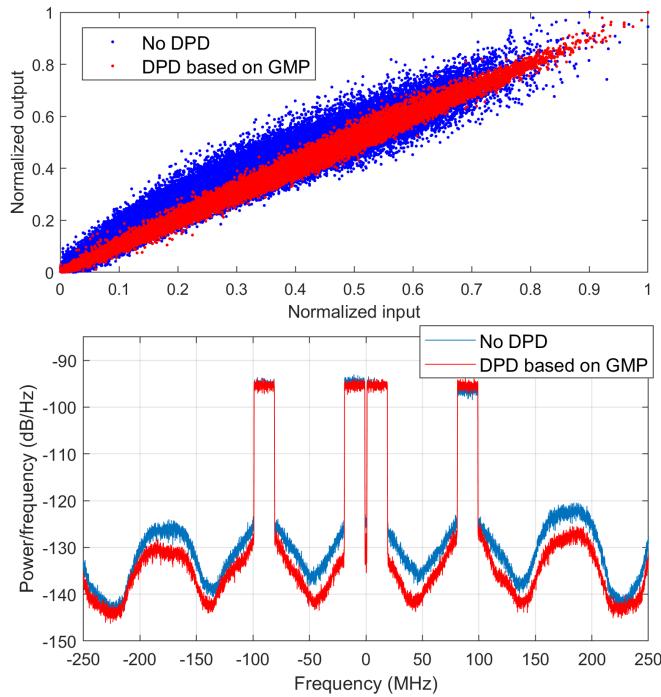


Fig. 16. AM-AM characteristics (top) and output power spectra (bottom) before and after GMP-based DPD linearization, considering a NC 4× LTE-20 signal at 2-GHz RF center frequency.

of the coefficients. In general, the bidimensional kernels of the GMP architecture represent a good trade-off between computational complexity and nonlinear dynamic modeling capabilities. But yet, the GMP is still a simplified version of the Volterra series covering a much more reduced area in the kernel space than the original Volterra series. Therefore, when dealing with wideband signals (e.g., hundreds of megahertz), the inherent limitation of the GMP architecture for covering certain cross-memory products can prevent meeting the required ACPR levels, regardless of whether regularization or feature selection techniques are applied. Consequently, in this particular demanding situation, the use of ANNs is justified. Note that the ANN with a memory depth of seven taps still cannot meet the targeted ACPR level, and thus, the memory depth has to be increased up to nine taps to go beyond the minimum  $-45$  dB of ACPR. Figs. 16 and 17 show the AM-AM characteristics and output power spectra before and after DPD linearization using GMP and ANN DPD, respectively. It is

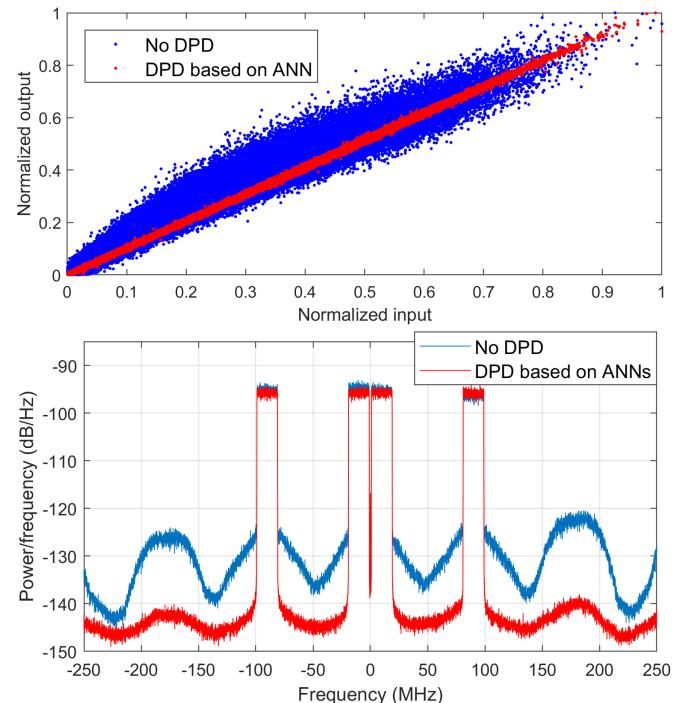


Fig. 17. AM-AM characteristics (top) and output power spectra (bottom) before and after ANN-based DPD linearization, considering a NC 4× LTE-20 signal at 2-GHz RF center frequency.

evident that the ANN DPD clearly outperforms the GMP DPD for wide-band signals where memory effects are critical, and the ANN can take advantage of the internal interactions that occur in its hidden layers.

### C. Linear and Power-Efficient LMBA Amplification From 1.8 to 2.4 GHz Considering Different Transmitted Signals

To guarantee linear and power-efficient amplification for dual-input PAs such as the DUT used in this work, we have proposed the methodology described in the flow diagram of Fig. 2. Therefore, assuming a dynamic environment where the transmitted signal to be amplified by the LMBA can present different BWs and operate at different center frequencies in the range of 1.8–2.4 GHz, we will proceed as follows (see Fig. 2).

- 1) First, assuming that the delay value has been previously determined and fixed, the optimum phase shift between

TABLE II  
LINEARITY AND POWER EFFICIENCY RESULTS OF THE LMBA OPERATED WITH TRANSMITTED SIGNALS WITH DIFFERENT BW CONFIGURATIONS IN THE RF RANGE OF 1.8–2.4 GHz

Signal configuration	DPD	Nº Coeff.	PAPR	Worst ACPR	NMSE	Pout	Efficiency	EVM
F <sub>c</sub> = 2.2 GHz, BW = 200 MHz	ANN	1982	9 dB (with CFR)	-46.1 dB	-32.9 dB	33.0 dBm	16.7 %	1.2 %
F <sub>c</sub> = 1.9 GHz, BW = 20 MHz	GMP	248	9 dB (with CFR)	-48.0 dB	-37.7 dB	38.3 dBm	38.5 %	0.8 %
F <sub>c</sub> = 2.1 GHz, BW = 60 MHz	GMP	248	10.2 dB	-51.2 dB	-38.3 dB	37.1 dBm	31.1 %	0.7 %
F <sub>c</sub> = 1.8 GHz, BW = 200 MHz	ANN	1982	8 dB (with CFR)	-46.1 dB	-27.2 dB	32.4 dBm	18.6 %	2.7 %

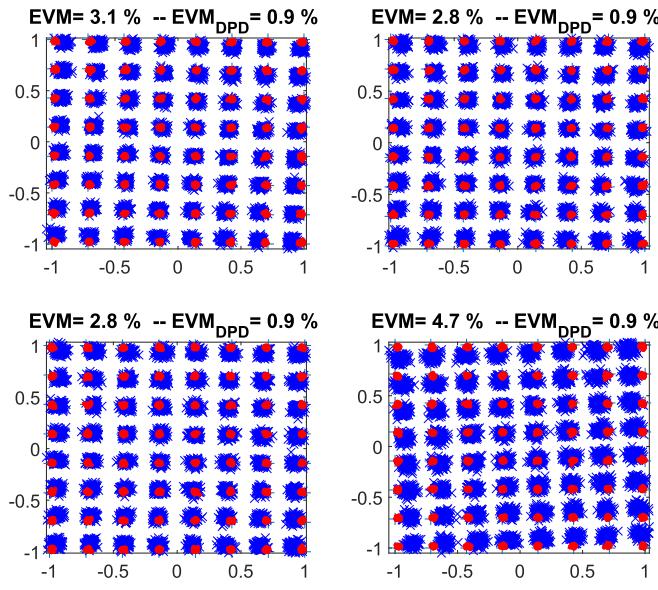


Fig. 18. 64-QAM constellations of the NC 4×LTE-20 signal at 2-GHz RF center frequency before and after ANN-based DPD linearization.

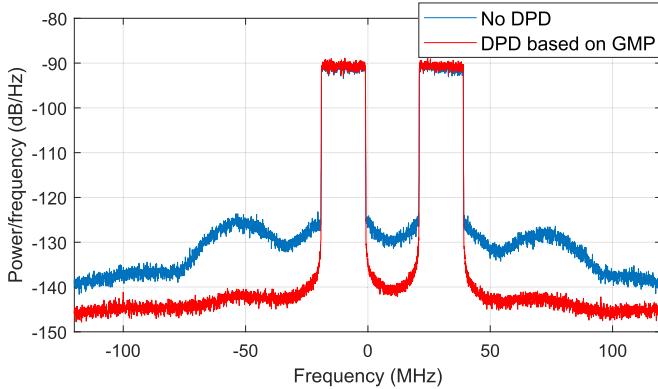


Fig. 19. Linearized and unlinearized output power spectra with GMP-based DPD considering a NC 2×LTE-20 signal at 2.1-GHz RF center frequency.

the BPA and CSP signals is calculated. For a given input signal with a certain BW and at a specific center frequency, we will use the extracted phase-shift model to determine the best phase shift between the LMBA input signals.

- 2) Then, it is necessary to determine the values of  $p$  and  $V_{GG,2}$  for trading-off linearity and power efficiency as described in Section III-C. For example, for the experimental results presented in Table II, we have considered  $p = 3.5$  and  $V_{GG,2} = -4.2$  V.

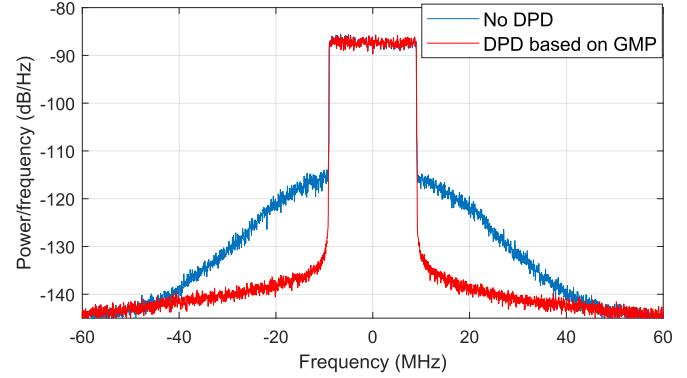


Fig. 20. Linearized and unlinearized output power spectra with GMP-based DPD considering an LTE-20 signal at 1.9-GHz RF center frequency.

- 3) A CFR technique is used to limit the PAPR of the signal. A well-known strategy commonly used in literature is to reduce the PAPR of the signal to be able to operate the PA with less back-off and thus more power efficiently. In this work, we have considered the peak cancellation CFR technique as described in [23].
- 4) Then, according to the transmitted signal characteristics we decide the DPD strategy. For example, in our particular test case, for signals exceeding 60-MHz total BW, the linearization is carried out by means of an ANN-based DPD. Otherwise, for signals with up to 60-MHz total BW, the less computationally intensive GMP DPD is considered. It is worth mentioning that the ANN-based DPD would be enough to linearize any type of the test signals considered, however, at the price of introducing additional computational complexity (with respect to GMP) when it is not strictly required.
- 5) Finally, from a given initial condition (obtained in an offline training), the adaptive DPD periodically updates its coefficients to meet the required linearity specifications. If the BW or the center frequency of the transmitted signal changes (for example, because other LTE-20 channels are aggregated), then the LMBA linearization system has to be reconfigured again, as schematically described in Fig. 2.

To validate the proposed methodology, Table II summarizes how the linearity specifications are met under different transmitted signal configurations in terms of BW and center frequency ( $F_c$ ) of operation. As an example, Figs. 19 and 20 show the spectra of the NC 2×LTE-20 signal of 60-MHz total BW and the LTE-20 signal, respectively, before and after

GMP-based DPD linearization. When limiting the maximum PAPR of the signal (by means of the peak cancellation CFR technique), it is possible to push the input signal harder into compression to slightly improve the overall drain efficiency, but at the price of introducing in-band distortion (as reflected in the NMSE and EVM figures in Table II). Notice that the LMBA linear mean output power and power efficiency change not only with the BW of the transmitted signal, but also with the center frequency of operation, as depicted in Fig. 7.

## VI. CONCLUSION

In this article, we propose a methodology to ensure linear amplification of a dual-input PA that can operate in a dynamic environment over a frequency range from 1.8 to 2.4 GHz. In a first step, some parameters defining the LMBA operation mode are properly tuned, taking into account the center frequency of operation to ensure linearizability with maximum power efficiency. In a second step, CFR and DPD linearization techniques are used to meet the required linearity specifications. An ANN-based DPD is previously trained offline to be able to cope with different signal BWs. By including adaptation, the ANN is able to meet the linearity specifications for any signal BW configuration. However, given the computational complexity introduced by the ANN, for less challenging scenarios (e.g., in our particular case, we considered signals with  $BW \leq 60$  MHz) a less computational complex GMP-based DPD is selected.

In the path toward allowing total reconfigurability of the amplification system, machine-learning techniques will be included in future works to allow fine tuning of the LMBA free parameters oriented at maximizing power efficiency for each frequency of operation and transmitted signal characteristics.

## REFERENCES

- [1] Z. Wang, "Demystifying envelope tracking: Use for high-efficiency power amplifiers for 4G and beyond," *IEEE Microw. Mag.*, vol. 16, no. 3, pp. 106–129, Apr. 2015.
- [2] V. Camarchia, M. Pirola, R. Quaglia, S. Jee, Y. Cho, and B. Kim, "The Doherty power amplifier: Review of recent solutions and trends," *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 2, pp. 559–571, Feb. 2015.
- [3] R. Pengelly, C. Fager, and M. Ozen, "Doherty's legacy: A history of the Doherty power amplifier from 1936 to the present day," *IEEE Microw. Mag.*, vol. 17, no. 2, pp. 41–58, Feb. 2016.
- [4] D. J. Shepphard, J. Powell, and S. C. Cripps, "An efficient broadband reconfigurable power amplifier using active load modulation," *IEEE Microw. Wireless Compon. Lett.*, vol. 26, no. 6, pp. 443–445, Jun. 2016.
- [5] R. Quaglia and S. Cripps, "A load modulated balanced amplifier for telecom applications," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 3, pp. 1328–1338, Mar. 2018.
- [6] Z. Popovic and J. A. Garcia, "Microwave class-E power amplifiers: A brief review of essential concepts in high-frequency class-E PAs and related circuits," *IEEE Microw. Mag.*, vol. 19, no. 5, pp. 54–66, Jul. 2018.
- [7] M. Litchfield and T. Cappello, "The various angles of outphasing PAs: Competitiveness of outphasing in efficient linear PA applications," *IEEE Microw. Mag.*, vol. 20, no. 4, pp. 135–145, Apr. 2019.
- [8] R. Ma *et al.*, "Machine-learning based digital Doherty power amplifier," in *Proc. IEEE Int. Symp. Radio-Freq. Integr. Technol. (RFIT)*, Melbourne, VIC, Australia, Aug. 2018, pp. 1–3.
- [9] E. Guillena, W. Li, P. L. Gilabert, and G. Montoro, "Prediction of the optimal phase shift between control signals in dual-input power amplifiers," in *Proc. Int. Workshop Integr. Nonlinear Microw. Millimetre-Wave Circuits (INMMiC)*, Cardiff, U.K., Jul. 2020, pp. 1–3.
- [10] K. Vivien, P. E. de Falco, G. Baudoine, O. Venard, P. P. C. Felix, and T. Barton, "Load modulated balanced amplifier designed for AM-PM linearity," in *Proc. 50th Eur. Microw. Conf. (EuMC)*, Utrecht, The Netherlands, Jan. 2021, pp. 304–307.
- [11] R. Quaglia, J. Powell, D. Shepphard, P. Tasker, and S. Cripps, "Analysis and characterization of a load modulated balanced amplifier for base-station applications," in *Proc. 11th German Microw. Conf. (GeMiC)*, Freiburg, Germany, Mar. 2018, pp. 1–4.
- [12] D. Collins, R. Quaglia, J. Powell, and S. Cripps, "Experimental characterization of a load modulated balanced amplifier with simplified input power splitter," in *Proc. Asia-Pacific Microw. Conf. (APMC)*, Kyoto, Japan, Nov. 2018, pp. 461–463.
- [13] T. Cappello, P. H. Pednekar, C. Florian, Z. Popovic, and T. W. Barton, "Supply modulation of a broadband load modulated balanced amplifier," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Philadelphia, PA, USA, Jun. 2018, pp. 304–307.
- [14] T. Cappello, P. Pednekar, C. Florian, S. Cripps, Z. Popovic, and T. W. Barton, "Supply- and load-modulated balanced amplifier for efficient broadband 5G base stations," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 7, pp. 3122–3133, Jul. 2019.
- [15] J. Pang *et al.*, "Analysis and design of highly efficient wideband RF-input sequential load modulated balanced power amplifier," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 5, pp. 1741–1753, May 2020.
- [16] D. Wang, M. Aziz, M. Helaoui, and F. M. Ghannouchi, "Augmented real-valued time-delay neural network for compensation of distortions and impairments in wireless transmitters," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 1, pp. 242–254, Jan. 2019.
- [17] M. Rawat, K. Rawat, and F. M. Ghannouchi, "Adaptive digital predistortion of wireless power amplifiers/transmitters using dynamic real-valued focused time-delay line neural networks," *IEEE Trans. Microw. Theory Techn.*, vol. 58, no. 1, pp. 95–104, Jan. 2010.
- [18] D. Lopez-Bueno, P. L. Gilabert, and G. Montoro, "Dataset reduction for neural network based digital predistorters under strong nonlinearities," in *Proc. IEEE Topical Conf. RF/Microw. Power Modeling Radio Wireless Appl. (PAWR)*, San Diego, CA, USA, Jan. 2021, pp. 1–4.
- [19] 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, Oct. 2006.
- [20] P. Jueschke and G. Fischer, "Machine learning using neural networks in digital signal processing for RF transceivers," in *Proc. IEEE AFRICON*, Cape Town, South Africa, Sep. 2017, pp. 384–390.
- [21] S. S. Haykin, *Neural Networks and Learning Machines*, 3rd ed. Upper Saddle River, NJ, USA: Pearson, 2009.
- [22] T. Wang and P. L. Gilabert, "Mesh-selecting for computational efficient PA behavioral modeling and DPD linearization," *IEEE Microw. Wireless Compon. Lett.*, vol. 31, no. 1, pp. 37–40, Jan. 2021.
- [23] D. Lopez, P. L. Gilabert, G. Montoro, and N. Bartzoudis, "Peak cancellation and digital predistortion of high-order QAM wideband signals for next generation wireless backhaul equipment," in *Proc. Int. Workshop Integr. Nonlinear Microw. Millimetre-Wave Circuits (INMMiC)*, Apr. 2014, pp. 1–3.



**Estefanía Guillena** (Graduate Student Member, IEEE) received the double bachelor's degree in aerospace systems engineering and telecommunications systems engineering from the Universitat Politècnica de Catalunya, Barcelona, Spain, in 2020.

She joined the Components and Systems for Communications Research Group in 2019 to develop her bachelor's thesis in the field of power amplifier linearization. She is currently working for Dares Technology, Barcelona, in the field of image processing as a Remote Sensing Engineer.

Ms. Guillena was awarded with the Microwave Theory and Techniques Society (MTT-S) Undergraduate Scholarship in April 2019.



**Wantao Li** (Graduate Student Member, IEEE) was born in Canton, China, in 1995. He received the B.E. degree in electronic information science and technology from Beijing Union University, Beijing, China, in 2017, and the M.S. degree in unmanned aircraft systems from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 2020, where he is currently pursuing the Ph.D. degree at the Department of Signal Theory and Communications. He joined the Components and Systems for Communications (CSC) Research Group in 2020. His research interests include signal processing for communication systems, power amplifier linearization, and the efficient RTL implementations in FPGAs.



**Roberto Quaglia** (Member, IEEE) was born in Casale Monferrato, Italy, in 1984. He graduated (*cum laude*) in electronic engineering from the Politecnico di Turin, Turin, Italy, in 2008, and he received the Ph.D. degree in electronic devices from the Politecnico di Turin.

He is currently a Lecturer at the School of Engineering, Cardiff University, Cardiff, U.K. His research interests concern the design, modeling, and predistortion of high-efficiency MMIC power amplifiers.

Dr. Quaglia is a member of the IEEE MTT-TC 12. He was a European Union Marie Skłodowska Curie Fellow in 2015 and a recipient of the 2009 Young Graduated Research Fellowship presented by the GAAS Association.



**Gabriel Montoro** (Member, IEEE) received the M.Sc. degree in telecommunication engineering and the Ph.D. degree from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 1990 and 1996, respectively.

He joined the Department of Signal Theory and Communications (TSC), UPC, in 1991, where he is currently an Associate Professor. His first research works were done on the area of adaptive control, and now his main research interest is in the use of signal-processing strategies for power efficiency improvement in communications systems.



**Pere L. Gilabert** (Senior Member, IEEE) received the M.Sc. degree in telecommunication engineering and the Ph.D. degree from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 2002 and 2008, respectively. He developed his master's thesis at the University of Rome "La Sapienza" with an Erasmus Exchange Grant.

He joined the Department of Signal Theory and Communications, UPC, in 2003. He is an Associate Professor at the Castelldefels School of Telecommunications and Aerospace Engineering. His research activity is in the field of linearization techniques and digital signal-processing solutions for highly efficient transmitter architectures.

Dr. Gilabert received the Extraordinary Doctoral Prize for his Ph.D. degree.