Transmitter Linearization Adaptable to Power-Varying Operation

Carlos Crespo-Cadenas, *Senior Member, IEEE,* María J. Madero-Ayora, *Member, IEEE,*
Javier Reina-Tosina, *Senior Member, IEEE,* and Juan A. Becerra-González, *Student Member, IEEE*

Abstract—This paper presents the design of a power-scalable digital predistorter (DPD) for transmitter architectures. The target is to accomplish the joint compensation of impairments due to the I/Q modulator and nonlinearities associated with the power amplifier (PA), and procure a maintained linearization performance in a range of average working operation levels. The identification method for the linearizer parameters enriches the standard least-squares procedure with a synergistic integration with sparsity-based model pruning strategies. The method has been tested with a general complex-valued Volterra model applied to the linearization of two communications transmitters operating at 3.6 GHz. The linearizers designed for the two transmitters effectively provide the joint compensation of the nonlinear behavior. In addition to their good performance in terms of adjacent channel power ratio, the DPDs exhibit a wide range of power-varying adaptation.

Index Terms—Behavioral modeling, Volterra series, power amplifiers, nonlinear model identification, digital predistortion.

I. Introduction

The evident interest of the current published work about the linearization of wireless communications transmitters reveals the convenience of adequate behavioral models for the design of digital predistorters (DPD). The conventional baseband Volterra models, such as the full Volterra (FV) [1], the memory polynomial (MP) [2], or the generalized MP (GMP) [3], are specific for power amplifiers (PA), and have insufficient accuracy to represent more general nonlinear systems with complex-valued input signals.

To comply with the objective of linearizing a transmitter, designers address several challenging issues. In particular, the joint compensation of the I/Q-modulator and the PA impairments [4]–[7], and the problem of concurrent dual-band PAs [8] have been subject of concern in the last years. Another important goal is the design of a DPD that is capable of following the changes in the PA operating conditions. This has been performed by updating the DPD parameters in order to follow the power level changes [9]–[10].

Joint compensation of the I/Q modulator impairments and the PA nonlinearities are addressed in [5] by using a structure composed of a MP model dependent on the signal $x(k)$ and a filter, augmented with a similar second branch dependent

on the image signal $x^*(k)$. A similar approach is presented in [6] for the case of a multiple-input multiple-output (MIMO) transmitter. The generalized twin-box model [7] incorporates parallel-Hammerstein (PH)-based branches that ignore the out-of-diagonal kernels and may contain non-significant diagonal terms. In these approaches, some terms belonging to the GMP or the FV models, which can be important, are not included.

In [9], a power-scalable DPD based on a two-box architecture with a set of precalculated memoryless look-up table and a MP function to compensate the memory effects is proposed. The need of a power-adaptive DPD to compensate distortion in level-variable conditions was solved in [10] by adjusting the model parameters in accordance with the changes in the input power level. The procedure introduces extended scaling factors to track the varying signal power with an interpolation approach. The model extraction in both proposals requires PA measurements at multiple power levels and the number of coefficients is fixed independently of the varying PA nonlinear operation.

In this paper, we focus on a DPD design to linearize the transmitter in a range of output power levels incorporating a model with a complete set of parameters to jointly compensate the I/Q modulator impairments and the PA nonlinearities. The mathematical deduction of the general complex-valued Volterra series (CVS) representation [11] is based on Wirtinger calculus by considering the nonlinear system dependent on the input $x(k)$ and its complex-conjugate $x^*(k)$, operating these variables as real-valued. Then, the transmitter can be viewed as a two-input system and analyzed using a double Volterra series approach. Under the assumption of a Volterra representation, the CVS model is able to describe the nonlinear behavior of a wireless communications transmitter. The present proposal requires signal acquisition at only a single power operating point and the identified model is directly extended to a wide range of power levels. The rise of the number of coefficients in the CVS model makes necessary a suitable procedure to identify the model parameters in an efficient and robust manner, as in [12]–[13], where a technique for model reduction using the sparse structure of Volterra kernels was introduced. A thresholding procedure contributes to the model effectiveness by reducing the number of coefficients as the system enters into weakly nonlinear modes. In [14], the authors confirmed the robustness of the procedure by verifying that the set of parameters identified at a given power level are applicable to the accurate estimation of the system output over a wide dynamic range. Here, the proposed approach is applied to the design of a DPD to accomplish the joint linearization of

transmitter I/Q modulator and PA impairments, under power-varying conditions.

The next sections of this paper are organized as follows: After this introduction, Section II presents the rationale for the proposed identification method in regression models. First, subsection II-A describes the framework of the CVS model in the context of general discrete-time complex-valued non-linear systems. Next, subsection II-B gives a detailed theoretical justification of the proposed identification procedure, extending the material presented by the authors in [14], and establishes the power-scalable law for the model parameters. Subsection II-C reviews related works for the joint compensation of I/Q modulator and PA impairments. The application of the proposed identification procedure to power-adaptive transmitter linearization is covered in Section III. Two DPDs are designed and their performance is experimentally characterized: applied to a basic transmitter lacking pre-amplification in subsection III-A and to a realistic transmitter incorporating a preamplifier to increase the PA operation working level in subsection III-B. The power-scalable characteristics of the latter are compared with a DPD based on [6] in subsection III-C. The concluding remarks are summarized in Section IV.

II. DPD STRUCTURE AND IDENTIFICATION PROCEDURE

Compensation of transmitter RF impairments, originated predominantly by I/Q imbalance and nonlinearities, is one of the most important challenges for DPD designers. This problem can be approached by following two different points of view. The first approach is based on the knowledge of how the different blocks are assembled inside the transmitter (see Fig. 1a) and an approximate input-output representation is deduced for the whole system. In this paper an alternative perspective, where the predistorter is designed considering the transmitter as a black box, is presented below. For comparison, the first approach is discussed at the end of this section.

A. Proposed DPD Model for Joint Mitigation in Transmitters

In the context of Volterra series representation, any given system with complex-valued input can be modeled by the CVS model [11]. Viewed as a black box, the relationship for the input and output complex envelopes, $x(k)$ and $y_0(k)$, of the DPD in a wireless communications system can be expressed as

$$y_0(k) = h_{0,0} + \sum_{n=1}^{\infty} \left\{ \sum_{m=0}^{n} \sum_{\mathbf{q}_n=\mathbf{0}}^{\mathbf{Q}} \sum_{\mathbf{p}_m=\mathbf{0}}^{\mathbf{Q}} h_{n-m,m}(\mathbf{q}_{n-m}, \mathbf{p}_m) \times \right.$$
$$\left. \times \prod_{r=1}^{n-m} x(k-q_r) \prod_{s=1}^{m} x^*(k-p_s) \right\}. \quad (1)$$

The vector of delays are $\mathbf{q}_n = [q_1, q_2, \cdots, q_n]^T$ and $\mathbf{p}_m$, defined in a similar way, with a maximum delay Q for all indices. The product of the input signal samples is denoted as $\prod_{r=1}^{n} x(k-q_r) = x(k-q_1)x(k-q_2)\cdots x(k-q_n)$ and the same notation is used for the product of the image samples $x^*(k)$. For $m = 0$ and $m = n$, $h_{n,0}(\mathbf{q}_n)$ and $h_{0,n}(\mathbf{p}_n)$ are standard Volterra kernels and the products do not contain $x^*(k-p_s)$
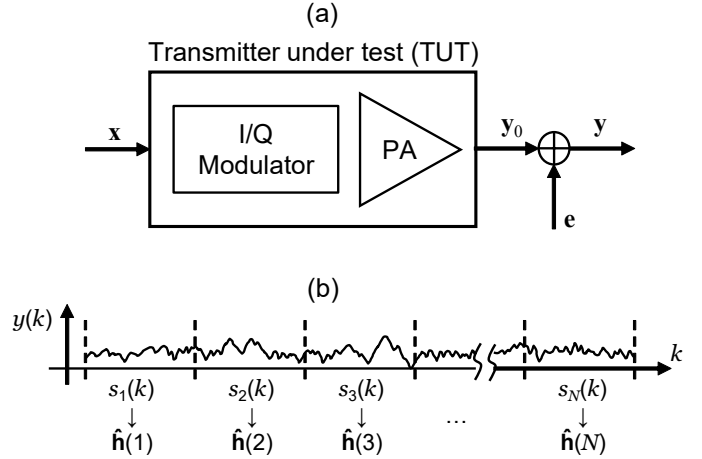

Fig. 1. Acquisition model with additive noise (a) and signal segmentation (b).

and $x(k-q_r)$, respectively. The factor $h_{n-m,m}(\mathbf{q}_{n-m}, \mathbf{p}_m)$ is a Volterra kernel of order n dependent on the two kinds of indices.

The CVS model (1) with $h_{0,0} = 0$ can be viewed also as a linear combination of the Volterra regressors $\prod_{r=1}^{n-m} x(k-q_r) \prod_{s=1}^{m} x^*(k-p_s)$. The M samples of the input signal can be disposed to form the column vector $\mathbf{x} = [x(0), x(1), \cdots, x(M-1)]^T$ and, if the Volterra regressors are likewise organized to construct the regressor vectors $\boldsymbol{\xi}_i$, the truncated version of (1) can be rewritten as

$$\mathbf{y}_0 = \sum_{i=1}^{N_R} h_i \boldsymbol{\xi}_i, \quad (2)$$

where the ith regression coefficient h_i comes from the Volterra kernels arranged in an ordered-fashion, $\mathbf{y}_0$ is a vector with the output samples disposed in the same way as $\mathbf{x}$, and N_R is the number of regressor vectors of the model. Examples of Volterra regressor vectors are $\boldsymbol{\xi}_1 = \mathbf{x}$, corresponding to the memoryless linear regressor, and $\boldsymbol{\xi}_{i_0} = [x(-q)|x(-q)|^2, x(1-q)|x(1-q)|^2, \cdots, x(M-1-q)|x(M-1-q)|^2]^T$, corresponding to the third-order memory polynomial regressor with a delay q. The regressor vectors constitute a complete set in the case of the CVS model. Although $\boldsymbol{\xi}_i$ are non-orthogonal, a feasible recursive algorithm to estimate the N_R coefficients of (2) is the Orthogonal Matching Pursuit (OMP) [15].

The so-called observation matrix $\mathbf{X}$ and the coefficients vector $\mathbf{h}$ are constructed by stacking all column vectors $\boldsymbol{\xi}_i$ and the model coefficients h_i, respectively. Therefore, (2) can be expressed in a compact matrix form [3],[13].

B. Identification Procedure

In an experimental setup, the vector of the acquired samples, $\mathbf{y}$, is contaminated by equipment distortion and noise. Assuming an approximately distortionless setup, for instance in laboratory conditions with high-performance equipment, the schematic of the acquisition experiment can be modeled as in Fig. 1. Therefore, the input-output relationship is

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{e}, \quad (3)$$

with a zero-mean additive white Gaussian (AWGN) noise $\mathbf{e}$, stationary and complex-valued. The model parameters can be identified by using a direct least-squares (LS) solution, given by

$$\hat{\mathbf{h}} = \left(\mathbf{X}^H\,\mathbf{X}\right)^{-1}\mathbf{X}^H\,\mathbf{y} + \mathbf{w}\,, \qquad (4)$$

where H represents the Hermitian transpose operation. The difference between the estimated vector $\hat{\mathbf{h}}$ and the true vector $\mathbf{h}$ is originated by the noise process $\mathbf{w}$. Observe that the selection of a model with lacking regressors is another important source of identification error. For example, the FV and the GMP representations, as particular cases of (1), are insufficient to model a transmitter with impairments.

The general character of the CVS model provides a complete set of regressors with a high number of coefficients and, at the same time, an associated suitability for a pruning procedure without a presumption on the significant regressors. The application of the Bayesian information criterion (BIC) was proposed in [13], [14], [16] to select the significant parameters. Whereas [16] is based on a simulated annealing algorithm with all the possible variants of the model, the approach in [13] is based on the application of the OMP algorithm to represent the PA output as its projection onto the span of Volterra regressors, and on the BIC rule to discard the irrelevant coefficients, maintaining only the active regressors. It is worth noticing that while LS identification is affected by regressor correlation, the OMP algorithm guarantees the recovery of the exact value of the coefficients in a noiseless environment in a given number of iterations [17].

The alternative procedure applied in [14] is mainly a parameters detection based on statistical hypotheses testing. Referred to Fig. 1b), the entire set of acquired samples is divided in N segments. Focusing on the ith coefficient of the estimated vector $\hat{\mathbf{h}}$, each segment returns an independent measurement of this particular coefficient. We can define the vector $\tilde{\mathbf{h}}_i$ constructed with the N different realizations of this random variable to decide whether or not the ith coefficient has to be incorporated as an active parameter of the model. We use the Neyman-Pearson (NP) approach to make the decision, based on two hypotheses [18]:

- The measurement is produced by noise,

$$\mathcal{H}_0\text{: if } \tilde{\mathbf{h}}_i = \mathbf{w}_i,$$

- The measurement is produced by the presence of a model coefficient plus noise,

$$\mathcal{H}_1\text{: if } \tilde{\mathbf{h}}_i = \mathbf{h}_i + \mathbf{w}_i.$$

The probability density function under $\mathcal{H}_1$ is

$$p(\tilde{\mathbf{h}}_i;\mathcal{H}_1) = \frac{1}{\pi^N\sigma^{2N}}e^{-\frac{1}{\sigma^2}(\tilde{\mathbf{h}}_i-\mathbf{h}_i)^H(\tilde{\mathbf{h}}_i-\mathbf{h}_i)}\,, \qquad (5)$$

where σ^2 is the variance of the complex-valued AWGN $\mathbf{w}_i$. Likewise, the probability density function under $\mathcal{H}_0$ is expressed as

$$p(\tilde{\mathbf{h}}_i;\mathcal{H}_0) = \frac{1}{\pi^N\sigma^{2N}}e^{-\frac{1}{\sigma^2}\tilde{\mathbf{h}}_i^H\tilde{\mathbf{h}}_i}\,. \qquad (6)$$

For a particular probability of erroneous measurement induced by noise, the NP theorem states that the probability of true detection is maximized if the detector decides the hypothesis $\mathcal{H}_1$ when the likelihood ratio $L(\tilde{\mathbf{h}}_i)$ exceeds a given threshold γ, i.e.,

$$L(\tilde{\mathbf{h}}_i) = \frac{p(\tilde{\mathbf{h}}_i;\mathcal{H}_1)}{p(\tilde{\mathbf{h}}_i;\mathcal{H}_0)} > \gamma. \qquad (7)$$

Substituting (5) and (6) in the likelihood ratio and taking $\ln L(\tilde{\mathbf{h}}_i)$, we obtain

$$\ln L(\tilde{\mathbf{h}}_i) = \frac{2}{\sigma^2}\text{Re}\left(\mathbf{h}_i^H\tilde{\mathbf{h}}_i\right) - \frac{1}{\sigma^2}|\mathbf{h}_i|^2 \qquad (8)$$

and the decision

$$\text{Re}(\mathbf{h}_i^H\tilde{\mathbf{h}}_i) = \text{Re}\left(\sum_{r=1}^N h_i^*(r)\tilde{h}_i(r)\right) > \gamma' \qquad (9)$$

is equivalent to the likelihood ratio test (7). Under steady conditions, any coefficient is an unknown constant $h_i(r) = h_i$, with its estimator being the average value

$$h_i \approx \bar{\bar{h}}_i = \frac{1}{N}\sum_{r=1}^N \tilde{h}_i(r). \qquad (10)$$

Operating with (9), we can decide $\mathcal{H}_1$ if

$$\text{Re}\left(\bar{\bar{h}}_i^*\frac{1}{N}\sum_{r=1}^N \tilde{h}_i(r)\right) = \text{Re}\left(\bar{\bar{h}}_i^*\bar{\bar{h}}_i\right) = |\bar{\bar{h}}_i|^2 > \frac{\gamma'}{N}. \qquad (11)$$

There is a trade off between the deficit of model coefficients (if the threshold is high), and the inclusion of undesired noise (if the threshold is chosen too low).

The proposed procedure is as follows. Once the set of coefficients has been estimated by the OMP algorithm for each segment of the acquired signal (see Fig. 1b), the average $\bar{\bar{h}}_i$ and the hypothesis test (11) are computed for all coefficients. The lower the threshold level, the more coefficients (regressors) will be incorporated to the model and the BIC criterion is used to decide the optimum number of regressors n_c [13]. Due to the sparse character of the Volterra kernels, many coefficients can be discarded without a significant loss in accuracy.

In this paper, we express the BIC rule with its explicit dependence on the normalized mean square error (NMSE). If the NMSE is expressed in dB, the variance is given by

$$\hat{\sigma}_e^2 = \left(\frac{1}{M}\sum_{m=0}^{M-1}|y(m)|^2\right) \times 10^{\text{NMSE}/10} \qquad (12)$$

and substituting in (23) of [13], the BIC rule becomes

$$n_{c0} = \arg\min_{n_c}\left\{\text{NMSE} + \frac{n_c}{M}10\log(2M)\right\}. \qquad (13)$$

This procedure, summarized in Table I, and that published in [13], start with the OMP, a method that estimates the set of coefficients by iteratively adding new components to the coefficients vector. The difference is that in this work the coefficients computed with several segments of the signal are averaged and a further thresholding step is implemented to get the sparse model structure.

According to the results in [14], the normalized coefficients originated in the nonlinearities of the I/Q branches are not dependent on level variations at the modulator output and the coefficients associated to the FV regressors are normalized

TABLE I
SUMMARY OF THE PROPOSED IDENTIFICATION PROCEDURE

1: Divide the input $x(k)$ and measured output $y(k)$ signal into N segments, each with M samples.
2: Estimate the parameter vectors $\hat{\mathbf{h}}(r)$, $r = 1, \ldots, N$ using the OMP algorithm with the same pre-defined N_R number of components for all models.
3: For all the CVS components, compute the average coefficients $\bar{\bar{h}}_i$.
4: Sweep the thresholding level to apply the hypothesis test to all coefficients.
5: Take the thresholding level that optimizes the BIC rule.

according to the PA input level [13]. For example, if $h^{(n)}\big|_{P_{i0}}$ is a FV nth-order normalized coefficient (n odd) for an input signal with a power level P_{i0} and this level changes to $P = A^2 P_{i0}$ ($A > 0$, real-valued), the corresponding normalized coefficients follow the relation

$$\frac{h^{(n)}\big|_P}{h^{(n)}\big|_{P_{i0}}} = \left(\frac{1}{A}\right)^n = \left(\frac{P_{i0}}{P}\right)^{n/2}. \quad (14)$$

Once the normalized parameters of the DPD have been computed at a given input level, P_{i0}, they can be straightforwardly scaled to adapt the coefficients to other level P. Odd-order normalized parameters shared with the FV model follow an exponential scaling with the average input power, so that the magnitudes at P_{i0} and P are related as:

$$h^{(n)}\big|_P = h^{(n)}\big|_{P_{i0}} \cdot 10^{-\frac{n \cdot \Delta P(\text{dB})}{20}}, \quad (15)$$

where $\Delta P(\text{dB}) = P(\text{dBm}) - P_{i0}(\text{dBm})$ and n is the coefficient order. Those normalized coefficients not included in the FV model (e.g., the image or the even-order parameters associated with impairments of the I/Q modulator), are not dependent on the power level. In that case, the way the DPD is adapted to a decrement in power level is by excluding parameters with values below a given threshold, and not recalculating the coefficients [9], [10].

C. Preceding DPD Models for Joint Mitigation

In the previous procedure, the transmitter is viewed as a black box and the CVS model is adopted for the DPD. Another perspective is to assume the internal architecture of the transmitter (Fig. 1a) and deduce the DPD structure to join the mitigation of PA and I/Q-modulator impairments under reasonable approximations [4]-[6]. In [4], the rationale is to compensate the impairments in the reverse order that they appear. Based on the knowledge of the ensemble arranged by the modulator and the PA, a first MP structure is proposed for the PA DPD and then the modulator compensation is implemented with a widely linear (WL) model. This point of view is further extended to a general FV model [5] and to the augmented complex conjugate (ACC) model for joint mitigation of distortion in MIMO transmitters [6]. The output for the ACC model in a single-input single-output (SISO) transmitter is written as

$$y(k) = \sum_{n=1}^{N}{}' \bar{H}_n\{x(k)\} + \sum_{n=1}^{N}{}' \bar{H}_{cn}\{x^*(k)\} + h_0, \quad (16)$$
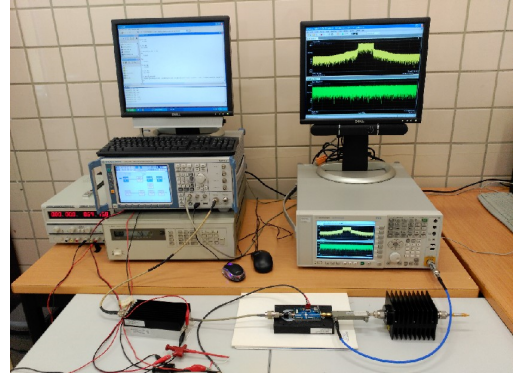

Fig. 2. Photograph of the measurement test bench. From left to right: (back) power supply for the preamplifier, SMU200A on top of PA power supply, PXA-N9030A; (front) preamplifier, Cree's CGH40010 evaluation board, attenuator.

where the prima symbols in the sums indicate that only odd orders are considered,

$$\bar{H}_n\{x\} = \sum_{q=0}^{Q} h_n(q)|x(k-q)|^{n-1}x(k-q) \quad (17)$$

and $\bar{H}_{cn}\{x^*(k)\}$ is defined in the same form.

The two viewpoints discussed above have strengths and weaknesses. When the transmitter composition is exploited, the principal advantage is the achievement of a simpler *a priori* structure for the DPD [6]. This favourable feature is based on the assumption of the particular modulator model, in this case a WL transformation. The WL supposition does not consider the contribution of the nonlinearities in the modulator baseband [19]-[20], but it is a good approximation in cases where the modulator nonlinearity can be neglected. On the other hand, making no assumptions about the internal architecture of the transmitter, the CVS model gives a complete (and huge) set of regressors. However, many of them are negligible because of the inherent sparse characteristics of the DPD. If the CVS representation is complemented with the compressed-sensing technique described in Section II-B, the result is a reduced model with only the indispensable number of active coefficients.

III. DPD DESIGN AND EXPERIMENTAL PERFORMANCE

To illustrate the proposed method, the joint compensation and linearization of two transmitters, with adaptable capability in a wide dynamic range, is demonstrated.

A. Linearization of a Basic Transmitter

The first case of study is the basic transmitter modeled in [14], referred to as TUT-1. The TUT-1 was arranged with the I/Q modulator integrated in the commercial generator SMU200A of Rohde & Schwarz, and a PA based on Cree's board for the evaluation of the power GaN HEMT CGH40010, operated at a carrier frequency of 3.6 GHz. The test bench was completed with the vector signal analyzer PXA-N9030A of Agilent Technologies (see Fig. 2). Later on a DPD linearizer for a more realistic transmitter (TUT-2) will be tested.
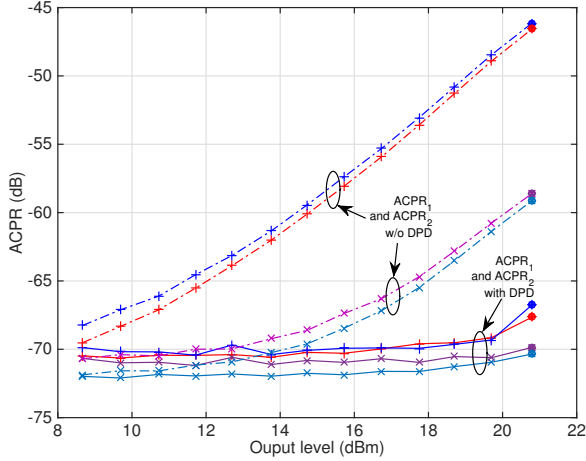
Fig. 3. Adjacent channel power ratios of the output spectrum with (solid lines) and without (dashed lines) DPD. First adjacent channel (+ marks) and second adjacent channel (× marks). Basic transmitter TUT-1.

Fig. 4. Normalized mean-square error of the output signal without DPD (dashed line) and with DPD (solid line). Basic transmitter TUT-1.

The probing signal was designed with an OFDM format and 15-MHz bandwidth, according to the LTE-downlink standard. The input vector **x** containing over 300000 samples of the complex envelope was uploaded into the generator and transmitted with a peak power of about +30 dBm, corresponding to a measured output average level of +19 dBm. Since the peak-to-average power ratio (PAPR) level is 11 dB, the upper limit to the transmitter power is fixed by the maximum modulator peak level, about +15 dBm (+6 dBm of average power). The samples of the output signal were gathered by averaging 300 acquisitions in the vector signal analyzer. Driven with this signal, the TUT-1 serves as a first proof of concept of the proposed linearization approach. To demonstrate the DPD capability to compensate the nonlinerities as well as the I/Q impairments, the settings of the modulator were configured with a quadrature error of 1 deg. Whereas commercial I/Q modulators show quadrature errors in the range of 1–3 deg [21]–[23], the choice of 1 deg was intended to show to which extent a reduced error can have a significant impact on the linearization capability of a DPD. The acquired baseband samples were subsequently off-line postprocessed with Matlab and used in a conventional indirect learning architecture to obtain the DPD coefficients at a maximum average output power of +20.8 dBm (+31.8 dBm of peak output power).

The model parameters were chosen to reduce the error between $y(k)/G$, the measured output scaled with the target gain of the linearized TUT-1, and $x(k)$. The underlying CVS structure was configured with thirteenth order, a maximum delay $Q = 3$ for orders 1 to 5, and memoryless (ML) for the higher orders, comprising 1370 coefficients. The OMP algorithm was executed to provide the ordering of 200 coefficients per segment. The synthesized DPD was composed of 22 normalized coefficients above a selected −68 dB threshold. The results of the linearization produced by the DPD at the output power of +20.8 dBm are plotted with filled circles in Fig. 3, where an adjacent channel power ratio (ACPR) of about −66 dB in the first adjacent channel is satisfactorily
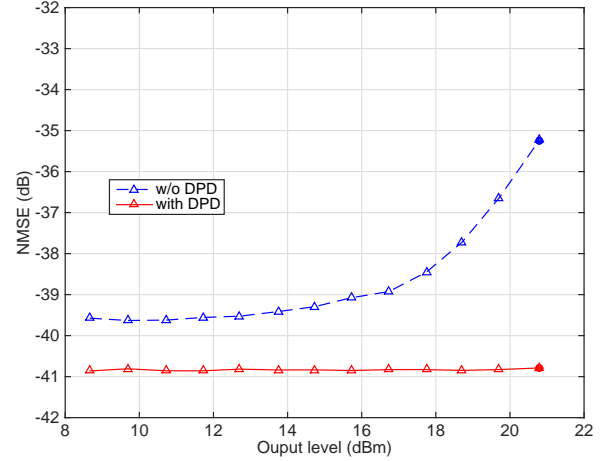
compared to the value of −46 dB obtained with the non-linearized transmitter. A reduction of about 10 dB in ACPR is also observed in the second adjacent channel. Likewise, about 6 dB of NMSE reduction was achieved with the TUT-1 driven at the maximum level, as it is shown with filled marks in Fig. 4.

The normalized coefficients of the DPD were extended at other power levels following the explained procedure, and those below the threshold were discarded. A second series of measurements was then performed with the new predistorted signals calculated according to the model adjusted at the different output levels, yielding the ACPR displayed with solid lines in Fig. 3. To have a reference of the DPD performance, the ACPRs of the transmitter without DPD are also shown (dashed lines), demonstrating a reduction over 21 dB in the first adjacent channel at $P_o = +20.8$ dBm. The DPD performs in all the dynamic range, demonstrating values of ACPR$_1$ and ACPR$_2$ better than −65 dB and −70 dB, respectively, without any further adjustment of the coefficients. This capability is also observed in the results of the NMSE displayed in Fig. 4 for the transmitter without DPD, in dashed line, and with DPD, in solid line. NMSE levels of about −41 dB in a wide range of operating powers demonstrate a flexible functioning of the linearizer.

The results of the DPD performance for the TUT-1 were encouraging and led to apply the procedure to a new transmitter where the power capability of the output device was exploited thoroughly.

B. Linearization of a Realistic Transmitter

The objective of this subsection is the joint compensation and linearization of a transmitter operating near its maximum level of +40 dBm. In this transmitter, referred to as TUT-2, a ZHL42W preamplifier of MiniCircuits is connected at the input of the Cree's evaluation board to drive it to a higher nonlinear operating point and delivering a maximum average output power of +26 dBm (+37 dBm of peak power), with
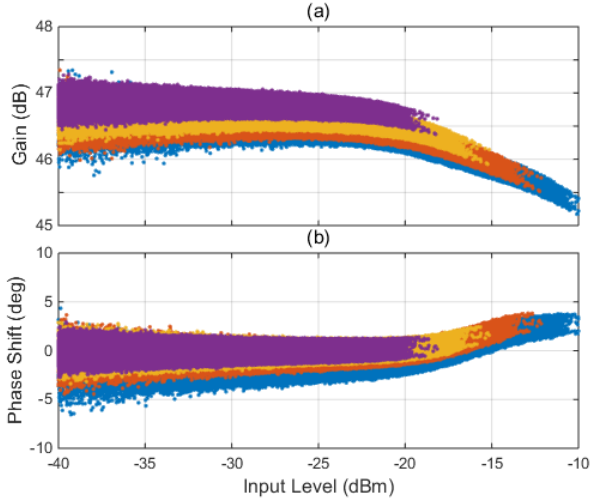
Fig. 5. Compression curves of the TUT-2 transmitter for average output levels of $+26$ dBm (blue), $+23.4$ dBm (red), 20.6 dBm (orange), and $+17.8$ dBm (magenta). Gain compression (a) and AM/PM (b) characteristics.
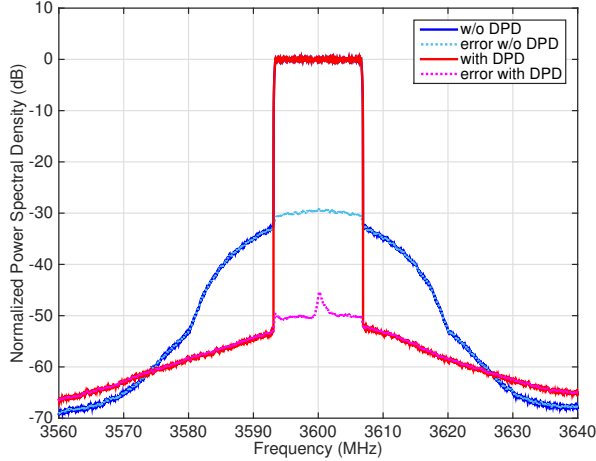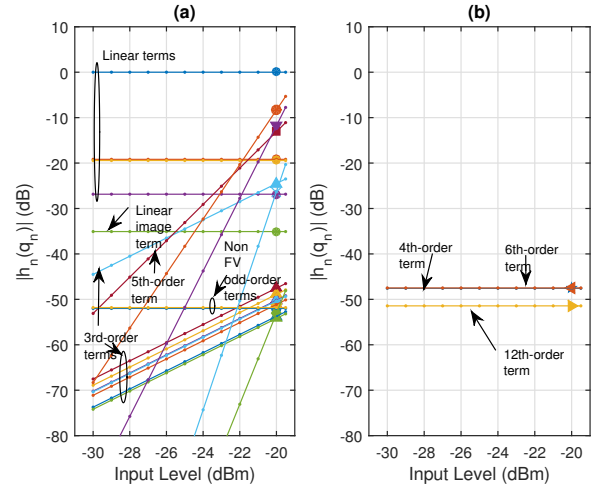

Fig. 7. Measured normalized coefficients at $P_o = +26$ dBm (large marks) and computed by using (15) (straight lines). (a) Odd-order and (b) even-order coefficients. The PA input power is represented in the abscissas axis. Realistic transmitter TUT-2.

Fig. 6. Normalized output spectrum and spectrum of the error for an average output level of $P_o = +26$ dBm ($+37$ dBm, peak), with and without DPD. Realistic transmitter TUT-2.
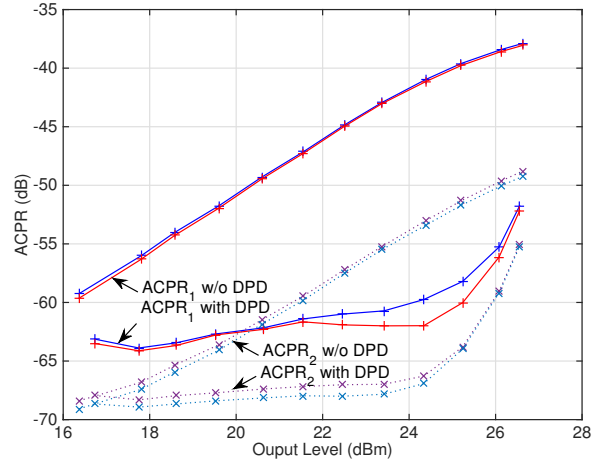

Fig. 8. Adjacent channel power ratios of the output spectrum with and without DPD. First adjacent channel (solid lines) and second adjacent channel (dotted lines). Realistic transmitter TUT-2.

a small-signal chain gain of about 46.7 dB. Again, 1 deg of quadrature error was considered for the I/Q modulator. The AM/AM and AM/PM characteristics are plotted in Fig. 5, revealing a gain compression of about 1.6 dB. The output spectrum is shown in Fig. 6, where we observe an important spectral regrowth that fails to comply with the standard ACPR of -45 dB. Let us also remark that not only the PA is contributing to the nonlinear distortion, but the I/Q modulator is also driving the preamplifier beyond its linear operation level. Again, a thirteenth-order CVS structure with 1370 coefficients was pruned. After repeating the procedure described in Section III-A, the corresponding DPD was designed. The linearization capability is also shown in the same figure, demonstrating a spectral regrowth reduction of about 20 dB in the first adjacent channel. Although a second iteration will usually provide better results, this initial design is adequate enough for the objectives of the present paper.

The normalized magnitudes of the 23 most relevant coefficients at $P_o = +26$ dBm ($P_i = -20$ dBm) are plotted with large marks in Fig. 7 and, in agreement with our previous discussion, the values adapted to other levels are displayed with straight lines. Notice that all the values are referred to the ML linear coefficient and that the number of parameters can be reduced as the PA is operated at lower levels. Since the precision requirements allow neglecting the normalized coefficients below a -50 dB threshold, the model at a PA input of $P_i = -20$ dBm is composed of only 14 coefficients, a value that is reduced to 8 coefficients at $P_i = -30$ dBm. Next, a set of predistorted signals were calculated according to the extended parameters.

In order to evaluate the performance of the DPD, the ACPR of the linearized signals are displayed in Fig. 8 with $+$ and $\times$ marks for the first and second adjacent channels, respectively.
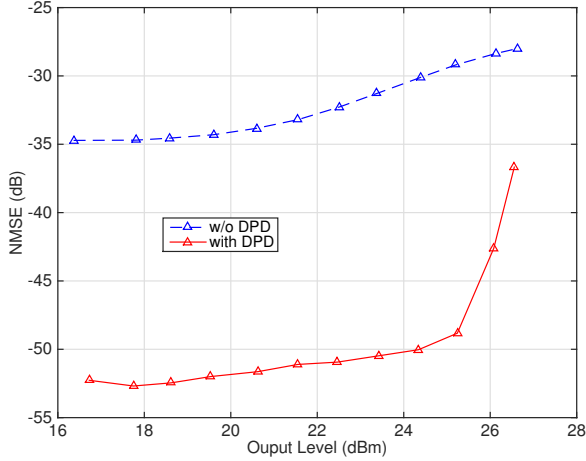
Fig. 9. Normalized mean-square error of the output signal without DPD (dashed line) and with DPD (solid line). Realistic transmitter TUT-2.
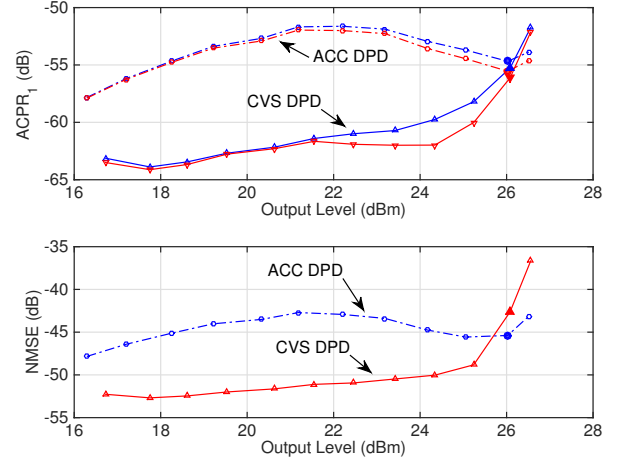

Fig. 11. Comparison of a) lower and upper ACPR$_1$ and b) NMSE for the SISO ACC model [6] (circles) and the present proposal with the CVS model (triangles). The filled marks correspond to performance at the identification level ($P_o = +26$ dBm). Realistic transmitter TUT-2.
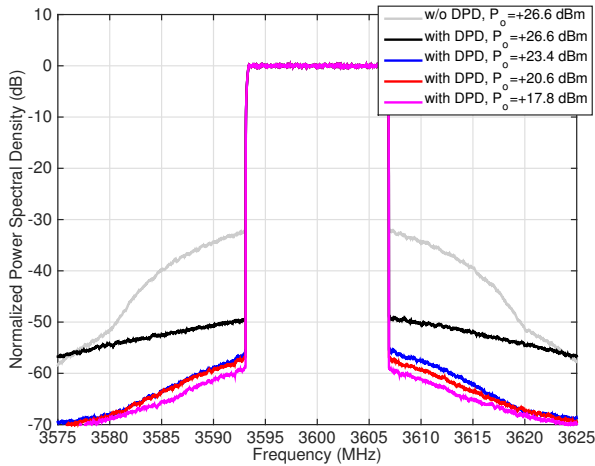

Fig. 10. Normalized spectral density of the linearized output for power levels ranging from $+17.8$ dBm to $+26.6$ dBm. As a reference, the trace of the output spectrum without DPD is also shown. Realistic transmitter TUT-2.

The ACPRs of the transmitter without DPD are also shown, demonstrating a reduction of 17 dB in the first adjacent channel at $P_o = +26$ dBm. The DPD linearization makes the ACPR$_1$ better than -55 dB and the ACPR$_2$ better than -60 dB, keeping the flexibility of compensating the linear and nonlinear impairments in the range from $P_o = +16$ dBm to $+26$ dBm without any further modification of the coefficients. This additional capability is also observed in the results of the NMSE displayed in Fig. 9 for the transmitter without DPD (dashed line) and with DPD (solid line). NMSE levels of about -50 dB in a wide range of operating powers demonstrate a maintained performance of the linearizer. Lastly, the normalized output spectra of the linearized transmitter are represented in Fig. 10 for average power levels ranging from $+17.8$ dBm to $+26.6$ dBm ($+28.8$ dBm to $+37.6$ dBm of peak power) in 3 dB steps.

C. Power Adaptability of ACC and CVS Linearizers

For TUT-2, we also implemented the DPD for joint mitigation of the modulator and the PA using the ACC model for SISO transmitters (16). The corresponding DPD was designed with equivalent settings (i.e, thirteenth order and $Q = 3$ for orders 1-5), yielding 32 coefficients that were estimated following the conventional LS algorithm. For the PA delivering $P_o = +26$ dBm, the attained ACPR$_1$ shown with filled circles in Fig. 11 a) has improved from -38.6 dB without DPD (see Fig. 8) to -55 dB. This value is very similar to that obtained with the present proposal, represented again here for comparison (triangles). These results indicate a good optimization of the performance at the operational point where both models were estimated (filled marks). In terms of model-order reduction, notice that the CVS DPD needs 14 coefficients, which compares favourably with the 32 coefficients necessary for a similar performance of the ACC DPD. When the ACC DPD is extended to other drive levels, its ACPR$_1$ deteriorates despite the PA is entering into a weakly nonlinear regime. On the contrary, the ACPR$_1$ of the CVS DPD shows a progressive reduction with a maximum improvement of about 9 dB with respect to the ACC DPD, demonstrating a better adaptability to output power variations. The NMSE results plotted in Fig. 11 b) repeat the performance deterioration of the ACC DPD and the adaptability of the CVS DPD under power level variations.

The ACPR$_1$ and NMSE behavior displayed in the figure is an example of overfitting in the case of the ACC DPD. The local character of the optimization is originated by the *a priori* assumption of the DPD structure. Therefore, the values estimated by the LS algorithm for the unneeded coefficients are 'coupled' to the indispensable ones and a variation of operational conditions disturbs the well functioning of the DPD at $P_o = +26$ dBm. The flexible performance exhibited by the CVS DPD is understood recalling that the method presented in this paper relies on a complete set of non-

orthogonal regressors, and on the OMP, a recursive algorithm to compute representations of the system output with respect to these non-orthogonal regressors. After a selection of the coefficients that surpass a given threshold, the resultant CVS DPD structure retains only the indispensable regressors.

IV. CONCLUSIONS

This paper has reported a reliable procedure to identify the model coefficients of a power-scalable DPD for the joint compensation of wireless communications transmitters. It has been applied to a DPD based on the CVS model and has been approvingly compared with other recent DPD proposals that have been advanced to overcome the limitations of conventional models under the presence of I/Q modulator impairments. The proposed DPD design uses a regressor search algorithm and a threshold to limit the number of parameters, allowing a significant reduction guaranteed by discarding negligible coefficients. Several causes are behind the model-order reduction: irrelevant kernel types, truncated nonlinear order, truncated memory depth, sparse memory delays, etc. The coefficients were identified at a power level where the PA is near saturation, and the invariance of the denormalized kernels in the measured dynamic range was exploited to estimate the values at other operating conditions without the need of interpolation or a new series of measurements. Tailoring of the number of parameters at other levels is performed with a simple rejection of regressors with normalized coefficients below the predetermined threshold. In consequence, the number of coefficients is further reduced as the transmitter enters into the less nonlinear operating range. In terms of model-order reduction and precision, a better performance with respect to other alternative approaches has been demonstrated in a wide range of measured power levels below the DPD upper operating point.

REFERENCES

[1] S. Benedetto, E. Biglieri, and R. Daffara, "Modeling and performance evaluation of nonlinear satellite links – A Volterra series approach," in *IEEE Trans. Aerosp. Electron. Syst.*, vol. 15, no. 4, pp. 494–507, Jul. 1979.

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

[3] 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.

[4] L. Anttila, P. Händel, and M. Valkama, "Joint mitigation of power amplifier and I/Q Modulator impairments in broadband direct-conversion transmitters," *IEEE Trans. Microw. Theory Techn.*, vol. 58, no. 4, pp. 730–739, Apr. 2010.

[5] B. Schubert, A. Gokceoglu, L. Anttila, and M. Valkama, "Augmented Volterra predistortion for the joint mitigation of power amplifier and I/Q modulator impairments in wideband flexible radio," in *Proc. Global Conf. Signal Inform. Process.*, Austin, TX, 3-5 Dec. 2013.

[6] Z. A. Khan, E. Zenteno, P. Händel, and M. Isaksson, "Digital predistortion for joint mitigation of I/Q imbalance and MIMO power amplifier distortion," *IEEE Trans. Microw. Theory Tech.*, vol. 65, no. 1, pp. 322-333, Jan. 2017.

[7] M. Younes and F. M. Ghannouchi, "Generalised twin-box model for compensation of transmitters radio frequency impairments," *IET Communications*, vol. 8, no. 4, pp. 413–418, Mar. 2014.

[8] H. Qian, S. Yao, H. Huang, X. Yang, and W. Feng, "Low complexity coefficient estimation for concurrent dual-band digital predistortion," *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 10, pp. 3153–3163, Oct. 2015.

[9] O. Hammi, A. Kwan, and F. Ghannouchi, "Bandwidth and power scalable digital predistorter for compensating dynamic distortions in RF power amplifiers," *IEEE Trans. Broadcast.*, vol. 59, no. 3, pp. 520–527, Sep. 2013.

[10] Y. Guo, C. Yu, and A. Zhu, "Power adaptive digital predistortion for wideband RF power amplifiers with dynamic power transmission," *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 11, pp. 3595–3607, Nov. 2015.

[11] C. Crespo-Cadenas, M. J. Madero-Ayora, J. Reina-Tosina, and J. A. Becerra-González, "Formal deduction of a Volterra series model for complex-valued systems," in *Signal Processing*, Vol. 131, pp. 245-248, (2017).

[12] J. Reina-Tosina, M. Allegue-Martínez, M. J. Madero-Ayora, C. Crespo-Cadenas, and S. Cruces, "Digital predistortion based on a compressed-sensing approach," in *Proc. 2013 Eur. Microw. Conf.*, Nuremberg, Germany, 6-10 Oct. 2013, pp. 408–411.

[13] J. Reina-Tosina, M. Allegue-Martinez, C. Crespo-Cadenas, Y. Chao, and S. Cruces, "Behavioral modeling and predistortion of power amplifiers under sparsity hypothesis," *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 2, pp. 745–753, Feb. 2015.

[14] C. Crespo-Cadenas, M. J. Madero-Ayora, J. Reina-Tosina, and J. A. Becerra-González, "Identification of Volterra model parameters in wireless systems," in *2017 IEEE Topical Conf. Power Amplifiers for Radio and Wireless Applicat. (PAWR)*, Phoenix, AZ, USA, 15-18 Jan. 2017, pp. 96–99.

[15] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal Matching Pursuit: Recursive Function Approximation with Applications to Wavelet Decomposition," in *27th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, 1-3 Nov. 1993.

[16] M. V. Amiri, S. A. Bassam, M. Helaoui, and F. M. Ghannouchi, "New order selection technique using information criteria applied to SISO and MIMO systems predistortion," *International Journal of Microwave and Wireless Technologies*, vol. 5, no. 2, pp. 123–131, Apr. 2013.

[17] Y. C. Eldar and G. Kutyniok (eds.), "Compressed sensing: theory and applications," *Cambridge University Press*, Cambridge, UK, p. 40, 2012.

[18] S. M. Kay, "Fundamentals of Statistical Signal Processing: Detection Theory," *Prentice Hall*, 1998.

[19] D. Wissel, "Identification and Measurement of Transmitter Non-Linearities," in *56th ARFTG Conference Digest*, Vol. 38, pp. 1-6 AZ, USA, 30 Nov-1 Dec. 2000.

[20] H. Cao, A. S. Tehrani, C. Fager, T. Eriksson, and H. Zirath, "I/Q imbalance compensation using a nonlinear modeling approach," *IEEE Trans. Microw. Theory Techn.*, vol. 57, no. 3, pp. 513–518, Mar. 2009.

[21] J. K. Cavers and M. Liao, "Adaptive compensation for imbalance and offset losses in direct conversion transceivers," *IEEE Trans. Vehicular Techn.*, vol. 42, no. 4, pp. 581–588, Nov. 1993.

[22] J. K. Cavers, "The effect of quadrature modulator and demodulator errors on adaptive digital predistorters for amplifier linearization," *IEEE Trans. Vehicular Techn.*, vol. 46, no. 2, pp. 456–466, May 1997.

[23] Intersil application note AN1022, "Operation and performance of the ISL5239 predistortion linearizer," July 2002.

Carlos Crespo-Cadenas (M'73–SM'15) was born in Madrid, Spain. He received the degree in Physics in 1973 and Doctor degree in 1995 from the Polytechnique University of Madrid.

Since 1998 he has been Associate Professor and currently he teaches lectures on Radio Communications in the Department of Signal Theory and Communications, University of Seville. His current interests are nonlinear analysis applied to the linearization of wireless communications subsystems.

María J. Madero-Ayora (S'06–M'09) received the Telecommunication Engineering and Doctor degrees from the University of Seville, Seville, Spain, in 2002 and 2008, respectively.

Since 2003 she has been with the Department of Signal Theory and Communications, University of Seville, and is currently Associate Professor. Her research interests include the nonlinear analysis of active microwave devices, compensation of impairments in modulators and power amplifiers, and measurement techniques of nonlinear communication systems.

Juan A. Becerra-González (S'12) received the B.S. and M.Sc. degrees in Telecommunication Engineering in 2009 and 2012, respectively, from University of Seville, Spain. He spent some time in the industry working for Accenture and Canon. He is currently enrolled in the Ph.D. program in Automatics, Electronics and Telecommunication Engineering at the University of Seville and he is also pursuing a Ph.D. in Electrical and Computer Engineering at the University of Delaware, Newark, DE, USA.

His main research areas include behavioral modeling of power amplifiers, linearization techniques, compressive-sensing and high dimensional signal processing.

Javier Reina-Tosina (S'99-M'06-SM'09) was born in Seville, Spain. He received the Telecommunication Engineering and Doctor degrees from the University of Seville, Seville, Spain, in 1996 and 2003, respectively.

Since 1997 he has been with the Department of Signal Theory and Communications, University of Seville, where he is currently an Associate Professor. His research interests include the nonlinear analysis of active microwave devices, behavioral modeling of power amplifiers and linearization techniques, the integration of information technologies in biomedicine, intelligent devices for homecare, and bioelectromagnetics.