Spectral Weighting Orthogonal Matching Pursuit Algorithm for Enhanced Out-of-Band Digital Predistortion Linearization

Pere L. Gilabert, *Senior Member, IEEE*, David López-Bueno, *Member, IEEE* and Gabriel Montoro, *Member, IEEE*

Abstract—This paper presents a new variant of the orthogonal matching pursuit (OMP) algorithm for reducing the computational complexity of the digital predistortion (DPD) behavioral model in the forward path. The proposed spectral weighting OMP (SW-OMP) algorithm focuses on selecting the most relevant basis functions to compensate for the out-of-band residual distortion which may eventually be masked by the dominant in-band residual error. This basis selection is carried out in an off-line process that does not affect the computational complexity of the real-time closed-loop DPD but, on the contrary, reduces its complexity while enhancing the robustness. Experimental results show that by selecting the DPD coefficients with the SW-OMP, the inherent ACLR and NMSE degradation suffered when reducing the number of coefficients is mitigated under strong nonlinear operation, when compared to using the basis functions selected by the classical OMP algorithm.

Index Terms—5G mobile communication, power amplifiers, digital predistortion linearization.

I. INTRODUCTION

DIGITAL predistortion (DPD) linearization is the most common and widespread solution today to cope with the power amplifier's (PA) inherent linearity versus efficiency trade-off. The use of 5G spectrally efficient waveforms featuring high peak-to-average power ratio (PAPR) and occupying wider bandwidths in multiple antenna transmitters [1] only aggravates such compromise. When considering wide bandwidth signals, carrier aggregation or multi-band configurations [2] in high efficient transmitter architectures, such as Doherty PAs, envelope tracking PAs or outphasing transmitters, the number of parameters required in the DPD model to compensate for both static nonlinearities and dynamic memory effects can be unacceptably high. This has a negative impact in the DPD model extraction/adaptation process, because it increases the computational complexity which may provoke over-fitting and uncertainty in the DPD estimation stages [3]. However, by applying regularization techniques [4] we can both avoid the numerical ill-conditioning of the estimation and reduce the

number of coefficients of the DPD function in the forward path, which ultimately impacts the baseband processing computational complexity and power consumption.

The objective of such regularization techniques is to enforce the sparsity constraint on the vector of parameters by minimizing the number of active components (i.e., ℓ_0-norm) subject to a constraint on the ℓ_2-norm squared of the identification error. Unfortunately, this is a non-deterministic polynomial-time hard (NP-hard) combinatorial search problem. Therefore, in the field of DPD linearization, several sub-optimal approaches have been proposed targeting both robust identification and model order reduction such as: the least absolute shrinkage and selection operator (LASSO), used for example by Wisell et al. in [5] and consisting in a ℓ_1-norm regularization; the Ridge regression, used for example by Guan et al. in [6] and consisting in a ℓ_2-norm regularization; the sparse Bayesian learning (SBL) algorithm, used by Peng et al. in [7]; or the orthogonal matching pursuit (OMP), a greedy algorithm for sparse approximation used in [8] by Reina et al. to select the most relevant basis functions of the DPD function. Considering the use of these algorithms in an off-line search to obtain the most relevant basis functions (which thus obviates the computational complexity needed for an online implementation), Mozos compares in [9] the DPD performance when reducing the number of coefficients of the DPD function by using LASSO, Ridge or OMP algorithms. For a certain PA nonlinear behavior, the basis sets selected by these algorithms yielded similar coefficient reduction versus NMSE/ACLR degradation, although the OMP algorithm presented slightly better performance.

This paper is not aimed at providing an overview or comparison on the existing model order reduction techniques as found in [10]. Instead, in this paper an alternative approach to the classical OMP algorithm which is based on a spectral weighting strategy is proposed. This new approach is oriented to select the most relevant DPD basis functions by paying special attention to the out-of-band distortion compensation. When the mean squared error is mainly dominated by the linear part, by adding spectral weighting (through in-band notching of the reference signal and the DPD basis functions) the OMP algorithm can focus on the subset of parameters that better contribute to minimize the mean squared error which is dominated by the out-of-band distortion.

The remainder of this paper is organized as follows. In Section II, the proposed spectral weighting orthogonal matching
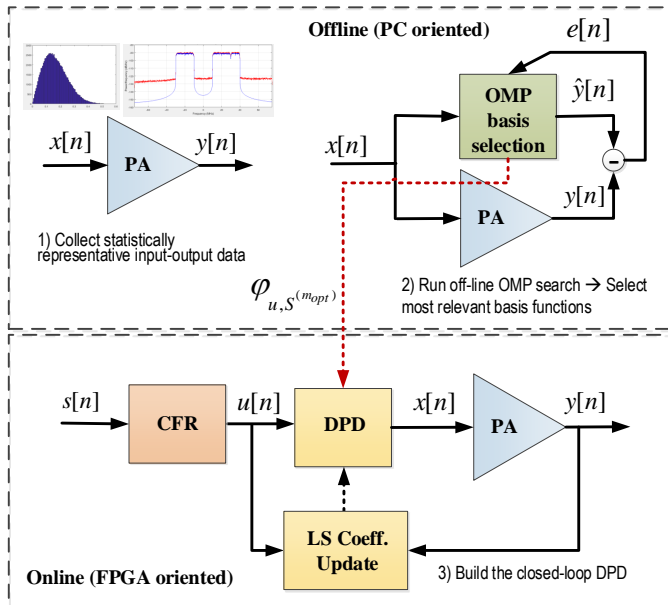
Fig. 1. Block diagram of the SW-OMP off-line search process and DPD linearization architecture.

pursuit (SW-OMP) algorithm is described in detail. Section III shows the experimental test bench and discusses the experimental results, showing the advantages of the proposed SW-OMP algorithm in regard to the order reduction of the DPD function. Finally, conclusions are given in Section IV

II. SPECTRAL WEIGHTING ORTHOGONAL MATCHING PURSUIT (SW-OMP) ALGORITHM

A. Description of the off-line basis functions selection process

The objective is to build a robust and accurate DPD linearizer that meets the linearity levels specified in communications standards making use of the minimum necessary number of coefficients. Reducing the number of coefficients has a beneficial impact on the amount of FPGA resources [11] (e.g., less DSP48 slices and RAM memory blocks in Xilinx FPGAs) required for the implementation of the DPD in the forward path. An off-line study of the most relevant basis functions involved in the characterization of the PA nonlinear distortion is conducted a priori in order to be later used for DPD purposes. The search of the most relevant components, given a general behavioral model descriptor (e.g., full Volterra series, generalized memory polynomial, etc.), is carried out only once in an off-line process. To further enhance the DPD linearization performance obtained with the subset of the selected components, we propose to use the SW-OMP, which is oriented toward emphasizing the out-of-band distortion part of the mean squared error. The whole process is schematically depicted in Fig. 1 and described in the following:

1) Collect the input-output data of the PA with a test signal having the same statistical properties as the one that will be later used for the transmission.
2) Run off-line (e.g., in a PC running Matlab) the proposed SW-OMP algorithm to obtain a sorted set of the most relevant basis functions that better characterize the PA

nonlinear distortion with special emphasis on the out-of-band distortion. This search is done only once.
3) Build the closed-loop DPD taking into account the basis functions previously selected by the SW-OMP algorithm. The identification of the DPD will be well-conditioned, robust and may take only few iterations to converge.

In the following subsection we will provide a detailed description of the SW-OMP algorithm.

B. PA Behavioral Modeling Identification

Following the notation in Fig. 1, the estimated PA behavioral model output $\hat{\mathbf{y}}$ can be defined as:

$$\hat{\mathbf{y}} = \boldsymbol{\Phi}_x \hat{\mathbf{w}}_{pa} \tag{1}$$

where $\hat{\mathbf{w}}_{pa}$ is the $O \times 1$ vector with the estimated coefficients and $\boldsymbol{\Phi}_x$ is the $L \times O$ data matrix containing the basis functions or components.

As explained before, the objective is to minimize the number of active components (i.e., ℓ_0-norm) subject to a constraint on the ℓ_2-norm squared of the identification error,

$$\min_{\hat{w}_{pa}} \|\hat{\mathbf{w}}_{pa}\|_0 \tag{2}$$

$$\text{subject to} \quad \|\mathbf{y} - \boldsymbol{\Phi}_x \hat{\mathbf{w}}_{pa}\|_2^2 \leq \varepsilon$$

where $\mathbf{y}$ is the $L \times 1$ vector of measured data at the PA output. To overcome this NP-hard combinatorial search problem, greedy algorithms, such as the OMP, have been proved useful to deal with this minimization problem.

In order to minimize the number of coefficients being required by the DPD function in the forward path, we assume that the optimal subset of selected basis functions of the DPD function will be the same as that used for PA behavioral modeling. In the following, we consider the generalized memory polynomial (GMP) behavioral model in [12],

$$\hat{y}[n] = \sum_{i=0}^{N_a-1} \sum_{p=0}^{P_a-1} \alpha_{pi} \cdot x[n-\tau_i^a] \big| x[n-\tau_i^a] \big|^p +$$
$$\sum_{j=1}^{M_b} \sum_{i=0}^{N_b-1} \sum_{p=1}^{P_b} \beta_{pij} \cdot x[n-\tau_i^b] \big| x[n-\tau_i^b - \tau_j^b] \big|^p + \tag{3}$$
$$\sum_{j=1}^{M_c} \sum_{i=0}^{N_c-1} \sum_{p=1}^{P_c} \gamma_{pij} \cdot x[n-\tau_i^c] \big| x[n-\tau_i^c + \tau_j^c] \big|^p$$

where α_{pi}, β_{pij} and γ_{pij} are the complex coefficients describing the model, and τ^a, τ^b and τ^c (with $\tau \in \mathbb{Z}$ and $\tau_0 = 0$) are the most significant non-consecutive delays of the input signal $x[n]$ that better contribute to characterize memory effects. The most relevant delays can be either found empirically or by using a greedy algorithm such as the proposed SW-OMP. The total number of coefficients is $O = P_a N_a + P_b N_b M_b + P_c N_c M_c$. From now on, the original coefficients of the GMP, α_{pi}, β_{pij} and γ_{pij}, are mapped for simplicity into w_r coefficients, with $r = 0, 1, \cdots, O - 1$.

Consequently, the $L \times O$ data matrix in (1) containing the basis functions or components (intended either for PA
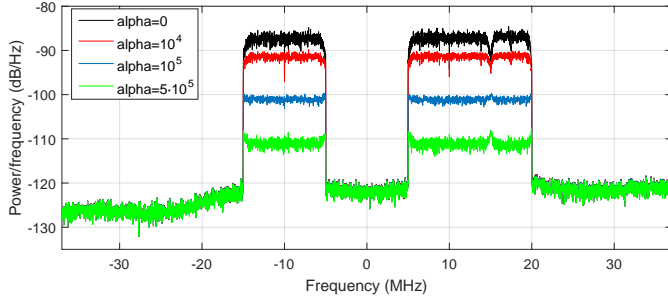
Fig. 2. Spectra of the filtered PA outputs for different alpha values.

Algorithm 1 Spectral Weighting Orthogonal Matching Pursuit

1: **procedure** SW-OMP $(\mathbf{Y}_f, \mathbf{\Psi}_{x,f})$
2:     initialization:
3:     $\mathbf{E}^{(0)} = \mathbf{Y}_f - \hat{\mathbf{Y}}^{(0)}$; with $\hat{\mathbf{Y}}^{(0)} = 0$
4:     $\mathbf{S}^{(0)} = \{\}$
5:     **for** $m = 1$ **to** m_{max} **do**
6:         $i^{(m)} = \arg\min_i \min_{\hat{w}_{pa,i}} \left\| \mathbf{E}^{(m-1)} - \mathbf{\Psi}_{x,f,\{i\}} \hat{w}_{pa,i} \right\|_2^2 \approx$
   $\arg\max_i \left| \mathbf{\Psi}_{x,f,\{i\}}^H \mathbf{E}^{(m-1)} \right|$
7:         $\mathbf{S}^{(m)} \leftarrow \mathbf{S}^{(m-1)} \bigcup i^{(m)}$
8:         $\hat{\mathbf{w}}_{pa}^{(m)} = \left(\mathbf{\Psi}_{x,f,S^{(m)}}^H \mathbf{\Psi}_{x,f,S^{(m)}} \right)^{-1} \mathbf{\Psi}_{x,f,S^{(m)}}^H \mathbf{Y}_f$
9:         $\hat{\mathbf{Y}}^{(m)} = \mathbf{\Psi}_{x,f,S^{(m)}} \hat{\mathbf{w}}_{pa}^{(m)}$
10:         $\mathbf{E}^{(m)} = \mathbf{Y}_f - \hat{\mathbf{Y}}^{(m)}$
11:     **end for**
12:     return $\mathbf{S}^{(m_{max})}$
13: **end procedure**

---

identification or for DPD linearization), $\mathbf{\Phi}_x$, can be defined as

$$\mathbf{\Phi}_x = \left(\boldsymbol{\varphi}_x[0], \boldsymbol{\varphi}_x[1], \cdots, \boldsymbol{\varphi}_x[n], \cdots, \boldsymbol{\varphi}_x[L-1] \right)^T \quad (4)$$

where $\boldsymbol{\varphi}_x[n] = \left(\phi_0^x[n], \phi_1^x[n], \cdots, \phi_r^x[n], \cdots, \phi_{O-1}^x[n] \right)^T$ is the $O \times 1$ vector with the basis functions at time n.

In the following subsection we will describe the SW-OMP algorithm capable to find the most relevant basis functions from $\mathbf{\Phi}_x$ and to be later used in the DPD linearizer.

C. Description of the SW-OMP algorithm

In order to select the most relevant basis functions by paying special attention to the out-of-band distortion compensation, the proposed SW-OMP targets the minimization of an in-band notched version of the mean square error. Consequently, the PA output and all the basis functions will be notched in-band.

Therefore, as described in [13], the PA output is filtered in the frequency domain as follows,

$$Y_f[k] = \Gamma\{Y[k]\} = \frac{Y[k]}{1 + \alpha |Y[k]|^2} \quad (5)$$

with α being a parameter to control the notching level (see Fig. 2) and with $Y[k]$ being the discrete Fourier transform (DFT) of $y[n]$,

$$Y[k] = DFT_N\{y[n]\} = \sum_{n=0}^{N-1} y[n] e^{-j \frac{2\pi}{N} kn} \quad (6)$$

where $k = 0, \cdots N - 1$ with $N \geq L$. Similarly, each basis function in the frequency domain is filtered as follows,

$$\mathbf{\Upsilon}_{x,f}[k] = \Gamma\{\mathbf{\Upsilon}_x[k]\} = \quad (7)$$
$$\left(\Gamma\{\psi_0^x[k]\}, \cdots, \Gamma\{\psi_r^x[k]\}, \cdots, \Gamma\{\psi_{O-1}^x[k]\} \right)^T$$

with $\psi_r^x[k]$ $(r = 0, \cdots, O-1)$ being the DFT of $\phi_r^x[n]$ and $\Gamma\{\cdot\}$ the filtering operator. In the frequency domain, the $N \times O$ matrix of filtered basis functions is defined as

$$\mathbf{\Psi}_{x,f} = \left(\mathbf{\Upsilon}_{x,f}[0], \cdots, \mathbf{\Upsilon}_{x,f}[k], \cdots, \mathbf{\Upsilon}_{x,f}[N-1] \right)^T. \quad (8)$$

Now, $\mathbf{S}^{(m)}$ is defined as the support set containing the indices of the basis functions building the model. Considering that m_{max} is the number of basis functions under study (i.e., $m_{max} = O$), the proposed SW-OMP algorithm is defined in Algorithm 1.

At every iteration, the basis function that better contributes to minimize the filtered residual error is selected and added to the support set $\mathbf{S}^{(m)}$. The elements of $\mathbf{\Psi}_{x,f,S^{(m)}}$ have been normalized in power to simplify the index $i^{(m)}$ calculation in line 6 of the Algorithm, which can be obtained by maximizing the absolute value of the correlation between the filtered basis function $\mathbf{\Psi}_{x,f,\{i\}}$ (i.e, $\mathbf{\Psi}_{x,f,\{i\}} = \Gamma\{DFT_N\{\mathbf{\Phi}_{x,\{i\}}\}\}$) and the filtered residual error $\mathbf{E}^{(m-1)}$ of the previous iteration. After a complete SW-OMP search, we obtain a vector $\mathbf{S}^{(m_{max})}$ with the indices of all the original basis functions (active components) sorted according to their relevance. Then, by using some information criterion, such as the Akaike (AIC) or the Bayesian (BIC) [8], it is possible to determine the optimum number of coefficients (m_{opt}), where $m_{opt} < m_{max}$. Finally, the subset of selected basis functions, $\boldsymbol{\varphi}_{x,S^{(m_{opt})}}[n]$, is used in (9) to carry out the DPD (i.e., $\boldsymbol{\varphi}_u[n] = \boldsymbol{\varphi}_{x,S^{(m_{opt})}}[n]$).

D. Closed-loop Digital Predistortion

Thanks to the SW-OMP algorithm we can define the closed-loop DPD function in the forward path by properly selecting the most relevant basis functions that contribute to minimizing the out-of-band distortion. The input-output relationship at the DPD block in Fig. 1 is defined as

$$\mathbf{x} = \mathbf{u} + \mathbf{\Phi}_u \hat{\mathbf{w}}_{dpd} \quad (9)$$

where $\hat{\mathbf{w}}_{dpd} = \left(\hat{w}_0, \hat{w}_1, \cdots, \hat{w}_{P-1} \right)^T$ is the $P \times 1$ (with $P < O$) estimated vector of coefficients of the DPD behavioral model, $\mathbf{\Phi}_u$ is the $L \times P$ data matrix $(n = 0, 1, \cdots, L - 1)$,

$$\mathbf{\Phi}_u = \left(\boldsymbol{\varphi}_u[0], \boldsymbol{\varphi}_u[1], \cdots, \boldsymbol{\varphi}_u[n], \cdots, \boldsymbol{\varphi}_u[L-1] \right)^T \quad (10)$$

and where $\boldsymbol{\varphi}_u[n] = \left(\phi_0^u[n], \phi_1^u[n], \cdots, \phi_r^u[n], \cdots, \phi_{P-1}^u[n] \right)^T$ is the $P \times 1$ vector containing the P (with $P < O$) previously selected basis functions using the SW-OMP algorithm (i.e., $\boldsymbol{\varphi}_u[n] = \boldsymbol{\varphi}_{x,S^{(m_{opt})}}[n]$). Following the direct learning approach, the DPD coefficients can be estimated iteratively

by using the least squares (LS) criterion,

$$\hat{\mathbf{w}}_{dpd}^{i+1} = \hat{\mathbf{w}}_{dpd}^{i} + \mu\left(\mathbf{\Phi}_u^H \mathbf{\Phi}_u\right)^{-1}\mathbf{\Phi}_u^H \mathbf{e} \qquad (11)$$

where μ is a weighting factor. The $L \times 1$ error vector is defined as $\mathbf{e} = \frac{\mathbf{y}}{G_0} - \mathbf{u}$, where G_0 is the desired PA linear gain.

E. Other considerations

The benefit of using the SW-OMP is to reduce the computational complexity of the DPD behavioral model in the forward path. The SW-OMP algorithm is inherently stable and the index search is repeated until the upper-bound m_{max} on the number of active coefficients is reached, or until the norm of the filtered residual error $\left\|\mathbf{E}^{(m)}\right\|_2$ falls below the noise tolerance threshold. The running time depends on both the number of coefficients considered in the search and on the digital signal processor used for running the algorithm. However, by comparing the SW-OMP algorithm with the Less Relevant Basis Removal (LRBR) brute-force technique presented in [10] in terms of computational time, the OMP is 21 times faster, while the accuracy of the search is similar to that obtained with OMP. The computational complexity of the SW-OMP in comparison to the OMP is slightly higher due to the FFT transformations and filtering operations. However, when running an off-line SW-OMP search (e.g., in a PC running Matlab instead of in a FPGA) considering a high number of coefficients (e.g., > 300 coeff.), the computational time devoted to calculate the Fourier transforms and filtering operations is negligible. Instead, thanks to the SW-OMP search, some benefits in the DPD forward path are obtained in terms of robustness (e.g., fast convergence in just 3 iterations when considering the use of different input data featuring different PAPR), flexibility (e.g., trade-off between the number of selected coefficients and the DPD linearization performance) and reliability (by selecting the most relevant basis the over-parametrization is avoided which yields to a properly conditioned and accurate coefficient estimation).

III. EXPERIMENTAL TEST-BENCH AND RESULTS

A. Experimental Test-bench

The test bench shown in Fig. 3 was used to conduct experimental validation of the SW-OMP technique. To emulate a LTE carrier-aggregation transmission, two 20 MHz bandwidth channels and one 10 MHz channel were grouped as shown in Fig. 2. The PAPR of the waveform has been limited to 9.5 dB by using the peak cancellation crest factor reduction technique. The DPD processing blocks run in a PC with Matlab, which is also employed to interface the waveform generation and acquisition instruments. The direct RF generation of the LTE carrier-aggregated signal was carried out through the arbitrary waveform generator M8190A from Keysight, considering a clock rate of 7.968 GHz and 14 bits. The device under test was a class AB-biased GaN pHEMT 6 W PA delivering 28.6 dBm mean output power at 2140 MHz (downlink base-station LTE band). The gate bias was chosen to maximize efficiency at the price of showing severe nonlinear behavior. Finally, a digital storage oscilloscope (DSO Keysight 90404A) was used

to acquire the RF output signal with 8-bit resolution and at sampling rate of 20 GS/s.

B. Experimental Results

For the experimental results, we considered a GMP behavioral model with a configuration of polynomial orders and memory terms that resulted in an initial dictionary of $m_{max} = 596$ basis functions. The use of 596 coefficients in the GMP model resulted in an ill-conditioned estimation. By using the Moore-Penrose inverse (i.e., $(\mathbf{\Phi}_u^H \mathbf{\Phi}_u)^{-1}\mathbf{\Phi}_u^H$) the coefficients estimation was inaccurate and some regularization was required. With the SW-OMP we addressed the over-parametrization problem by properly selecting the most relevant basis functions that minimize the out-of-band distortion.

After applying the SW-OMP search and the BIC approximation, we obtained a reduced set (being sorted according to their relevance) of around $m_{opt} = 350$ active components, which resulted in a perfectly well-conditioned estimation. In addition, as shown in Fig. 4, further pruning was considered by selecting some of the most relevant components of $\mathbf{S}^{(m_{opt})}$ and by taking into account different notching levels (i.e., different α values in (5) as shown in Fig. 2). We empirically found that $\alpha = 5 \cdot 10^5$ is the maximum notching level that guarantees the minimum in-band level required for accurate selection of the most relevant basis. The DPD linearization performance was evaluated in terms of ACLR and NMSE. As expected (see Fig. 4), under severe nonlinear behavior of the PA, the selection made by using the SW-OMP with the highest α outperforms the classical OMP approach and is more robust against the inherent ACLR degradation suffered when reducing the number of coefficients of the DPD function.

In addition, in order to show the convergence speed and robustness of the closed-loop DPD when considering different data bursts with different PAPR values at each iteration, Fig. 5 and Fig. 6 show, respectively, the linearization performance of the DPD operating with 50 and 150 properly selected coefficients and considering different notching levels. By taking into account this particular PA nonlinear behavior, the target of -45 dBc of ACLR is only met when using a minimum of 150 coefficients of the GMP behavioral model (see Fig. 6). Fig. 7 shows the output power spectra before and after DPD linearization when considering 150 coefficients being properly selected with the SW-OMP algorithm and when considering the highest notching level (i.e., $\alpha = 5 \cdot 10^5$). However, the advantage given by conducting a previous SW-OMP search in comparison to the classical OMP search, is more evident when fewer coefficients are considered in the DPD (e.g., only 50 coefficients in Fig. 5) to compensate for the PA severe nonlinear behavior.

IV. CONCLUSION

The proposed SW-OMP algorithm emphasizes the importance of the out-of-band distortion in the OMP residual error, forcing the selection of the DPD components that better compensate for the adjacent channel spectral regrowth. Experimental results showed that, thanks to the proposed SW-OMP technique, the inherent ACLR and NMSE degradation suffered
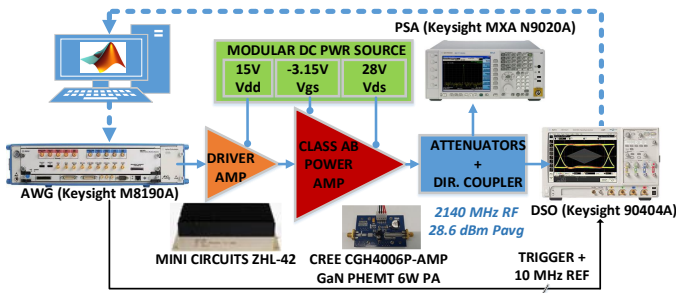
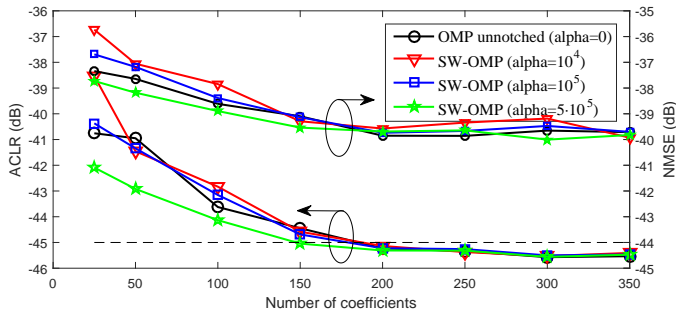Fig. 3. Experimental test bench employed for SW-OMP DPD profiling.

Fig. 4. ACLR and NMSE versus number of DPD selected coefficients (using SW-OMP) for different alpha values of the notch filter.
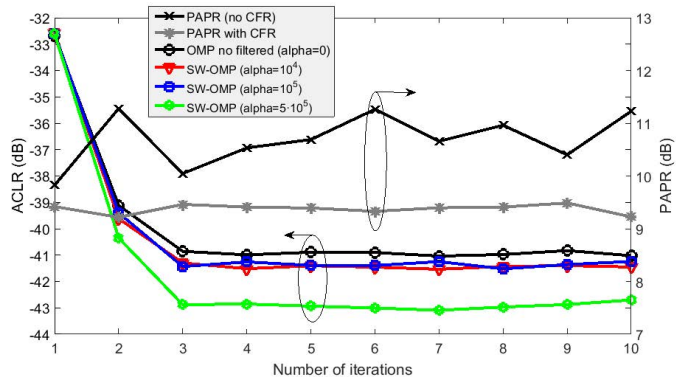

Fig. 5. PAPR and ACLR vs # of iterations, with 50 coeff. SW-OMP DPD.

Fig. 6. PAPR and ACLR vs # of iterations, with 150 coeff. SW-OMP DPD.

Fig. 7. PA output power spectra before and after SW-OMP DPD.

when significantly reducing the number of coefficients can be mitigated (and improved in comparison to the classical OMP under severe PA nonlinear behavior), while the robustness of the DPD is enhanced, since no further adaptation is required after 3 iterations.

REFERENCES

[1] P. M. Suryasarman and A. Springer, "A comparative analysis of adaptive digital predistortion algorithms for multiple antenna transmitters," *IEEE Trans. on Circ. and Sys. I*, vol. 62, no. 5, pp. 1412–1420, May 2015.

[2] P. Jaraut, M. Rawat, and F. M. Ghannouchi, "Harmonically related concurrent tri-band behavioral modeling and digital predistortion," *IEEE Trans. on Circ. and Sys. II: Express Briefs*, pp. 1–5, Oct. 2018.

[3] J. Chani-Cahuana, M. Ozen, C. Fager, and T. Eriksson, "Digital pre-distortion parameter identification for RF power amplifiers using real-valued output data," *IEEE Trans. on Circ. and Sys. II: Express Briefs*, vol. 64, no. 10, pp. 1227–1231, Oct. 2017.

[4] R. N. Braithwaite, "Digital predistortion of an RF power amplifier using a reduced volterra series model with a memory polynomial estimator," *IEEE Trans. on Microw. Theory and Tech.*, vol. 65, no. 10, pp. 3613–3623, Oct. 2017.
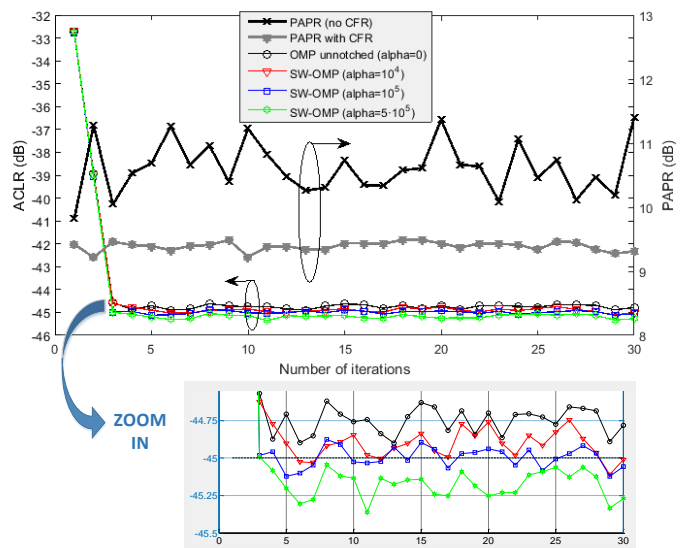
[5] D. Wisell, J. Jalden, and P. Handel, "Behavioral power amplifier modeling using the lasso," in *2008 IEEE Inst. and Meas. Tech. Conf.*, May 2008, pp. 1864–1867.

[6] L. Guan and A. Zhu, "Optimized Low-Complexity Implementation of Least Squares Based Model Extraction for Digital Predistortion of RF Power Amplifiers," *IEEE Trans. on Microw. Theory and Tech.*, vol. 60, no. 3, pp. 594–603, Jan. 2012.

[7] J. Peng, S. He, B. Wang, Z. Dai, and J. Pang, "Digital predistortion for power amplifier based on sparse bayesian learning," *IEEE Trans. on Circ. and Sys. II: Express Briefs*, vol. 63, no. 9, pp. 828–832, Sep. 2016.

[8] J. Reina-Tosina, M. Allegue et al., "Behavioral modeling and predistortion of power amplifiers under sparsity hypothesis," *IEEE Trans. on Microw. Theory and Tech.*, vol. 63, no. 2, pp. 745–753, Feb. 2015.

[9] J. M. Mozos, "Model order reduction techniques for power amplifier digital predistortion linearization," Jul. 2016. [Online]. Available: http://hdl.handle.net/2117/96850

[10] P. L. Gilabert, G. Montoro, et al., "Comparison of model order reduction techniques for digital predistortion of power amplifiers," in *46th European Microw. Conf. (EuMC)*, Oct. 2016, pp. 182–185.

[11] J. Wood, "System-level design considerations for digital pre-distortion of wireless base station transmitters," *IEEE Trans. on Microw. Theory and Tech.*, vol. 65, no. 5, pp. 1880–1890, May 2017.

[12] D. R. Morgan, Z. Ma et al., "A Generalized Memory Polynomial Model for Digital Predistortion of RF Power Amplifiers," *IEEE Trans. on Signal Processing*, vol. 54, no. 10, pp. 3852–3860, Oct. 2006.

[13] R. N. Braithwaite, "Reducing estimator biases due to equalization errors in adaptive digital predistortion systems for RF power amplifiers," in *2012 IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2012, pp. 1–3.