

A Doubly Orthogonal Matching Pursuit Algorithm for Sparse Predistortion of Power Amplifiers

Juan A. Becerra, *Student Member, IEEE*, María J. Madero-Ayora, *Member, IEEE*,
 Javier Reina-Tosina, *Senior Member, IEEE*, Carlos Crespo-Cadenas, *Senior Member, IEEE*,
 Javier García-Frías, *Senior Member, IEEE*, Gonzalo Arce, *Fellow, IEEE*

Abstract—This letter presents a new method for the digital predistortion of power amplifiers (PAs) based on sparse behavioral models. The Gram-Schmidt orthogonalization is synergistically integrated into the orthogonal matching pursuit algorithm to decorrelate the selected model regressors against the components still to be selected. Experiments conducted on a test bench based on a GaN PA driven by a 15-MHz orthogonal frequency division multiplexing signal were conducted in order to validate the algorithm. Experimental results in a digital predistortion application and a comparison with other state-of-the-art algorithms highlight the enhancement of its pruning capabilities, reducing the number of coefficients while maintaining the performance.

Index Terms—Behavioral modeling, compressive-sensing, digital predistortion, orthogonal matching pursuit, power amplifier.

I. INTRODUCTION

THE evolution of wireless communication systems is pushing the design of power amplifiers (PAs) towards challenging constraints in terms of linearity and efficiency. The power range where the PA is more efficient is also where nonlinearity occurs, leading to the need of techniques as digital predistortion (DPD) to mitigate this inconvenience.

The pruning of behavioral models is being extensively researched by the community, coming along with a wide set of techniques that range from the application of the principal component analysis (PCA) method [1] to matching pursuit algorithms like the compressive sampling matching pursuit (CoSaMP) [2] and the orthogonal matching pursuit (OMP) [3]. Compressed-sensing techniques are characterized by their simplicity and flexibility, due to the low computational complexity of the greedy algorithms these techniques are based on. Greedy algorithms do not perform any *a priori* decision about the selection of components, becoming suitable for the pruning of Volterra series. OMP [4] is a greedy algorithm that makes a hard decision based on a local optimal criterion whereby the estimated output of the model is always orthogonal to the residual [5], orthogonalization from which OMP takes its name. Since OMP selects one component of the model

at each iteration, a pseudoinverse of the data matrix with a number of columns that is equal to the iteration value has to be performed. The work in [6] proposed a simplified sparse parameter identification resulting in a lower computational complexity and in [7], a greedy algorithm ensures that the selected element has maximum energy in every iteration.

The estimation of Volterra coefficients is intricate since the basis functionals of the Volterra series are highly correlated. This correlation leads to a large condition number in the model matrices, implying that the equations system is ill-conditioned, affecting the least-squares (LS) solution.

A doubly orthogonal matching pursuit (DOMP) algorithm is proposed in this letter to enhance the selection of coefficients in a sparse parameter identification of the model. The main contribution of this work is the addition of the Gram-Schmidt orthogonalization at one step of the OMP algorithm, decorrelating the selected regressors and those still to be selected.

The remainder of this letter is organized as follows. First, Section II introduces the details of model selection and derives the DOMP algorithm. DPD experimental design and results are detailed in Section III. Finally, Section IV summarizes the main results and concludes the paper.

II. THE DOUBLY ORTHOGONAL MATCHING PURSUIT

The recovery of a high-dimensional vector from a set of measured observations arises in many different fields. The measurement process equation models a linear relation between the Volterra kernel vector $\mathbf{h} \in \mathbb{C}^n$ and the system output $\mathbf{y} \in \mathbb{C}^m$,

$$\mathbf{y} = \mathbf{X} \cdot \mathbf{h} + \mathbf{w}, \quad (1)$$

through the measurement matrix $\mathbf{X} \in \mathbb{C}^{m \times n}$, which contains one component of the model in each of its columns. The vector \mathbf{w} represents the measurement noise.

Greedy algorithms aim at recovering the Volterra kernel vector through an iterative approach following the ℓ_1 -norm minimization. In this section, the DOMP algorithm is developed. Its pseudocode instructions are shown in Algorithm 1.

A. Initialization

The algorithm returns a support set, S , whose elements are sorted in decreasing impact over the output. The initial state of the support set is empty, $S^{(0)} = \emptyset$, since no components have been added to it yet. Prior to the algorithm iterations, the matrix $\mathbf{Z}^{(0)} = \mathbf{X}$ is defined. This matrix will be used to keep

Manuscript received April 19, 2018; accepted May 17, 2018. (*Corresponding author: Juan A. Becerra.*)

J. A. Becerra, J. García-Frías and G. Arce are with the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE, 19716 USA (e-mail: becerra@udel.edu; jgf@udel.edu, arce@udel.edu).

J.A. Becerra, M.J. Madero-Ayora, J. Reina-Tosina and C. Crespo-Cadenas are with the Departamento de Teoría de la Señal y Comunicaciones, Escuela Técnica Superior de Ingeniería, Universidad de Sevilla, Sevilla, 41092, Spain (e-mail: jabcerra@us.es, mjmadero@us.es, jreina@us.es, ccrespo@us.es).

the information of the orthogonalized regressors, and after the algorithm execution, it will be equal to the result of applying the Gram-Schmidt procedure to the regressors \mathbf{X} in the order of the final support set $S^{(\text{end})}$. The residual is set to $\mathbf{r}^{(0)} = \mathbf{y}$, since it keeps track of the remainder left to be captured by the selected regressors of the model and initially there are no components in the support set.

B. Loop of the Algorithm

At each iteration t , the algorithm normalizes the basis components by dividing each column of the matrix \mathbf{Z} by its ℓ_2 -norm

$$\mathbf{Z}_{\{i\}}^{(t-1)} \leftarrow \frac{i \notin S^{(t-1)}}{\|\mathbf{Z}_{\{i\}}^{(t-1)}\|_2} \mathbf{Z}_{\{i\}}^{(t-1)}. \quad (2)$$

Then, the component with the highest normalized scalar resolution in the direction of the residual is selected

$$i^{(t)} \leftarrow \arg \max_{i \notin S^{(t-1)}} \left| \mathbf{Z}_{\{i\}}^{(t-1)H} \mathbf{r}^{(t-1)} \right|, \quad (3)$$

where H is the Hermitian transpose.

Please note that the condition $i \notin S^{(t-1)}$ is always met since the residual is orthogonal to the selected basis set $S^{(t-1)}$. This condition decreases the computational complexity of the algorithm in a real implementation. The selected component is then included into the support set

$$S^{(t)} = S^{(t-1)} \cup i^{(t)}. \quad (4)$$

Next, the algorithm performs the Gram-Schmidt orthogonalization by first obtaining

$$\mathbf{p}^{(t)} = \mathbf{Z}_{\{i^{(t)}\}}^{(t-1)H} \mathbf{Z}^{(t-1)}, \quad (5)$$

where the components of $\mathbf{p} \in \mathbb{C}^n$ are the vector projections of the selected regressor onto each one of the components of the basis. This projection is subtracted from each regressor, hence the selected component is orthogonal to the remaining of the basis set

$$\mathbf{Z}^{(t)} = \mathbf{Z}^{(t-1)} - \mathbf{p}^{(t)} \otimes \mathbf{Z}_{\{i^{(t)}\}}^{(t-1)}, \quad (6)$$

where the Kronecker product is denoted by \otimes .

With the inclusion of the Gram-Schmidt process a double orthogonalization is ensured: the selected basis set is always orthogonal to the residual and, at the same time, it is orthogonal to the regressors that do not belong to the selected support set. The estimation of the Volterra kernel vector is computed through LS

$$\hat{\mathbf{h}} = \mathbf{X}_{S^{(t)}}^+ \mathbf{y}, \quad (7)$$

where $\mathbf{X}_{S^{(t)}}^+ = \tilde{\mathbf{X}}_{S^{(t)}} \mathbf{X}_{S^{(t)}}^H$ is the Moore-Penrose pseudoinverse. The matrix $\tilde{\mathbf{X}}_{S^{(t)}} = (\mathbf{X}_{S^{(t)}}^H \mathbf{X}_{S^{(t)}})^{-1}$ may be used for attaining a lower complexity version of the algorithm as in [6]. The estimation of the output $\hat{\mathbf{y}}^{(t)}$ can be obtained through

$$\hat{\mathbf{y}}^{(t)} = \mathbf{X}_{S^{(t)}} \hat{\mathbf{h}}. \quad (8)$$

Finally, the residual is updated according to

$$\mathbf{r}^{(t)} = \mathbf{y} - \hat{\mathbf{y}}^{(t)}. \quad (9)$$

Algorithm 1 Summary of the Doubly Orthogonal Matching Pursuit (DOMP) Algorithm

Input: $\mathbf{X} \in \mathbb{C}^{m \times n}$, $\mathbf{y} \in \mathbb{C}^m$

Output: $S^{(t)}$

- 1: *Initialization* : $\mathbf{Z}^{(0)} = \mathbf{X}$, $\mathbf{r}^{(0)} \leftarrow \mathbf{y}$, $S^{(0)} \leftarrow \emptyset$
 - 2: **for** $t = 1$ till stopping criterion is met **do**
 - 3: $\mathbf{Z}_{\{i\}}^{(t-1)} \leftarrow \frac{i \notin S^{(t-1)}}{\|\mathbf{Z}_{\{i\}}^{(t-1)}\|_2} \mathbf{Z}_{\{i\}}^{(t-1)}$
 - 4: $i^{(t)} \leftarrow \arg \max_{i \notin S^{(t-1)}} \left| \mathbf{Z}_{\{i\}}^{(t-1)H} \mathbf{r}^{(t-1)} \right|$
 - 5: $S^{(t)} \leftarrow S^{(t-1)} \cup i^{(t)}$
 - 6: $\mathbf{p}^{(t)} \leftarrow \mathbf{Z}_{\{i^{(t)}\}}^{(t-1)H} \mathbf{Z}^{(t-1)}$
 - 7: $\mathbf{Z}^{(t)} \leftarrow \mathbf{Z}^{(t-1)} - \mathbf{p}^{(t)} \otimes \mathbf{Z}_{\{i^{(t)}\}}^{(t-1)}$
 - 8: $\hat{\mathbf{h}} \leftarrow \mathbf{X}_{S^{(t)}}^+ \mathbf{y}$
 - 9: $\hat{\mathbf{y}}^{(t)} \leftarrow \mathbf{X}_{S^{(t)}} \hat{\mathbf{h}}$
 - 10: $\mathbf{r}^{(t)} \leftarrow \mathbf{y} - \hat{\mathbf{y}}^{(t)}$
 - 11: **end for**
-

III. DPD EXPERIMENTAL DESIGN AND RESULTS

For the experimental validation of the algorithm generating a DPD, the following test bench was used. The transmitter under test is composed by the commercial I/Q modulator SMU200A of Rohde & Schwarz followed by a ZHL42W preamplifier of MiniCircuits and the evaluation board of a PA based on the Cree CGH40010 GaN HEMT. A vector signal analyzer PXA-N9030A from Keysight Technologies is employed. A test signal based on a standard orthogonal frequency division multiplexing (OFDM) was generated by modulating 16-QAM random data into 900 carriers, resulting in a band of 15 MHz. The signal exhibits a peak-to-average power ratio (PAPR) of 10 dB and an oversampling rate of 6 was applied. The operation point is characterized by a gain of 48.9 dB and a gain compression of 1 dB. The target output power was fixed to 31.9 dBm and the normalized mean square error (NMSE) between the input and the output had a value of -28.3 dB. The adjacent channel power ratio (ACPR) values are -37.1 dBc and -36.9 dBc for the first left channel and the first right channel, respectively.

In order to test the pruning capabilities of the algorithm, a generalized memory polynomial (GMP) [8] model was selected, in which a configuration of thirteenth order was set. A fading memory with order (from 15 to 1) in the memory polynomial branch was fixed. The non-diagonal branches were set to seventh order and memory length of 1. The test Volterra model before pruning contains 77 components. A number of 5500 samples were used for the DPD calculation, producing a signal length to number of coefficients ratio of about 89.

The OMP, DOMP, CoSaMP and PCA algorithms were run over an indirect learning scheme in order to perform a benchmark. In each iteration of the algorithms, a DPD was obtained, allowing to show the effect of adding new components in all the schemes. The Bayesian information criterion (BIC) proposed in [3] was also calculated along with the OMP error evolution to attain its optimal number of coefficients. Since CoSaMP requires a sparsity target, it was run sweeping the sparsity level up to the number of

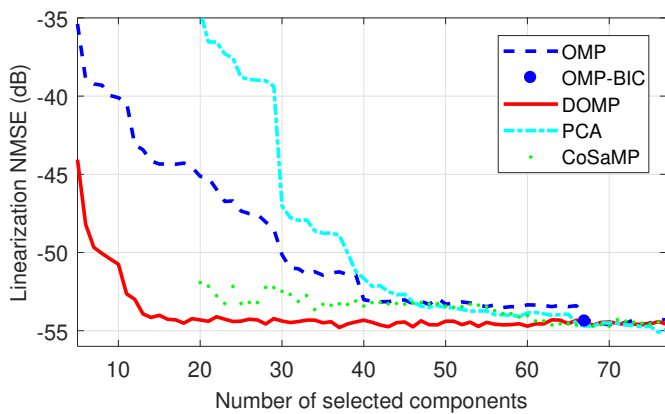


Fig. 1. Linearization NMSE for the set of algorithms in comparison. The DOMP algorithm shows better predistortion capabilities while reducing the number of coefficients of the model.

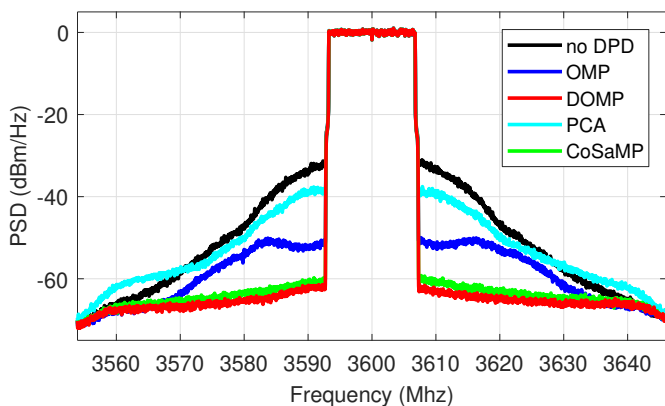


Fig. 2. Power spectral density for the output without DPD, and the DPD generated by the OMP, DOMP, PCA and CoSaMP algorithms for 20 coefficients.

components.

The linearization NMSE, i.e., the error between the input signal after the linear gain of the PA and the output of the predistorted PA, is shown in Fig. 1. The DOMP outperforms the rest of algorithms for every number of selected components. For a fixed linearization error, the number of components selected by the rest of the algorithms is always higher compared to the DOMP set.

With the aim of comparing their linearization performance, the power spectral density (PSD) of the output without DPD and with DPD attained through the set of the tested algorithms for a model of 20 coefficients is shown in Fig. 2. The linearization capability of each algorithm is evidenced by its reduction in spectral regrowth, obtaining a better linearization performance with the DOMP DPD.

Finally, Table I shows the linearization ACPR for a fixed number of 20 coefficients in each algorithm. The best predistortion performance, corresponding to the DOMP algorithm, improves in 28 dB the ACPR of the first adjacent channels with respect to the case without DPD. Notice that the DOMP DPD meets the spectral mask required for LTE.

TABLE I
ACPR VALUES OF THE SYSTEM WITHOUT DPD AND THE LINEARIZED SIGNAL FOR THE OMP, DOMP, PCA AND CoSaMP ALGORITHMS FOR 20 COEFFICIENTS.

	ACPR-2 (dBc)	ACPR-1 (dBc)	ACPR+1 (dBc)	ACPR+2 (dBc)
no DPD	-46.6	-37.1	-36.9	-46.3
OMP	-54.7	-52.1	-51.6	-53.8
DOMP	-66.9	-65.1	-65.1	-66.0
PCA	-49.9	-42.1	-42.4	-49.7
CoSaMP	-64.4	-62.6	-62.3	-63.8

IV. CONCLUSION

A novel method based on the OMP algorithm for the pruning of PA behavioral models has been presented. The addition of the Gram-Schmidt orthogonalization in the iteration of the algorithm enables the selection of a better set of components in a highly-correlated regressors scenario. The framework for the use of this new technique is provided with a comprehensive explanation of the algorithm steps. The paper has illustrated the ability of producing a model with less components than the classical method while still maintaining the performance. A benchmark against state-of-the-art algorithms through the results of a DPD produced with each technique has shown the outperforming of the DOMP algorithm, giving an overview of the quality of the results that can be obtained.

The coefficient reduction capabilities of the proposal have been demonstrated with the predistortion of a commercial PA, showing a better linearization performance for the same number of components. Its faster pruning and error minimization make DOMP a forward-looking proposal in the pruning of Volterra series models.

REFERENCES

- [1] P. L. Gilbert, G. Montoro, D. López, N. Bartzoudis, E. Bertran, M. Payar, and A. Hourtane, "Order reduction of wideband digital predistorters using principal component analysis," in *2013 IEEE MTT-S Int. Microw. Symp. Dig.*, June 2013, pp. 1–7.
- [2] A. Abdelhafiz, A. Kwan, O. Hammi, and F. M. Ghannouchi, "Digital predistortion of LTE-A power amplifiers using compressed-sampling-based unstructured pruning of Volterra series," *IEEE Trans. Microw. Theory Techn.*, vol. 62, no. 11, pp. 2583–2593, Nov 2014.
- [3] J. Reina-Tosina, M. Allegue-Martínez, C. Crespo-Cadenas, C. Yu, 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.
- [4] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec 2007.
- [5] Y. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*, ser. Compressed Sensing: Theory and Applications. Cambridge University Press, 2012.
- [6] J. Peng, S. He, B. Wang, Z. Dai, and J. Pang, "Digital predistortion for power amplifier based on sparse Bayesian learning," *IEEE Trans. Circuits Syst. II, Express Briefs*, vol. 63, no. 9, pp. 828–832, Sept 2016.
- [7] M. Li, Z. Yang, Z. Zhang, R. Li, Q. Dong, and S. Nakatake, "Sparsity adaptive estimation of memory polynomial based models for power amplifier behavioral modeling," *IEEE Microw. Wireless Compon. Lett.*, vol. 26, no. 5, pp. 370–372, May 2016.
- [8] D. Morgan, Z. Ma, J. Kim, M. 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.