Optimization of a pruned 2-D Digital Predistortion Model Structure for Power Amplifiers Linearization

Siqi Wang, Morgan Roger, Caroline Lelandais-Perrault

Optimization of a pruned 2-D Digital Predistortion Model Structure for Power Amplifiers Linearization

Siqi Wang, Morgan Roger, Caroline Lelandais-Perrault

GeePs — Group of electrical engineering - Paris, UMR CNRS 8507,
CentraleSupélec, Univ. Paris-Sud, Université Paris-Saclay, Sorbonne Université
3 & 11 rue Joliot-Curie, Plateau de Moulon 91192 Gif-sur-Yvette CEDEX, France

Abstract—**Two-dimensional digital predistortion (2-D DPD) is one of the most commonly used approaches to linearize the concurrent dual-band radio frequency (RF) power amplifiers (PA). This paper explores the use of a hill-climbing optimization heuristic to determine the optimal structure for 2-D DPD. To improve convergence and reduce computation time, we propose a new parameterization for the 2-D DPD model. The proposed search criterion is based on the generalized information criterion, which represents the trade-off between the DPD linearization accuracy in both bands and the model complexity. A comparison against a compressed-sensing-based method is also made. The effectiveness of the proposed method is validated with two 20 MHz long-term-evolution (LTE) signals on two carriers with 100 MHz frequency separation.**

Index Terms—**Digital predistortion, information criterion, non-linear distortion, opitmization, RF power amplifiers**

I. INTRODUCTION

Multi-band power amplifiers (PA) are designed for the compatibility of different standards in modern wireless communication systems [1]. The conventional single-band digital predistortion (DPD) becomes difficult for the linearization since the feedback bandwidth is largely increased if the dual-band signal is considered as a single broadband signal. Bassam *et al.* [2] proposed a 2-D memory polynomial (MP) model DPD by processing the signals in different bands separately and adding cross terms to preserve accuracy.

Even for single-band DPD, determining the model structure according to a trade-off between model accuracy and complexity is an important step before estimating the model coefficients. Some optimization heuristic (OH) methods have been proposed to address this problem [3], [4]. In [3], an algorithm based on a hill-climbing heuristic was developed for sizing single-band generalized memory polynomial (GMP) models.

These methods determine the optimal full models with minimal nonlinearity orders and memory depths for an acceptable linearization performance. Sparse models are considered in [5] for the case of 2-stage cascade MP models. Determining a sparse model structure is equivalent to pruning a complete model whose nonlinearity order and memory depth are large. A compressed sensing (CS) method under sparsity hypothesis to select only important model kernels has been proposed for the model pruning in [6]. This pruning technique is applied for the case of 2D-MP model in [7].
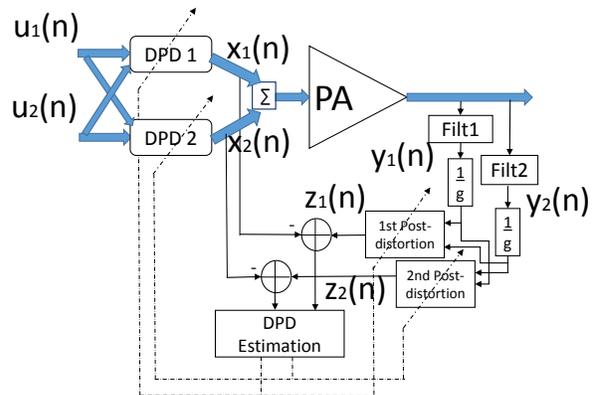

Fig. 1. Dual-band system architecture

In this paper, we use an OH-based method to determine an optimal pruned structure for the 2-D MP model. To improve convergence and reduce computation time, we propose a new form for the 2-D MP model. Besides, a comparison against the CS-based pruning technique is given. In the simulations, the stimulus is composed of two 20 MHz long term evolution (LTE) signals on different bands.

This paper is organized as follows. Section II presents the structure of the DPD. The OH-based method proposed to size the 2-D DPD model structure is introduced and compared with CS-based method in Section III. The simulation results are given in Section IV. Finally, the conclusion is given in Section V.

II. DPD STRUCTURE

A. Model Structure

The dual-band system is depicted in Fig 1. The lower and upper band signals are denoted by $u_1(n)$ and $u_2(n)$ respectively. The feedback signals $y_1(n)$ and $y_2(n)$ can be also extracted from the PA output with filters selecting the corresponding frequency band. In this paper, the DPD uses a 2-D MP model [2]. A post-distortion is first identified and then applied upstream of the PA as the DPD. The i-th band model coefficients are estimated using the PA input signal $x_i(n)$ and the filtered PA output signal $y_i(n)$ normalized by the chosen gain g. Denoting the normalized signal by $\tilde{y}_i$, the post-distortion output is expressed as a function of both normalized

signals:

$$z_i(n) = \sum_{k=0}^{\mathcal{K}-1} \sum_{j=0}^{k} \sum_{l=0}^{\mathcal{L}-1} \gamma_{kjl}^{(i)} \tilde{y}_i(n-l)$$
$$\times |\tilde{y}_i(n-l)|^{k-j} |\tilde{y}_{3-i}(n-l)|^j \qquad (1)$$
$$= f_i(\tilde{y}_1(n), \tilde{y}_2(n))$$

where k is the nonlinearity index, l the memory index, and the γ_{kjl}'s are the complex coefficients. As different possible choices for g achieve about the same PA efficiency [8], we choose the small signal gain in this paper.

B. Model Identification

The identification is based on indirect learning architecture (ILA) [3] so that we can take the average of the two bands normalized mean square error (NMSE) between $x_i(n)$ and $z_i(n)$ to represent the modeling accuracy:

$$Y = 10 log_{10} \left[\frac{\text{NMSE}_1 + \text{NMSE}_2}{2} \right] \qquad (2)$$

where for $i = 1, 2$

$$\text{NMSE}_i = \frac{\sum_{n=1}^{N} |x_i(n) - z_i(n)|^2}{\sum_{n=1}^{N} |x_i(n)|^2}. \qquad (3)$$

The model coefficients can be estimated by solving a linear problem. For each band, we can express the post-distortion using matrix notation for a block of N samples:

$$\boldsymbol{z_i} = \boldsymbol{\Psi}_i \boldsymbol{\gamma_i} \qquad (4)$$

where $\boldsymbol{z_i} = [z_i(1), \ldots, z_i(N)]^T$, $\boldsymbol{\Psi}_i$ is the $N \times R$ matrix containing the basis functions and R is the total number of model coefficients. The LS estimation of $\boldsymbol{\gamma_i}$ is found by

$$\hat{\boldsymbol{\gamma_i}} = [\boldsymbol{\Psi}_i^H \boldsymbol{\Psi}_i]^{-1} \boldsymbol{\Psi}_i^H \boldsymbol{x_i} \qquad (5)$$

which minimizes the cost function (3).

III. MODEL COMPLEXITY REDUCTION METHODS

A. Optimization-heuristic-based method

For the full model with nonlinearity orders $\boldsymbol{k} = [0 : \mathcal{K} - 1]$ and memory depth $\boldsymbol{l} = [0 : \mathcal{L} - 1]$, the number of coefficients equals to $\frac{1}{2}\mathcal{K}(\mathcal{K}+1)\mathcal{L}$. Pruning the model, especially pruning the nonlinearity arrays $\boldsymbol{k}$, can sharply decrease the model complexity.

Hill-climbing heuristic is an optimization algorithm with a rapid progress converging to a local optimum. It has been confirmed robust for DPD model structure optimization since its local optima are very close to or even the same as the global optimum [3].

We propose to use this kind of optimization heuristic to determine the 2D-MP model structure. At the 1st iteration starting point $M_{0(1)}$ is chosen and its merit value $J(M_{0(1)})$ is computed as well as those of its neighbors. If the best neighbor element $M_{s(1)}$ is better than $M_{0(1)}$, it is taken as the starting point for the next iteration $M_{0(2)}$. Otherwise $M_{0(1)}$ is taken as the final solution and the procedure ends.

For the hill-climbing method to be effective, it is important to define relevant neighbours. To this end, we need here to characterize the sparsity in a 2-D DPD model. It is quite easy to characterize the sparsity in a vector. S can be used to denote that $S-1$ elements are selected from the $\mathcal{K}-1$ element array $[1 : \mathcal{K}-1]$ (element 0 is always in the array $\boldsymbol{k}$ and $\boldsymbol{l}$). All possible combinations of $S-1$ elements of vector $[1 : \mathcal{K}-1]$ can be listed in increasing order. Then each combination is indicated by its index I and its sparsity S. Since parameters k and l are independent of each other, they can be treated in the same way.

Noticing the range of values for parameter j in (1) varies according to k, it is more difficult to list the sparse arrays in the same way. Considering this, we propose to rewrite (1) as follows:

$$z_i(n) = \sum_{m=0}^{\mathcal{K}-1} \sum_{\substack{j=0 \\ m+j \leqslant \mathcal{K}-1}}^{\mathcal{K}-1} \sum_{l=0}^{\mathcal{L}-1} \gamma_{mjl}^{(i)} \tilde{y}_i(n-l)$$
$$\times |\tilde{y}_i(n-l)|^m |\tilde{y}_{3-i}(n-l)|^j, \qquad (6)$$

The structure of the 2-D MP model can then be represented by 6 parameters: $(S_m, I_m; S_j, I_j; S_l, I_l)$. Thus we can construct a 6-D discrete space where each point represents a model structure. We define two constraints for M_d to be a neighbor of M_i:

1) The parameters of M_d can be represented by $(S_m^{(i)} \pm 1, I_m^{(i)} \pm 1; S_j^{(i)} \pm 1, I_j^{(i)} \pm 1; S_l^{(i)} \pm 1, I_l^{(i)} \pm 1)$.
2) The total variation of the number of coefficients $|R_d - R_i| \leqslant d$, where d is a given threshold.

B. Compressed-sensing-based method

We take the orthogonal matching pursuit (OMP) algorithm as a reference in this paper to solve the CS problem. First, for the PA output signal filtered in the i-th band, we construct the basis matrix $\boldsymbol{\Psi_i}$ of the complete model M_{comp} with $\mathcal{K}$ and $\mathcal{L}$ as the maximal nonlinearity order and memory depth respectively. The model $M_{(q)}$ contains q basis functions. At the beginning, $M_{(0)}$ is an empty set. We select the most important basis function from $\boldsymbol{\Psi_i}$ and add it to our model $M_{(q)}$ at each iteration. Thus q represents also the number of iterations. The importance of each basis function is determined by the product

$$I(c) = \frac{|\boldsymbol{\Psi}_c^H \boldsymbol{r}_{(q-1)}|}{||\boldsymbol{\Psi}_c||} \qquad (7)$$

where $\boldsymbol{\Psi}_c$ is the column vector for the c-th basis function and $\boldsymbol{r}_{(q-1)}$ is the residual at the $(q-1)$-th iteration. The residual is initialized as $\boldsymbol{r}_{(0)} = \boldsymbol{z_i}$ and is updated at each iteration with the estimated model output:

$$\boldsymbol{r}_{(q)} = \boldsymbol{r}_{(q-1)} - \boldsymbol{\Psi_i}\hat{\boldsymbol{\gamma}}^i \qquad (8)$$

where the coefficients $\hat{\boldsymbol{\gamma}}^i$ are estimated using (5). The algorithm ends when a given stall condition is reached.
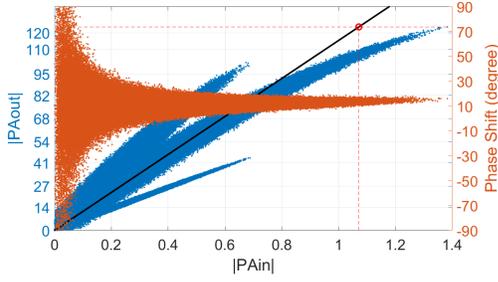
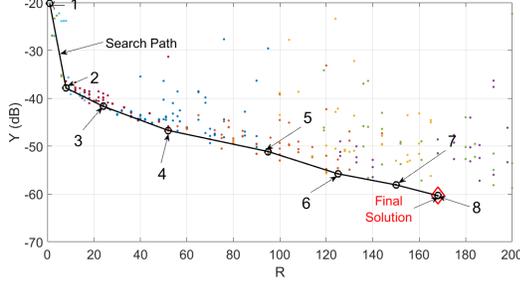Fig. 2. AM/AM & AM/PM curve of Wiener model PA for dual-band LTE

Fig. 4. Search path of New OH-based method (*d*=50)

Fig. 3. Search path of OH-based method for 2D-MP model structure (*d*=50)
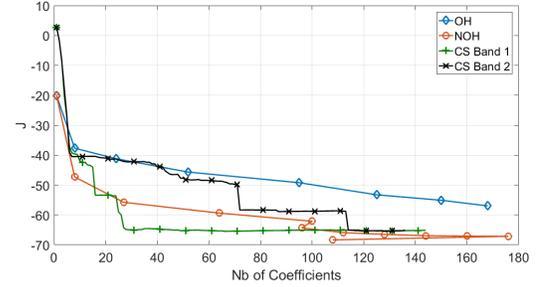
Fig. 5. Evolution of merit value for different methods

C. Merit function

In order to find the optimal model with respect to the trade-off between modeling accuracy and model complexity, we compute merit values according to the generalized information criterion (GIC) [9]. The GIC can be written as

$$GIC = -2\ln p(\boldsymbol{x}, \hat{\boldsymbol{\gamma}}^{(R)}) + (1+\rho)R \qquad (9)$$

where $p(\boldsymbol{x}, \hat{\boldsymbol{\gamma}}^{(R)})$ is the likelihood function of the data vector $\boldsymbol{x}$ which depends on the $1 \times R$ parameter vector $\hat{\boldsymbol{\gamma}}^{(R)}$. In the case of a 2D-MP model, we have in the i-th band

$$\begin{aligned} -2\ln p(\boldsymbol{x}_i, \hat{\boldsymbol{\gamma}}_i^{(R)}) &= N\ln(\frac{1}{N}||x_i(n) - z_i(n)||^2) + cst \\ &= \frac{N\ln 10}{10}\text{NMSE}_i + cst \end{aligned} \qquad (10)$$

where cst is a constant. Considering the performances in both bands, we propose the merit function as following:

$$J = Y + \frac{10 \cdot (1+\rho)}{N\ln 10}R = Y + \mu R. \qquad (11)$$

Since the importance of a basis function can be computed for the signal of only one band at a time, the CS-based method determines the model structure separately for the two bands and thus they can be different. We take $R = \frac{1}{2}(R_1 + R_2)$ in (11) for the merit value of the CS-based method solution.

IV. SIMULATION RESULTS

We create a baseband dual-band signal with 100 MHz frequency separation as the stimulus. The signal in each band is a 20 MHz LTE signal. The signal is fed to a Wiener model PA. The AM/AM & AM/PM (Amplitude Modulation/Amplitude Modulation & Amplitude Modulation/Phase Modulation) curves of the PA are illustrated in Fig 2.
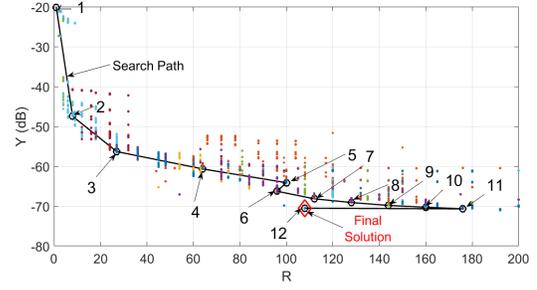
In the following tests, we use a dataset of $N = 30000$ samples. Considering the scenario as well as the value of N, we set $\mu = 0.02$ for the merit function which represents that we trade-off 1dB on NMSE with 50 coefficients. Computation times are given for an Intel Core i7-7500U CPU @ 2.70GHz.

The simulation results for the OH-based method and the CS-based method are listed and compared in Table I. The OH-based method is first applied on the original form of the 2-D DPD model (1), which is denoted by "OH". We set $d = 50$ for the neighborhood constraint. The OH-based method is then applied on the 6-D parameter space corresponding to the new proposed form (6), which is denoted by "NOH". The maximum orders are set in bth cases as $\mathcal{K}=\mathcal{L}=15$.

As the modeling accuracy is represented by the error between the post-distortion output $z_i(n)$ and the stimulus $u(n)$ as (3), the spectra of $z_1(n)$ and $z_2(n)$ given by different methods are depicted along with the spectrum of $u(n)$ in Fig 6 and Fig 7 respectively. The lower and upper adjacent channel power ratios (ACPR.L/U) for each band are also given in Table I.

1) Results for the OH-based method: The search path of the proposed OH-based method is shown in Fig 3. The colored points are tested models. The black circles represent solutions at each iteration (the iteration number is indicated beside), which trace the search path of the algorithm converging towards the red diamond. The final solution has R=168 coefficients and Y=-60.3 dB, $\boldsymbol{k}$=[0:4, 8], $\boldsymbol{l}$=[0:5, 8]. 323 models were tested in this case for a total computation time of 9.6 minutes.

2) Results for the new OH-based method (NOH): The search path of the NOH method is shown in Fig 4. The final solution has R=108 coefficients and Y=-70.5 dB, $\boldsymbol{m}$=[0:2],
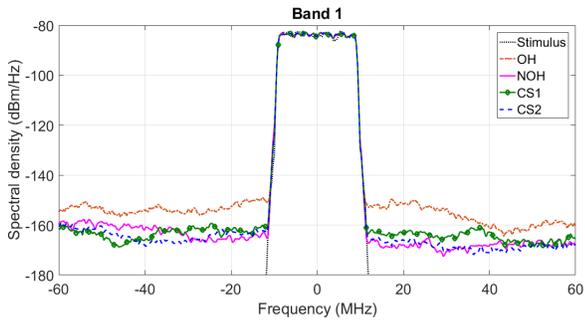
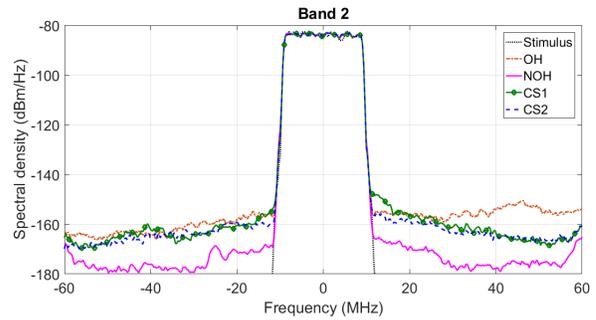Fig. 6. Spectra of the post-distortion output in the 1st band

Fig. 7. Spectra of the post-distortion output in the 2nd band

j=[0:2], l=[0:10, 12]. 2435 models were tested in this case for a total computation time of 90.9 minutes.

3) Results for the new OH-based method with d=25 (NOH$_{\frac{1}{2}}$): Since NOH has 6-D search space, the number of models tested is greater than that of OH. To compare both methods with a similar computation time, we test NOH with a reduced neighborhood radius $d = 25$ with results given in Table I. The final solution has R=48 coefficients and Y=-66.7 dB, m=[0:1], j=[0:2], l=[0:7]. 1345 models were tested in this case for a total computation time of 12.8 minutes.

4) Result for the CS-based method: We set $\mathcal{K}=\mathcal{L}$=15 the orders of the complete model M_{comp} same as in OH and NOH. We test the algorithm with different stall conditions. In Table I and Fig 6-7, To compare with the NOH solution "CS1" ends when R=108. "CS2" ends when $J(M_{(q)})$ better than $0.9 \times J(M_{(q-1)})$ at the q-th iteration.The solution given by CS1 reaches an NMSE value of Y=-62.9 dB. The computation time is 47.1 minutes. The solution given by CS2 reaches an NMSE value of Y=-67.9 dB with 144 coefficients for the 1st band and 136 for the 2nd band. The computation time is 54.6 minutes.

The evolution of merit values versus the number of coefficients is illustrated in Fig 5. The proposed NOH methods can reach better merit values than OH and CS methods.

Compared with the conventional OH method, NOH improves the linearization performance by about 10 dB with less coefficients. Comparing Fig. 3 and Fig. 4, we can see that equation (6) allows NOH to test models with NMSE lower than -60 dB and coefficients less than 100. The solution given by CS1 has the same number of coefficients than that of NOH but its linearization performance is very weak. The solution given by CS2 achieves the best linearization performance among the CS-based solutions. But it is still worse than the NOH solution even though with more coefficients.

V. CONCLUSION

In this paper, we extend the search algorithm proposed in [5] to a 2-D MP model. To trade-off model complexity and accuracy, a search criterion based on GIC is used. We propose a new form of 2-D MP model which improves convergence and leads to a more accurate and less complex model. A comparison against the CS-based pruning technique validates the advantages of the proposed NOH method. Future work includes studying the impact of d on the algorithm and

extending the method to allow for different models on each band.

TABLE I
PERFORMANCES OF DIFFERENT METHODS

		OH	NOH	NOH$_{\frac{1}{2}}$	CS1	CS2
J		-57.0	-68.4	-65.8	-60.7	-65.1
R (Band 1/2)		168/168	108/108	48/48	108/108	144/136
Y (dB)		-60.3	-70.5	-66.7	-62.9	-67.9
ACPR (dBc)	1.L	-65.5	-77.8	-76.1	-74.2	-76.2
	1.U	-65.4	-80.6	-79.3	-76.1	-79.2
	2.L	-70.9	-84.0	-76.3	-73.0	-73.8
	2.U	-69.4	-82.2	-72.0	-68.7	-72.1
Runtime (min)		9.6	90.9	12.8	47.1	54.6

NOH$_{\frac{1}{2}}$: NOH with d=25.

REFERENCES

[1] W. Chen, S. A. Bassam, X. Li, Y. Liu, K. Rawat, M. Helaoui, F. M. Ghannouchi, and Z. Feng, "Design and linearization of concurrent dual-band doherty power amplifier with frequency-dependent power ranges," *IEEE Trans. Microw. Theory Tech.*, vol. 59, no. 10, pp. 2537–2546, Oct 2011.

[2] S. Bassam, M. Helaoui, and F. Ghannouchi, "2-d digital predistortion (2-d-dpd) architecture for concurrent dual-band transmitters," *IEEE Trans. Microw. Theory Techn.*, vol. 59, no. 10, pp. 2547 –2553, oct. 2011.

[3] S. Wang, M. Abi Hussein, O. Venard, and G. Baudoin, "A novel algorithm for determining the structure of digital predistortion models," *IEEE Trans. Veh. Technol*, vol. 67, no. 8, pp. 7326–7340, Aug 2018.

[4] A. Abdelhafiz, L. Behjat, and F. M. Ghannouchi, "Generalized memory polynomial model dimension selection using particle swarm optimization," *IEEE Microw. Wireless Compon. Lett.*, vol. 28, no. 2, pp. 96–98, Feb 2018.

[5] S. Wang, M. Abi Hussein, O. Venard, and G. Baudoin, "Optimal sizing of two-stage cascaded sparse memory polynomial model for high power amplifiers linearization," *IEEE Trans. Microw. Theory Tech.*, vol. 66, no. 9, pp. 3958–3965, Sept 2018.

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

[7] S. Wu, J. Zhai, Z. Yu, and N. Zhang, "A compressed sensing method for pruning concurrent dual-band power amplifier models," in *2018 IEEE MTT-S International Wireless Symposium (IWS)*, May 2018, pp. 1–3.

[8] S. Wang, M. Abi Hussein, O. Venard, and G. Baudoin, "Impact of the normalization gain of digital predistortion on linearization performance and power added efficiency of the linearized power amplifier," in *2017 12th European Microwave Integrated Circuits Conference (EuMIC)*, Oct 2017, pp. 310–313.

[9] P. Stoica and Y. Selen, "Model-order selection: a review of information criterion rules," *IEEE Signal Process. Mag.*, vol. 21, no. 4, pp. 36–47, July 2004.