Digital predistortion of power amplifiers using structured compressed-sensing Volterra series

Juan Antonio Becerra-González✉, María José Madero-Ayora, Javier Reina-Tosina and Carlos Crespo-Cadenas

Digital predistortion has become an attractive technique for power amplifier linearisation whose limiting factor for using Volterra series as the underlying model is its computational complexity, since the number of components rapidly grows with the non-linear order and memory. Based on a previous reference algorithm, which consists on applying the orthogonal matching pursuit for the sorting of the model components and a Bayesian information criterion for the selection of the optimum number of components, a new technique to reduce the size of the support set taking into account the structural information within a model is presented. Experimental results of the predistortion of a commercial power amplifier are given as a proof of its capabilities, showing equivalent performance to the pruning with the reference algorithm while further reducing the number of components.

Introduction: Recent advances in digital communication standards require the management of the trade-off between efficiency and non-linearity. Digital predistortion comes up as a solution that allows the power amplifier (PA) to operate near saturation, mitigating the distortion created in new standard signals such as orthogonal frequency division multiplexing (OFDM), which are characterised by a high peak-to-average power ratio (PAPR). Digital predistorters (DPDs) rely on behavioral models [1], which usually take the form of Volterra series. The most general full Volterra (FV) [2] series has a rich set of terms to represent the modelled system, but this number of components is usually very large due to its inherent structure. Given the limited real-time computational capability of nowadays field-programmable gate arrays, as it is desirable to keep an easy-to-manage number of components, researchers have developed a set of pruning strategies – also called sparse recovery techniques –. These strategies can be either ad-hoc, which include a subset of the FV such as the generalised memory polynomial (GMP) or dynamic deviation reduction among others, or based on information theory, which do not include the information of the intrinsic structure of the model [3]. Structural information based on the algorithm in [4] was incorporated by the authors in [5]. In this Letter, we show the structural pruning of Volterra series and validate the method in the DPD application, obtaining a reduced complexity model while keeping the level of performance.

This Letter is organised as follows. The algorithm is first formulated. Then, the experimental design of a DPD application for a commercial PA is presented and results are discussed. Finally, conclusions are drawn.

Structured compressed-sensing for Volterra series models: The structural compressed-sensing algorithm presented in this Letter can be considered a particularisation of the stagewise orthogonal matching pursuit algorithm, which selects a fixed number of regressors in each iteration based on a threshold. The improvement consists on the inclusion of a priority function that assigns the significance of the coefficient within the model in the subset defined after the thresholding. The new greedy algorithm for pruning Volterra model matrices taking into account the structural information is summarised in Algorithm 1.

The initialisation, which corresponds to line 1 of Algorithm 1, consists on the definition of the residual $\boldsymbol{r}^{(0)} = \boldsymbol{y}$, that will be used for keeping the remaining part of the output still to be modelled. The support set $S(0)$ is empty in the first iteration, as no regressor is still selected. In each iteration t, the algorithm calculates the correlation between the residual $\boldsymbol{r}^{(t)}$ and each of the columns of the measurement matrix $\boldsymbol{X}$ normalised by its ℓ_2-norm (line 3). A first preselection is performed where all the regressors with absolute value of the correlation greater than a fraction $(1-\alpha) \in [0, 1]$ of the maximum are included in the subset $\boldsymbol{i}_{\text{pre}}^{(t)}$, shown in line 4. When the span α is equal to 0, the selection becomes that of the classic OMP, which chooses only the maximum correlation within all the regressors, and if it is equal to 1, no correlation-based sorting is made and only the structural information of the model is evaluated for this arrangement. Then, the regressor with the lowest score given by the priority function $f(\cdot)$ is included onto the support set. Then, the estimation of the Volterra vector $\hat{\boldsymbol{h}}$ is obtained by a least-squares regression and the estimated output of the model $\hat{\boldsymbol{y}}^{(t)}$ and the residual $\boldsymbol{r}^{(t)}$ are updated (lines 7–9). An example of this function is given in the experimental design section. Finally, when all the regressors are sorted or a fixed maximum of regressors to sort $n_{\max}$ has been reached, a Bayesian information criterion (BIC) is applied to obtain the optimum number of Volterra kernels n_{opt}. The model with the lowest BIC is selected according to line 12, where $\hat{\sigma}_{\text{e}}^2$ is the estimation of the error variance and n_{c} is the number of components.

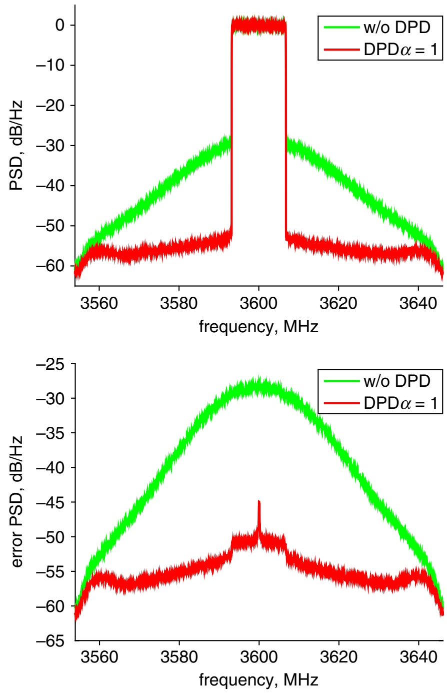

Fig. 1 *Spectral densities for one realisation of the experiment for the cases without DPD and with DPD in the case of $\alpha = 1$*

Algorithm 1: Summary of the structured compressed-sensing algorithm for Volterra series models

Input: $n_{\max} > 0$, $\alpha \in [0, 1]$, $f(\cdot)$, $\boldsymbol{X} \in \mathbb{C}^{m\times n}$, $\boldsymbol{y} \in \mathbb{C}^m$
Output: $S(t)$, n_{opt}
Initialisation:
1: $\boldsymbol{r}^{(0)} \leftarrow \boldsymbol{y}$, $S(0) \leftarrow \emptyset$
2: **for** $t = 1$ to $n_{\max}$ **do**
3: $\theta_t \leftarrow (1-\alpha)\cdot \max_{i\notin S(t)} \dfrac{\left|\boldsymbol{X}_{\{i\}}^{\text{H}} \cdot \boldsymbol{r}^{(t-1)}\right|}{\|\boldsymbol{X}_{\{i\}}\|_2}$
4: $\boldsymbol{i}_{\text{pre}}^{(t)} \leftarrow \left\langle i \,\middle|\, \dfrac{\left|\boldsymbol{X}_{\{i\}}^{\text{H}} \cdot \boldsymbol{r}^{(t-1)}\right|}{\|\boldsymbol{X}_{\{i\}}\|_2} > \theta_t \right\rangle$
5: $i^{(t)} \leftarrow \arg\min_{i\in \boldsymbol{i}_{\text{pre}}^{(t)}} f\left(\boldsymbol{X}_{\{i\}}\right)$
6: $S(t) \leftarrow S(t) \cup i^{(t)}$
7: $\hat{\boldsymbol{h}} \leftarrow \left(\boldsymbol{X}_{S(t)}^{\text{H}} \boldsymbol{X}_{S(t)}\right)^{-1} \boldsymbol{X}_{S(t)}^{\text{H}} \boldsymbol{y}$
8: $\hat{\boldsymbol{y}}^{(t)} \leftarrow \boldsymbol{X}_{S(t)} \hat{\boldsymbol{h}}$
9: $\boldsymbol{r}^{(t)} \leftarrow \boldsymbol{y} - \hat{\boldsymbol{y}}^{(t)}$
10: **end for**
11: $\hat{\sigma}_{\text{e}}^2 = \|\boldsymbol{y} - \hat{\boldsymbol{y}}^{(n_{\max})}\|_2^2$
12: $n_{\text{opt}} \leftarrow \arg\min_{n_{\text{c}}} \left\{2m \ln \hat{\sigma}_{\text{e}}^2 + 2n_{\text{c}} \ln(2m)\right\}$

Experimental design: A set of measurements were acquired in order to validate the predistortion capabilities of the algorithm. The weblab for PA digital predistortion and characterisation by Chalmers University of Technology was used [6]. The weblab setup consists of the PXI chassis PXIe-1082 from National Instruments Inc. The device under test was a GaN PA CGH40006P in its test board. The measurements were taken at an output power of +35.2 dBm. The test signal is compound of 56 OFDM symbols of a 15-MHz OFDM signal generated from 16-QAM symbols modulated onto 900 subcarriers and filtered by a raised-cosine with roll-off factor of 0.1, according to the LTE-downlink standard. This signal exhibits a PAPR of about 11 dB and a hard clipping to the seven samples with the highest absolute value was applied to reduce the PAPR to 10 dB. An oversampling of 1 to 6 applied to the original signal results in a sampling frequency of 92.16 MHz which was also used to record the measurements. The test signal contains over 360,000 samples.

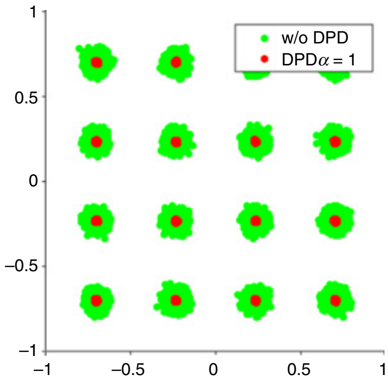

Fig. 2 *15-MHz bandwidth signal constellation for original signal without DPD and with DPD in the case of* $\alpha = 1$

Table 1: Performance results of DPD in a sweep of span α values

Case		ACPR-2, dBc	ACPR-1, dBc	ACPR +1, dBc	ACPR +2, dBc	EVM, %	NMSE, dB	# Coeff.
w/o DPD		−38.04	−32.82	−32.93	−38.00	5.00	−25.52	–
α	0.00	−39.76	−36.83	−36.82	−39.80	2.82	−30.46	200
	0.20	−42.57	−40.46	−40.51	−42.75	1.89	−33.90	198
	0.40	−44.15	−41.35	−41.29	−44.17	1.78	−34.49	190
	0.60	−44.65	−42.92	−42.91	−44.74	1.46	−36.18	148
	0.80	−50.04	−48.58	−48.71	−50.24	0.80	−41.52	17
	1.00	−55.09	−53.88	−54.13	−55.37	0.52	−45.62	194

A GMP model was selected to test the algorithm. This model has the structure shown in (1), where a configuration of seventeenth order and a maximum distance from the diagonal of 10 was set. This corresponds to $K_a = K_b = K_c = 16$, $L_a = 10$ and $L_b = M_b = L_c = M_c = 5$. The resulting Volterra model contains 1087 components.

$$y(k) = \sum_{p=0}^{K_a}\sum_{l=0}^{L_a} a_{pl}x(k-l)|x(k-l)|^p + \sum_{p=1}^{K_b}\sum_{l=0}^{L_b}\sum_{m=1}^{M_b} b_{plm}x(k-l)|x(k-l-m)|^p + \sum_{p=1}^{K_c}\sum_{l=0}^{L_c}\sum_{m=1}^{M_c} c_{plm}x(k-l)|x(k-l+m)|^p. \quad (1)$$

The thresholding function designed for this experiment was $f(l, m, p) = |l \mp m| + (p + 1)$ according to the form of the GMP regressors of $x(k-l)|x(k-l \pm m)|^p$. This function assigns a higher score – which corresponds to less priority – to high orders and lags, considering a memory fading behaviour [7]. The number of maximum components to be considered $n_{\max}$ was set to 200.

The predistorter was calculated through the indirect learning scheme [8]. The 30% of the signal with the highest maximum value at the output of the PA was used for modelling and the span value from 0 to 1 with an increment of 0.2 was swept. Once the predistorter was identified, the validation was carried out with the complete signal and the predistorted signal was obtained placing the previous PA input signal at the input of the DPD. The output of the DPD was then sent to the web platform and the performance parameters of normalised mean square error (NMSE), adjacent channel power ratio (ACPR) and error vector magnitude (EVM) were measured in the returned signal.

DPD performance: The performance parameters of the DPD are presented in Table 1. The ACPR, EVM and NMSE values experience a decrease with the span, which indicates that the introduction of the span enhances the predistortion capabilities of the model. The signal without predistortion is characterised by a first lower and upper ACPRs of −32.82 and −32.93 dBc. The predistortion enhances this values from about 4 dB for $\alpha = 0$ to 21 dB in the case of $\alpha = 1$, where the minimum values of −53.88 and −54.13 dBc are reached. Similar improvements are achieved for the second lower and upper ACPRs, where the minimum is also obtained in the same case and take the values of −55.09 and −55.37 dBc, corresponding to an improvement of about 17 dB with respect to the case without DPD. These values are in accordance with Fig. 1, where the power spectral density (PSD) of the predistorted signal for the best case is compared with the case without DPD. The reduction of the spectral regrowth achieved by the DPD agrees with the PSD of the error signal, which shows a low in-band and out-of-band error densities. The EVM is decreased from a 5% to a 2.8% for $\alpha = 0$ and a 0.52% in the case of $\alpha = 1$. This enhancement is clear in the constellation shown in Fig. 2, where the predistorted case shows a reduction in dispersion compared with the signal without predistorsion. The NMSE is reduced from −25.52 dBm without predistortion to −30.45 dB for the lowest value of the span and to −45.62 dB for $\alpha = 1$. Finally, the number of components in the pruned model results in roughly the same value for all the range of the span, which indicates that the pruning capabilities of the algorithm remains at the same performance level.

Conclusion: An improved method for the sparse recovery of Volterra series models has been presented. It adds structural information of the PA model in the selection process to achieve a reduction in the optimum number of coefficients while maintaining the fidelity. The benefits of the introduction of the span parameter has been experimentally demonstrated. This method is susceptible to be applied to any kind of Volterra series behavioral model. Predistortion with models given by the algorithm have been performed showing a high reduction of the complexity of the model with the same level of performance that the ones given by the complete model before pruning.

Acknowledgments: This work was partially supported by the Ministry of Economy and Competitiveness of Spain under Projects TEC2011-23559 and TEC2014-53103-P. The authors thank the RF WebLab team at the Chalmers University of Technology for having a web-accessible experimental setup open to everyone.

Submitted: *21 October 2016* E-first: *8 December 2016*
doi: 10.1049/el.2016.3879
One or more of the Figures in this Letter are available in colour online.

Juan Antonio Becerra-González, María José Madero-Ayora, Javier Reina-Tosina and Carlos Crespo-Cadenas (*Universidad de Sevilla, Sevilla, Spain*)

✉ E-mail: jabecerra@us.es

References

1 Kim, J., and Konstantinou, K.: 'Digital predistortion of wideband signals based on power amplifier model with memory', *Electron. Lett.*, 2001, **37**, pp. 1417–1418
2 Lima, E., Cunha, T., Teixeira, H., Pirola, M., and Pedro, J.: 'Base-band derived Volterra series for power amplifier modeling'. Proc. 2009 IEEE MTT-S Int. Microwave Symp. Digest (IMS 2009), June 2009, pp. 1361–1364
3 Abdelhafiz, A., Kwan, A., Hammi, O., and Ghannouchi, F.: 'Digital predistortion of LTE-A power amplifiers using compressed-sampling-based unstructured pruning of Volterra series', *IEEE Trans. Microw. Theory Tech.*, 2014, **62**, pp. 2583–2593
4 Reina-Tosina, J., Allegue-Martinez, M., Crespo-Cadenas, C., Yu, C., and Cruces, S.: 'Behavioral modeling and predistortion of power amplifiers under sparsity hypothesis', *IEEE Trans. Microw. Theory Techn.*, 2015, **63**, pp. 745–753
5 Becerra, J.A., Madero-Ayora, M.J., Reina-Tosina, J., and Crespo-Cadenas, C.: 'Structured compressed-sensing for Volterra series models'. Proc. 2016 IEEE 17th Annual Wireless Microwave Technology Conf., April 2016, pp. 1–4
6 Landin, P., Gustafsson, S., Fager, C., and Eriksson, T.: 'Weblab: A web-based setup for PA digital predistortion and characterization [application notes]', *IEEE Microw. Mag.*, 2015, **16**, pp. 138–140
7 Staudinger, J., Nanan, J.C., and Wood, J.: 'Memory fading volterra series model for high power infrastructure amplifiers'. Proc. 2010 IEEE Radio and Wireless Symp. (RWS), January 2010, pp. 184–187
8 Eun, C., and Powers, E.J.: 'A new Volterra predistorter based on the indirect learning architecture', *IEEE Trans. Signal Process.*, 1997, **45**, pp. 223–227