2022 IEEE Topical Conference on RF/Microwave Power Amplifiers for Radio and Wireless Applications (PAWR) | 978-1-6654-3472-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/PAWR53092.2022.9719839

Acceleration of Digital Pre-Distortion Training Using Selective Partitioning

Méabh Loughman, Declan Byrne, Ronan Farrell and John Dooley
Maynooth University, Ireland
Email: meabh.loughman.2018@mumail.ie

***Abstract*—In recent years model and Digital Pre-Distortion dimension reduction has been widely researched. The operations involved when running DPD are often far less than those needed during the training of the DPD coefficients. The proposed partitioned Least Squares (LS) adaptation allows a selected subset of DPD coefficients to be updated while the remaining coefficients are held constant. This technique allows a more adaptive training procedure, improved interpretability of the important DPD coefficient's during training and the ability to partition the DPD function into specific groups. The Frisch-Waugh-Lovell (FWL) theorem is exploited to partition the coefficients of a DPD basis function trained using LS regression. The proposed methodology was experimentally validated with a Generalized Memory Polynomial (GMP) DPD function, used to linearize a 5W power amplifier (PA) driven by a 40MHz 5G-NR signal.**

I. Introduction

The complexity of DPD solutions has grown considerably in recent years as the performance demands of cellular network communications become ever more challenging. Millimetre-Wave (mmWave) communications over 28GHz have become standardized with 5G. Additionally, more complex configurations of PA architectures such as MIMO and beamforming to improve the overall system performance at these frequencies. Substrates such as GaN also contribute to distortion and memory effects in PAs that have to be mitigated [1]. PA characteristics may also change depending on input stimulus. Therefore with arrays of PAs experiencing continuous changes in operating conditions, ideally DPD parameters must also be continuously adapted.

Several nonlinear dynamic structures have been derived which can accomplish DPD for PAs. The objective of DPD is to derive a concise model, inverse to a PA's characteristic nonlinearity. The DPD solution should ideally produce a highly linear PA output signal using the minimum number of coefficients. Overfitting and ill-conditioning is an area of concern when training DPD coefficients [2]. Feature selection and/or extraction has been a popular research topic concerning the reduction of DPD function dimension to avoid overfitting.

DPD coefficient estimation is typically performed using an iterative optimisation algorithm such as the LS. Typically DPD is over determined. Polynomial-based DPD exhibits structural multicollinearity between predictors, enabling researchers to intelligently prune DPD coefficients that do not contribute to the efficacy of the DPD linearisation [3][4]. Authors of [5] allow for function reduction and a change of basis by employing Principal Component Analysis (PCA).

Authors of [6] use a partial least squares (PLS) algorithm, which allows for the basis matrix used in the DPD estimation to be transformed at every iteration. The orthogonal matching pursuit algorithm was used to determine which basis functions contributed most to the DPD adaption. The drawback of these aforementioned methods is the complex sorting algorithms necessary in order to rank effectiveness of each coefficient on their respective models.

In this paper, we introduce a method for selectively partitioning the DPD coefficient updates. Basis functions within a DPD model can be selectively partitioned, and updated separately. The rationale of updating only a portion of the DPD coefficients is that all coefficients when calculated do not converge uniformly over training iterations [7]. The Frisch–Waugh–Lovell (FWL) theorem [8] can be applied to allow a local update of a single partition while the remaining coefficients were held constant. The rest of the paper is as follows: Section II describes the concept of DPD and the LS training process for a typical DPD function. Theory explaining the method for basis function partitioning and local updates is then introduced. Section III details the results of the aforementioned technique. Section IV concludes on the experimental results presented in Section III.

II. Digital Pre-Distortion

DPD is performed such that, the pre-distorted signal, u, is calculated by weighting a set of signal permutations, X, of the PA input signal, x, with a set of calculated DPD coefficients, h. The permutations are commonly derived from the Volterra series as shown in (1) [9], [10].

$$\begin{aligned} \mathrm{y}_{GMP}(n) &= \textstyle\sum_{m=0}^{M}\sum_{k=0}^{K} h_{mk}x(n-m) \mid x(n-m) \mid^{k-1} \\ &+ \textstyle\sum_{m=0}^{M}\sum_{k=0}^{K}\sum_{p=0}^{P} h_{mkp}x(n-m) \mid x(n-m-p) \mid^{k-1} \\ &+ \textstyle\sum_{m=0}^{M}\sum_{k=0}^{K}\sum_{q=0}^{Q} h_{mkq}x(n-m) \mid x(n-m+q) \mid^{k-1} . \end{aligned} \quad (1)$$

Where P is the lagging envelope parameter, Q is the leading envelope parameter, K the nonlinearity order and M memory depth . The LS algorithm aims to find a best fit set of coefficients, $\hat{h}$. An Indirect Learning Architecture (ILA) DPD system is utilised for the experimental validation of this paper. ILA this is defined as the minimisation problem (2).

$$\hat{h} = ||u - Yh||^2. \tag{2}$$

The matrix Y is an m by n matrix of signal permutations of the PA output signal, y, equivalent in selection to those in X. The DPD coefficients are extracted by solving (3).

$$\hat{h} = (Y^H Y)^{-1} Y^H e. \tag{3}$$

The error signal, e, for an ILA is calculated as (4) [11].

$$e = u - \hat{u}. \tag{4}$$

Where the post-distorted signal, $\hat{u}$, can be expressed as (5).

$$\hat{u} = Y\hat{h}. \tag{5}$$

The FWL manipulates the basis function Y as seen in (6), to be segmented in to two or more sections such that

$$Y = Y_1 h_1 + Y_2 h_2. \tag{6}$$

Where Y_1 and Y_2 are the partitioned basis functions of size $n \times k_1$ and $n \times k_2$. The partitioned sets of coefficients h_1 and h_2 are regression coefficients respectively. FWL asserts that it is possible to re-specify a linear regression model by manipulating the residual of only one partition of the basis function. The equation (6) can be re-specified according to the FWL theorem as

$$\hat{u}M_1 = M_1 Y_1 \hat{h_1} + M_1 Y_2 \hat{h_2} + M_1 e. \tag{7}$$

Where $\mathbf{M}_1$ is the residual maker of $\mathbf{Y}_1$, which encapsulates the variation of $\mathbf{Y}_1$ that cannot be resolved by $\mathbf{Y}_2$ and given by (8). $M_1 Y_1 \hat{h_1} = 0$, as the regression of Y_1 on itself yields no variance unexplained by Y_1

$$M_2 = I - Y_2 (Y_2^H Y_2)^{-1} Y_2^H. \tag{8}$$

A partitioned LS regression can now be performed. To do this, (2) can now be solved using FWL such that

$$\hat{h} = (Y_1^H M_2 Y_1)^{-1} Y_1^H M_2 e. \tag{9}$$

In this way only the coefficients from the first or targeted partition are updated. The second partition remain static. LS is a global estimator, employing all coefficients regardless of convergence rate in its estimation as seen in (2). Allowing for reduced computational complexity, i.e LS must perform a matrix inversion costing ($O(C^2 N)$), versus FWL ($O(C^3)$). DPD coeffecients do not converge globally. Partitioning Y supports select coefficients to be updated locally, supporting convergence of the partitioned basis functions, in turn directly influencing the trajectory of the global error.

III. Experimental Results

To validate the proposed technique an experimental test-bench was developed, as shown in Fig. 1. The testbench consisted of a Skyworks SKY66297-11 PA, an Analog Devices AD9375 transceiver board, MATLAB on a local PC and a Rohde & Schwarz FSL spectrum analyser. A 40 MHz 5G-NR signal was passed into the PA and the output signal captured using the observation receiver of the AD9375. The PA output was also attenuated and monitored on the spectrum analyser to confirm correct operation of the PA.

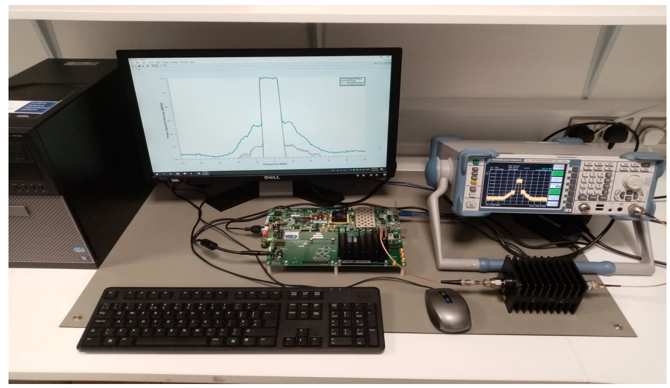

Fig. 1. Experimental testbench

As a reference, a 103 coefficient GMP based DPD function was used to linearise the PA. The LS and LS with FWL methods were both used to train the DPD coefficients and subsequently compared. The LS and FWL method achieved a superior normalised mean square error (NMSE) performance. The number of coefficients updated using the proposed LS with FWL DPD methodology was minimal. The relative performance of the two approaches over two iterations can be seen in Table I. The NMSE, error vector magnitude (EVM) and number of coefficients updated for both the LS and LS with FWL methods are displayed in Table I.

TABLE I
Performance Comparison

Method	LS	FWL
Coefficients Updated	206	107
NMSE (dB)	-40.4597	-44.8611
EVM (%)	.9484	.5714

The experimental results in Table I prove that updating only a select partition of the basis functions can more effectively linearise the dynamic nonlinear PA distortions. By applying LS with FWL to train the DPD coefficients allows for convergence of the partitioned basis functions to achieve a lower error. This is evidenced in the results by the decreased error value of the LS with FWL in Fig. 2. Thus the removal of DPD function coefficients is not needed in order to accelerate the speed of training iterations. Additional computational complexity due to pruning algorithms on the model are avoided. A frequency domain plot of the linearised PA outputs, and the original PA

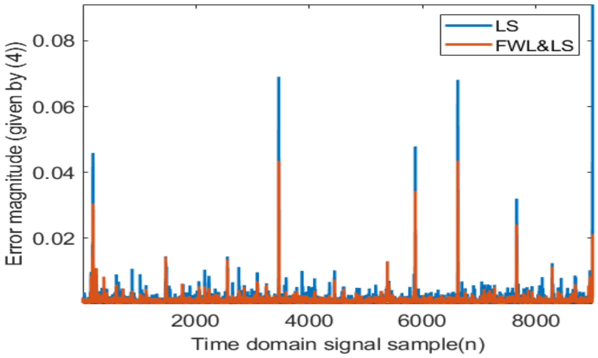

Fig. 2. Error signal for LS and FWL with LS

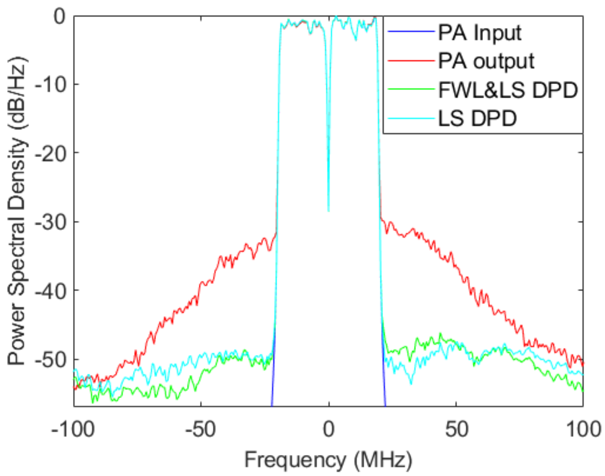

Fig. 3. Frequency domain representation

input and output is shown in Fig. 3. AMAM characteristics for the successful pre-distortion are also presented in Fig. 4.

Figs. 4 and 3 demonstrate the linearisation and experimental validation constructed utilising the LS with FWL and LS linearisation techniques on a 40MHz 5G-NR signal.

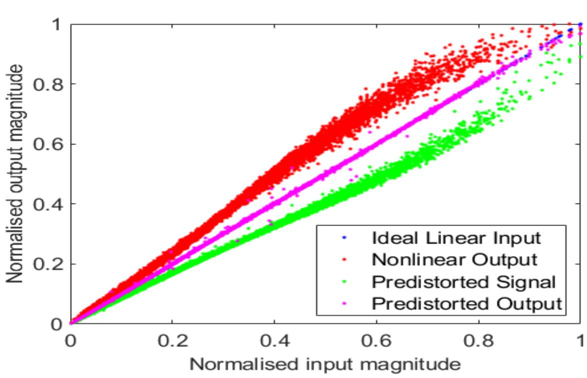

Fig. 4. AM/AM plot

IV. Conclusion

This paper demonstrates a method for selective partitioned adaptation for a LS trained DPD system. The novel methodology presented exploited the Frisch-Waugh-Lovell technique to enable a more accurate adaptation. The partitioned regression was accomplished by applying the FWL theorem to a DPD function trained using LS regression. The proposed methodology was experimentally validated by adapting a particular subset of a GMP DPD function which linearised a PA amplifying a 40MHz 5G-NR signal. The partitioning of model parameters is a research question unto itself. Partitioning specific effects introduced by complex PA architecture may be adapted in further work to allow for partitioning and updating segments of specific basis functions.

Acknowledgment

This publication has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) and is co-funded under the European Regional Development Fund under Grant Number 13/RC/2077.

References

[1] P. M. Tomé, F. M. Barradas, T. R. Cunha, and J. C. Pedro, "A multiple-time-scale analog circuit for the compensation of long-term memory effects in gan hemt-based power amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 68, no. 9, pp. 3709–3723, 2020.

[2] X. Lin, Y. Zhang, H. Li, G. Li, W. Qiao, and F. Liu, "Low computational complexity digital predistortion based on independent parameters estimation," in *2019 IEEE 19th International Conference on Communication Technology (ICCT)*, 2019, pp. 1501–1505.

[3] R. N. Braithwaite, "Pruning strategies for a volterra series model used in digital predistortion (dpd) of rf power amplifiers," in *2017 IEEE Topical Conference on RF/Microwave Power Amplifiers for Radio and Wireless Applications (PAWR)*, 2017, pp. 4–7.

[4] Y.-H. Kim, G. D. Jo, J.-H. Oh, J. H. Jung, J. H. Kim, C. Yu, and K. Lee, "An efficient simplified behavioral model for rf power amplifiers," in *2011 IEEE Topical Conference on Power Amplifiers for Wireless and Radio Applications*, 2011, pp. 65–68.

[5] P. L. Gilabert, G. Montoro, T. Wang, M. N. Ruiz, and J. A. García, "Comparison of Model Order Reduction Techniques for Digital Predistortion of Power Amplifiers," in *2016 46th European Microwave Conference (EuMC)*, Oct 2016, pp. 182–185.

[6] Q. A. Pham, D. López-Bueno, G. Montoro, and P. L. Gilabert, "Dynamic selection and update of digital predistorter coefficients for power amplifier linearization," in *2019 IEEE Topical Conference on RF/Microwave Power Amplifiers for Radio and Wireless Applications (PAWR)*, 2019, pp. 1–4.

[7] Y. Li, X. Wang, and A. Zhu, "Sampling Rate Reduction for Digital Predistortion of Broadband RF Power Amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 68, no. 3, pp. 1054–1064, 2020.

[8] H. J. Newton, C. F. Baum, N. Beck, a. C. Cameron, D. Epstein, J. Hardin, B. Jann, S. Jenkins, and U. Kohler, "The Stata Journal," *Stata Journal*, vol. 10, pp. 288–308, 2010. [Online]. Available: http://ideas.repec.org/a/tsj/stataj/v7y2007i4p465-506.html

[9] J. Kim and K. Konstantinou, "Digital Predistortion of Wideband Signals Based on Power Amplifier Model with Memory," *Electronics Letters*, vol. 37, no. 23, pp. 1417–1418, Nov 2001.

[10] 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 Transactions on Signal Processing*, vol. 54, no. 10, pp. 3852–3860, 2006.

[11] H. Paaso and A. Mammela, "Comparison of Direct Learning and Indirect Learning Predistortion Architectures," in *2008 IEEE International Symposium on Wireless Communication Systems*, Oct 2008, pp. 309–313.