MULTI-STAGE DIGITAL PREDISTORTION BASED ON INDIRECT LEARNING ARCHITECTURE

Mazen Abi Hussein [ξ1], Vivek Ashok Bohara [#2], Olivier Venard [#3]
[ξ] *ESYCOM, EA2552, Université Paris-Est*
[#] *LaMIPS, Laboratoire commun NXP-CRISMAT, UMR 6508 CNRS ENSICAEN UCBN Caen*

ESIEE Paris, Systems Engineering Dept.,
Cité Descartes, BP99, 93162 Noisy-Le-Grand, France
[1]m.abihussein@esiee.fr [2]v.bohara@esiee.fr [3]o.venard@esiee.fr

***Abstract*—In this paper, we propose a multi-stage indirect learning architecture (ILA) for digital predistortion. Unlike the conventional ILA, in multi-stage ILA the digital predistorter (PD) is implemented in two or more stages. We demonstrate that depending on the power amplifier (PA), multi-stage ILA can achieve same or even better performance than the conventional ILA but with significantly lower PD identification complexity. The identification complexity is measured by computing the number of multipliers needed for identification of PD. In addition to above, we also propose two algorithms for the identification of the multi-stage ILA coefficients. The performance of the multi-stage ILA is evaluated in terms of improvement in adjacent channel power ratio (ACPR) and error vector magnitude (EVM) at the output of PA when a Long Term Evolution-Advanced (LTE-Advanced) signal is applied at the input. The reference PA models used for simulation are the Wiener model and Wiener-Hammerstein model.**

***Index Terms*—Digital predistortion, nonlinear distortion, indirect learning architecture, high power amplifiers**

I. Introduction

The PA is one of the most nonlinear components in the radio frequency (RF) transceivers [1]. In order to achieve higher efficiency, PAs are usually driven towards saturation region, however this causes severe distortion on the transmitted signal resulting in out-of-band distortion (spectral regrowth beyond the signal bandwidth) and in-band distortion (Error Vector Magnitude degradation). Moreover, with the advent of multiple carrier transmissions such as 4 carrier-wideband code division multiple access (4x-WCDMA) or Long Term Evolution-Advanced (LTE-Advanced), with varying-envelope waveforms and wide bandwidth the PA also tends to exhibit memory effects [2]. These memory effects further deteriorate the transmitted signals resulting in more out-of-band and in-band distortions.

Over the years numerous authors have proposed many linearization techniques to increase the efficiency and reduce the distortions caused due to PA [1]. However owing to its flexibility of implementation, high performance improvement and cost-effectiveness, digital predistortion (DPD) stands out as one of the most popular methods to linearize the PA [3]. Basically in DPD, a digital predistorter (PD) which has the inverse nonlinear characteristics of that of the PA, is added in baseband. This way the cascaded digital predistorter-PA system becomes linear and the PA can be driven more towards the high efficiency saturation region without compromising much on linearity.

One of the most studied approaches to realize a DPD is the indirect learning architecture (ILA) [4]–[6]. In the conventional ILA, a post-inverse of the PA is first identified and then used as a PD. The post-inverse/PD is usually modeled as a memory polynomial (MP) [7] model and identified using the input and output of PA by applying simple least squares (LS) method [5]. The nonlinearity order and memory depth of the memory polynomial PD depends upon the PA nonlinearity to be compensated and varies with the PA. Usually there is a trade-off between the complexity of the MP and the nonlinear distortions that can be tolerated. Using a higher order and large memory depth MP model might improve the performance of PD and reduce nonlinear distortions at the output of PA however it will increase the computational complexity substantially [8], [9]. Moreover, there is always a limitation on the maximum performance that can be achieved by this conventional ILA-DPD approach. Once the ILA-DPD system converges to the best possible solution[1], there is no substantial improvement in the performance with any subsequent increase in nonlinearity order or memory depth [10].

In this paper, we propose a multi-stage ILA for digital predistortion that has lower computational complexity (CC) for the identification of PD parameters[2]. In this multi-stage architecture, the nonlinearity of PA is compensated by implementing PD in two or more stages. Each stage is modeled by a lower order memory polynomial thus having a lower identification complexity. Simulation results show that this architecture can achieve same or even better performance then the conventional single stage ILA. Moreover the identification complexity of this multi-stage architecture is much lower than the conventional architecture.

Apart from above, we also propose two different algorithms

*The research leading to these results has received funding from the European Seventh Framework Programme under grant agreement n° 230688.

[1]The best possible solution here refers to the identified MP model which achieves the best improvement in the performance of ACPR and EVM at the output of PA.

[2]In the following, the computational complexity always refer to the complexity involved in parameters identification.

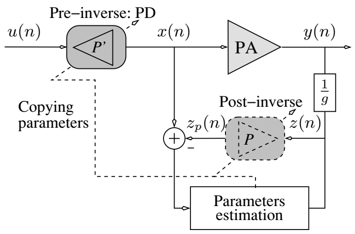

Fig. 1. Indirect Learning Architecture - ILA.

to identify the PD of this multi-stage ILA. In the first algorithm, the system level identification is done for each stage and once this stage converges to the best possible solution new stage is added to further improve the performance. For the second algorithm, the system level identification of each stage is done by taking all the stages into consideration. The system is optimized to achieve the best possible solution by simultaneously considering all the stages. We later show that for equivalent number of stages both the algorithms obtain approximately the same performance.

The performance of the proposed multi-stage ILA is evaluated in terms of adjacent channel power ratio (ACPR) and error vector magnitude (EVM) improvements using an LTE-Advanced signal. Two different PA models have been used: Wiener model given in [11] and Wiener-Hammerstein model given in [5]. The identification complexity for conventional as well as multi-stage ILA is measured by computing the number of multipliers needed.

The remainder of this paper is organized as follows. Section II gives the theoretical background related to the conventional ILA. Section III presents the proposed multi-stage ILA and discusses the identification algorithms for multi-stage ILA. In Section IV simulation results for conventional ILA and proposed multi-stage ILA are presented and discussed. Finally Section V concludes the paper. In the following, the vectors and matrices are denoted by bold lowercase letters (eg. $\boldsymbol{a}$) and bold uppercase letters (eg. $\boldsymbol{A}$) respectively. The superscripts $(.)^*$, $(.)^T$ and $(.)^H$ denote the conjugate, the transpose and the conjugate transpose, respectively.

II. Conventional ILA

The PD identification in ILA is done in a single step as shown in Fig. 1. A post-inverse of the PA is identified and used as a PD. If the post-inverse is modeled as a MP, then its output can be written as [6]

$$z_p(n) = \sum_{k\in K}\sum_{l\in L} c_{kl}\Phi_{kl}[z(n)] \tag{1}$$

where $z(n) = \frac{y(n)}{g}$ is the input to the post-inverse block as shown in Fig. 1, K is the index array for nonlinearity and L is the index array for memory. c_{kl}, $k \in K$ and $l \in L$ are the complex coefficients and $\Phi_{kl}[z(n)] = z(n-l)|z(n-l)|^k$. The total number of coefficients is $J = \bar{\bar{K}}\bar{\bar{L}}$ with $\bar{\bar{X}}$ denoting the cardinality (number of elements) of X.

After convergence, we should have $z_p(n) = x(n)$ and hence $z(n) = u(n)$. As (1) is linear in the parameters c_{kl}, we can rewrite it using matrix notation [6]

$$\boldsymbol{z}_p = \boldsymbol{Z}\boldsymbol{c} \tag{2}$$

where $\boldsymbol{z}_p = [z_p(1), \ldots, z_p(N)]^T$, $\boldsymbol{c}$ is $J \times 1$ vector containing the set of coefficients c_{kl}, $\mathbf{Z}$ is $N \times J$ matrix containing $\Phi_{kl}[\boldsymbol{z}]$ where $\boldsymbol{z} = [z(1), \ldots, z(N)]^T$ and N is the total number of samples. The least square solution of (2) will the solution of the normal equation

$$\mathbf{Z}^H\boldsymbol{z}_p = \mathbf{Z}^H\mathbf{Z}\hat{\boldsymbol{c}}. \tag{3}$$

The following briefly illustrates the steps in computation [3],of $\hat{\boldsymbol{c}}$.

- Step 1: Define a new compound matrix
$$[\mathbf{Z}^H\mathbf{Z}|\mathbf{Z}^H\boldsymbol{z}_p] = \mathbf{QR} = \mathbf{Q}[\mathbf{U}|w]. \tag{4}$$
- Step 2: Compute the QR decomposition of the compound matrix by Gram Schmidt process [12].
- Step 3: Substitute the result of Step 2 into Step 1 to obtain
$$w = \mathbf{U}\hat{\boldsymbol{c}} \tag{5}$$
which could be solved using back substitution [13].

TABLE I
Computational Complexity

Step Number	CC
Step 1	$N(J + J^2)$
Step 2	J^3
Step 3	J^2

Table I summarizes roughly the CC needed in the computation[4] of $\hat{\boldsymbol{c}}$. The CC is measured by computing the number of multiplication needed for each step. Hence, total CC needed for computation of $\hat{\boldsymbol{c}}$ would be $N(J + J^2) + J^3 + J^2$. However, since $\hat{\boldsymbol{c}}$ and $\mathbf{Z}$ are complex, the total number of real multiplication operations needed will be [12].

$$\text{CC} = 4(N(J + J^2) + J^3 + J^2). \tag{6}$$

In general, it is perceived that, as the number of terms of an MP model increases, its correction capability increases, although the amount of contribution might vary across the terms [6] [9], i.e an MP model with larger J might perform better than an MP model with smaller J. However, addition of each new term increases the computational complexity of both online as well as identification processing, and consequently the implementation cost [14]. As a result, there is a significant trade-off in the desired performance of the PD and the number of terms that can be included in the PD. To achieve a better performance we would like to have a larger J, however as the number of multiplication operations required is directly proportional to the square of J (6) it will substantially increase the computational complexity needed for identification of PD.

[3]Note that in this paper we are not taking into consideration the complexity involved in computing matrix $\mathbf{Z}$.

[4]Note that here we have used Gram-Schmidt method to compute the complexity by QR decomposition; other methods like Householder transformation or Givens rotation can also be used.

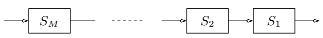

Fig. 2. Block diagram of a multi-stage (multi-box) nonlinear model.

III. Multi-stage ILA

Before proceeding to give a comprehensive overview of the proposed multi-stage ILA and the algorithms to identify the PD for this multi-stage ILA, it is important to discuss a multi-stage nonlinear model. A block diagram of a multi-stage nonlinear model is shown in Fig. 2. This multi-stage structure has been widely used for PA and PD modeling. In this multi-stage structure one or more linear time invariant (LTI) systems are cascaded with a static memoryless nonlinearity [15]. The most used model belonging to this category is the Wiener-Hammerstein (W-H) model, also called three-box model, with cascaded linear(L)-nonlinear(N)-linear(L) (LNL) operators [16], [17]. The Wiener (LN) and Hammerstein (NL) models which are two particular cases of the W-H model, has been also used [18], [19]. Different parameter estimation algorithms have been proposed in literature to extract the parameters of these model structures [16]. However, their implementation complexity and their low correction capability when compared to an MP model (1) limit their use as a PD.

We have tried to overcome the above drawback by generalizing the above multi-stage structure. We use an MP to model each stage of the above multi-stage structure. Other configurations like generalized MP (GMP) [6] or 2D-MSP proposed in [9] can also be used to model each stage. The proposed multi-stage model have the desirable property that the output at each stage is linearly dependent on its coefficients and hence LS algorithm can be used for parameter extraction as shown in Section II.

A. Identification algorithms for the proposed multi-stage ILA-DPD

One of the bottleneck in identifying a multi-stage system comes from the fact that the internal signals interconnecting the stages are inaccessible to measurements. In this paper, the proposed identification algorithms rely on the idea of gradually linearizing the PA, using an ILA. We start by identifying a first stage which partially compensates the nonlinearity and/or memory of the PA. Then we add a second stage to improve the linearity or to compensate for the residual distortion of the new system constituted by the cascade of the first stage and the PA, then a third one to linearize the cascade of the second and first stages with the PA, etc.. Thus, the n^{th} stage is implemented to improve the linearity of the cascade of the $n-1$ first stages with the PA. The complexity of the identification algorithm, which depends on the number of parameters, will substantially decrease since only one stage is processed by iteration. In the following, we present two identification algorithms for multi-stage DPD implementation based on an indirect learning architecture.

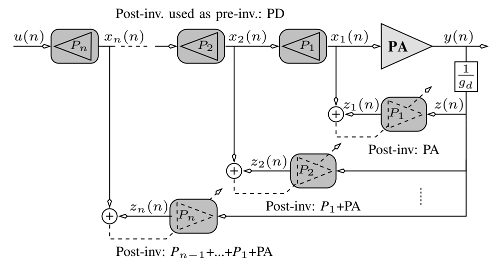

Fig. 3. Indirect learning architecture of a multi-stage DPD based on the first algorithm.

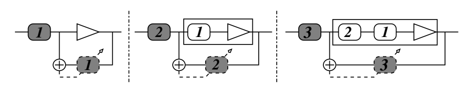

Fig. 4. Graphical illustration of algorithm 1 with three-stage DPD.

1) Algorithm 1: The most simple and straightforward identification approach is to identify stages from the closest to the farthest to the PA. Once a stage is identified, it will be considered as being part of a new system, the cascade with the already identified stages and the PA. So, in the identification of the current stage, P_i, only stages, $P_{i-1}, ..., P_1$, are used and their coefficients are kept constant until the system-level convergence is reached [10]. The block diagram of the ILA of the cascaded multi-stage DPD based on this first algorithm is shown in Fig. 3. It is important to note that for the identification of each stage more than one system level iteration is needed [10]. So after system convergence of the current stage, we can proceed with the identification of the next one. Fig. 4 illustrates graphically the steps of this algorithm for a three-stage DPD, where the dark boxes represent the stages that are currently being identified. If we denote N_i the number of iterations needed for stage P_i, the total number of iterations for the whole system to converge is equal to $\sum_{i=1}^{M} N_i$, where M is the total number of stages.

2) Algorithm 2: In the first algorithm, once P_i stage is identified, for subsequent identification of new stage P_{i+1}, it will be considered as being part of a newly constituted system (stage $P_i, P_{i-1}, ..., P_1$ PA) as shown in Fig. 4 and its parameters wont be affected by the identification of stage P_{i+1}. In order to have more interaction between stages, after the insertion of a new stage we can re-identify the other stages. This process may optimize the identification of all stages and lead to more accurate results. The block diagram of the ILA based on the second algorithm is shown in Fig. 5. In this case, stage P_i is processed at the $(Mk+i)^{th}$ iteration, with $k = 0, 1, 2, 3, ...$ using stages $P_M, ..., P_{i-1}$. The graphical description in Fig. 6 aids in explaining the steps for the identification of a three-stage DPD model using this algorithm.

IV. Simulation results

In this section, we present and discuss the simulation results for the multi-stage ILA. For this purpose, we use two different reference PA models, Wiener model as given in [11] and

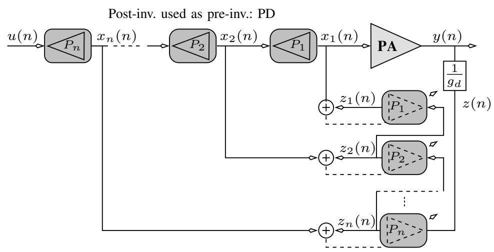

Fig. 5. Indirect learning architecture of a multi-stage DPD based on the second algorithm.

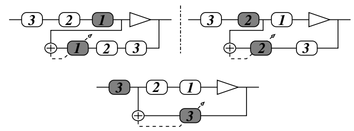

Fig. 6. Graphical illustration of algorithm 2 with three-stage DPD.

TABLE II
WIENER MODEL

Parameters	Without DPD	Two-stage ILA (Algorithm 1)	Two-stage ILA (Algorithm 2)	Conventional ILA
ACPRU(dBc)	-50.24	-111.2	-110.2	-109.9
ACPRL(dBc)	-48.26	-110.5	-110.6	-109.1
EVM(%)	17.95	0.0292	0.0291	0.0285
Index array for Nonlinearity and Memory	NA	Stage 1: K=[0 2 4 6], L=[0 2 4] Stage 2: K=[0 2 4 6 8], L= [0 2 4 6 8]	Stage 1: K=[0 2 4 6], L=[0 2 4] Stage 2: K=[0 2 4 6 8], L= [0 2 4 6 8]	K=[0 2 4 6 8 10] L=[0 2 4 6 8 10]
Number of Coefficients	NA	Stage 1: 12 Stage 2: 25	Stage 1: 12 Stage 2: 25	36
Normalized Complexity	NA	Stage 1: 0.1170 Stage 2: 0.4878	Stage 1: 0.1170 Stage 2: 0.4878	1

TABLE III
WIENER-HAMMERSTEIN MODEL

Parameters	Without DPD	Two-stage ILA (Algorithm 1)	Two-stage ILA (Algorithm 2)	Conventional ILA
ACPRU(dBc)	-44.79	-90.4	-89.8	-89.4
ACPRL(dBc)	-45.41	-90	-89.7	-89.3
EVM(%)	20.05	0.0161	0.016	0.0316
Index array for Nonlinearity and Memory	NA	Stage 1: K=[0 2 4], L=[0 1 2 3] Stage 2: K=[0 2 4 5 6 7], L= [0 1 2 4 6]	Stage 1: K=[0 2 4], L=[0 1 2 3] Stage 2: K=[0 2 4 5 6 7], L= [0 1 2 4 6]	K=[0 1 2 3 4 5 6 7] L=[0 1 2 4 6 7]
Number of Coefficients	NA	Stage 1: 12 Stage 2: 30	Stage 1: 12 Stage 2: 30	48
Normalized Complexity	NA	Stage 1: 0.0662 Stage 2: 0.3951	Stage 1: 0.0662 Stage 2: 0.3951	1

Wiener-Hammerstein model as given in [5].

The PA is driven by an LTE-Advanced signal with bandwidth 10 MHz, sampling frequency 122.88 MHz and peak-to-average power ration (PAPR) of approximately 11dB. The number of input samples for each system level iteration is 25000. As stated before, an MP model is used for the extraction of parameters of PD for both conventional as well as multi-stage ILA.

The simulation is done in two steps. In the 1st step, conventional ILA is used as a DPD for a particular reference PA. The nonlinearity order and memory depth of the MP are varied till the conventional ILA DPD converges to the best possible solution [10]. The computational complexity in terms of number of multipliers is computed for the identified MP model.

In the 2nd step, multi-stage ILA is used as DPD for the same reference PA. The nonlinearity order and memory depth of the MP for each stage are varied till the multi-stage ILA DPD achieves the same performance (in terms of ACPR and EVM at the output of the PA) as that with the best performing conventional ILA[5]. The CC in terms of number of multipliers is computed for the identified MP model of each stage. In the following, the performance of multi-stage ILA has been demonstrated by considering two-stage ILA, however it can be easily extended to more number of stages.

Table II shows the results for a PA modeled as Wiener Model. K and L are vectors containing k and l values as given in (1). ACPRU and ACPRL denotes the power in adjacent upper channel and lower channel with respect to main channel respectively. The CC of two-stage ILA has been normalized with respect to conventional ILA. As seen from Table II, both two-stage and conventional ILA DPD are able to achieve sufficient improvement in ACPR and EVM at the output of the PA. However, the two-stage algorithms 1 and 2 are able to achieve the same performance as best performing conventional ILA with significantly lower CC. It can also be seen from Table II that two-stage PD requires lower non linearity order or memory depth and as a consequence the CC for two-stage algorithms is approximately half of that needed with conventional ILA.

Table III shows the results for a PA modeled as Wiener-Hammerstein Model. The number of multipliers needed by two-stage algorithms is less than half of that needed by conventional ILA. Moreover with the two-stage algorithms, EVM is reduced to half of that obtained by the best performing conventional ILA. Memory depth has also been reduced by 1 for the case of two-stage ILA.

It is also worth noting from Table II and Table III that as each stage have fewer coefficients in two-stage PD than in the conventional PD, as a consequence the condition number of the matrix $\mathbf{Z}^H\mathbf{Z}$ in (3) is reduced and thus the sensitivity to noisy measurement is also reduced. Another point to note is that both the two-stage algorithms, 1 and 2 achieve approximately same ACPR and EVM performance for a given number of stages and same number of coefficients for each stage. This behavior might be due to the fact that both the algorithms although inherently different are trying to reach the same solution. Further investigation might be needed to validate this hypothesis and will be the subject of future work.

V. CONCLUSION

A multi-stage indirect learning architecture (ILA) with low identification complexity is proposed. In this multi-stage ILA, PD is implemented in two or more stages. It was shown that this multi-stage ILA was able to achieve same or even better performance than the conventional ILA but with significantly lower identification complexity. The performance of the multi-stage ILA was evaluated by measuring the ACPR and EVM at the output of the PA for an LTE-Advanced input signal. The reference PA models used for simulation were the Wiener model and Wiener-Hammerstein model.

[5]Note that the motivation here is to obtain same performance as that of conventional ILA with less CC. The optimal performance that can be achieved with multi-stage ILA can be higher than shown here.

REFERENCES

[1] S. C. Cripps, *RF Power Amplifiers for Wireless Communications, Second Edition (Artech House Microwave Library (Hardcover))*. Norwood, MA, USA: Artech House, Inc., 2006.

[2] Y. Ye, T. Liu, and F. M. Ghannouchi, "Linear and nonlinear memory effects of RF power amplifiers," in *Microwave Conference, 2008. APMC 2008. Asia-Pacific*. IEEE, Dec. 2008, pp. 1–4.

[3] P. Kenington, "Linearized transmitters: an enabling technology for software defined radio," *Communications Magazine, IEEE*, vol. 40, no. 2, pp. 156–162, Feb 2002.

[4] C. Eun and E. Powers, "A new volterra predistorter based on the indirect learning architecture," *Signal Processing, IEEE Transactions on*, vol. 45, no. 1, pp. 223–227, Jan 1997.

[5] L. Ding, G. Zhou, D. Morgan, Z. Ma, J. Kenney, J. Kim, and C. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," *Communications, IEEE Transactions on*, vol. 52, no. 1, pp. 159–165, Jan. 2004.

[6] D. Morgan, Z. Ma, J. Kim, M. Zierdt, and J. Pastalan, "A generalized memory polynomial model for digital predistortion of rf power amplifiers," *Signal Processing, IEEE Transactions on*, vol. 54, no. 10, pp. 3852–3860, Oct. 2006.

[7] 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.

[8] L. Guan and A. Zhu, "Low-cost fpga implementation of volterra series-based digital predistorter for rf power amplifiers," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 58, no. 4, pp. 866 –872, april 2010.

[9] M. Abi Hussein, V. A. Bohara, and O. Venard, "Two-dimensional memory selective polynomial model for digital predistortion," in *New Circuits and Systems Conference (NEWCAS), IEEE 10th International*, June 2012, pp. 469 –472.

[10] ——, "On the system level convergence of ila and dla for digital predistortion," in *Wireless Communication Systems (ISWCS), 2012 International Symposium on*, Aug. 2012, pp. 870 –874.

[11] E. Aschbacher and M. Rupp, "Modelling and identification of a nonlinear power-amplifier with memory for nonlinear digital adaptive predistortion," in *Signal Processing Advances in Wireless Communications, 2003. SPAWC 2003. 4th IEEE Workshop on*, june 2003, pp. 658 – 662.

[12] L. Trefethen and D. Bau III, *Numerical linear algebra*. Society for Industrial Mathematics, 1997, no. 50.

[13] M. Moonen and I. Proudler, "Introduction to adaptive signal processing," *Department of Electrical Engineering ESAT/SISTA KU Leuven, Leuven, Belgium*, pp. 105–107, 1998.

[14] A. Tehrani, H. Cao, S. Afsardoost, T. Eriksson, M. Isaksson, and C. Fager, "A comparative analysis of the complexity/accuracy tradeoff in power amplifier behavioral models," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 58, no. 6, pp. 1510 –1520, june 2010.

[15] J. Pedro and S. Maas, "A comparative overview of microwave and wireless power-amplifier behavioral modeling approaches," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 53, no. 4, pp. 1150–1163, April 2005.

[16] H.-W. Chen, "Modeling and identification of parallel nonlinear systems: structural classification and parameter estimation methods," *Proceedings of the IEEE*, vol. 83, no. 1, pp. 39 –66, jan 1995.

[17] A. Bolstad, B. Miller, J. Goodman, J. Vian, and J. Kalyanam, "Identification and compensation of wiener-hammerstein systems with feedback," in *Acoustics, Speech and Signal Processing (ICASSP),2011 IEEE International Conference on*, may 2011, pp. 4056 –4059.

[18] S. Chen, "An efficient predistorter design for compensating nonlinear memory high power amplifiers," *Broadcasting, IEEE Transactions on*, vol. 57, no. 4, pp. 856 –865, dec. 2011.

[19] T. Liu, S. Boumaiza, and F. Ghannouchi, "Augmented hammerstein predistorter for linearization of broad-band wireless transmitters," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 54, no. 4, pp. 1340–1349, June 2006.