

A low complexity pre-distortion scheme for power amplifier linearization in wideband applications

Pooria Varahram^{1,*†}, John Dooley², Keith Finnerty¹ and Ronan Farrell¹

¹*Callan Institute, Department of Electrical Engineering, National University of Ireland Maynooth, Maynooth, Ireland*

²*Electrical Engineering, National University of Ireland, Maynooth, Ireland*

SUMMARY

In this paper, a nonlinear autoregressive with exogenous inputs (NARX) digital pre-distortion scheme to linearize power amplifiers is proposed. The proposed NARX digital pre-distortion gives better accuracy and spectral leakage suppression compared with other commonly used Volterra-based techniques. The stability criterion of the NARX digital pre-distortion is derived from the frequency domain analysis. Simulation is carried out with a 20 MHz single carrier long-term evolution signal and two-carrier long-term evolution signal and instantaneous to average ratio of 6.2 dB at 0.01% complementary cumulative distribution function (CCDF). The results of simulation analysis show a slight improvement in adjacent channel leakage ratio performance with 30% reduction in the number of floating point operations compared with conventional pre-distortion techniques. Copyright © 2016 John Wiley & Sons, Ltd.

Received 3 March 2016; Revised 7 June 2016; Accepted 18 July 2016

KEY WORDS: POWER AMPLIFIER; NONLINEAR AUTOREGRESSIVE; PRE-DISTORTION; MEMORY POLYNOMIAL

1. INTRODUCTION

Power amplifiers (PAs) are inherently nonlinear devices in which by increasing the input power the output does not linearly vary with input. This causes limitation in the operating range of a PA. Some impacts of PA nonlinearity are spectral leakage and in-band distortion. Spectral leakage leads to superfluous emissions and adjacent channel interference which can disrupt signals in adjacent channels. One simple way to avoid generating adjacent channel leakage is to back-off the PA. In doing this, a linear operation can be achieved; however, this causes power efficiency degradation. Hence, a solution to efficiently linearize a PA is imperative.

Many techniques proposed in the literature to linearize power amplifiers by employing digital pre-distortion techniques. This is due to its several advantages over feedforward linearization techniques [1] such as high flexibility, low cost, and high performance. Popular techniques which have been presented in the past include Volterra series techniques [2], Hammerstein [3], memory polynomial (MP) pre-distortion [4], and generalized memory polynomial (GMP) pre-distortion [5]. A nonlinear autoregressive exogenous (NARX) model behaves like an infinite impulse response (IIR) filter [6]. This means that the output depends on both past inputs and outputs. This is in contrast to Volterra series techniques with a finite impulse response (FIR) architecture. The importance of the NARX model is the inclusion of the output signal in the model which gives additional information that leads to lower number of coefficients.

*Correspondence to: Pooria Varahram, Callan, Institute, Electrical Engineering, National University of Ireland Maynooth 8798, Maynooth, Ireland.

†E-mail pooria.varahram@nuim.ie

In [7], authors proposed an adaptive approach to model PA using a recursive filter based on IIR filter. The advantage of their recursive method was its greater accuracy with less number of coefficients compared with the conventional methods. Some other publications studied the stability and convergence of the digital pre-distortion (DPD) technique based on IIR structure. In [8], a nonlinear autoregressive moving average (NARMA) DPD based on multi-lookup table functions was proposed. They used a small-gain theorem based on bounded norms for stability test and derived a criterion to achieve the stability of the NARMA DPD. However, a comprehensive analysis was not presented for the approach and a comparison with other methods was not given. In [9], a neural network algorithm to implement digital pre-distortion was used. However, no discussion on stability was provided. In [10], a new recursive prediction technique is proposed. The proposed technique reduced the mean square error (MSE) and achieved good adjacent channel leakage ratio (ACLR) performance. However, the main drawback of the proposed approach is in its high complexity and low convergence speed which causes performance degradation in high data rate applications. In [11], a new adjacent channel leakage ratio DPD is proposed. Moreover, the NARX DPD model requires more coefficients compared with the one proposed in this paper, which gives lower complexity by using only linear terms of the output.

In this paper, a low complexity DPD technique based on the NARX model is proposed. The identification of the NARX DPD is based on least square (LS) algorithm. To ensure the stability of the proposed NARX DPD, a stability criterion is derived. The computational complexity is measured by evaluating the number of floating point operations (FLOPs).

The rest of the paper is organized as follows. The proposed NARX based DPD technique is presented in Section II. The LS solution to identify the pre-distortion coefficients is explained. The stability criterion of the NARX DPD is derived, and a peak cancelation technique to reduce high peaks is introduced. The computational complexity analysis is provided, and a discussion on the simulation results for the proposed approach compared with present solutions is compiled in Section III. Finally, the conclusions from this work are offered.

2. DIGITAL PRE-DISTORTION STRUCTURE

2.1. The nonlinear autoregressive exogenous digital pre-distortion model

The NARX DPD model in general [11] can be expressed by

$$u[n] = \sum_{k=1}^{K'} \sum_{i=1}^I a_{ki} u[n-i] |u[n-i]|^{k-i} + \sum_{k=1}^K \sum_{m=0}^M b_{km} y[n-m] |y[n-m]|^{k-1} + e[n], \quad (1)$$

where $u[n]$ and $y[n]$ denote the captured input samples and the captured output samples, respectively, $e[n]$ is the error of the modeling, a and b are the coefficients extracted from the inverse of the power amplifier, K' and K are the nonlinearity orders of autoregressive and exogenous inputs, and I and M are the input and output memory length, respectively.

Despite the fact that NARX provides better accuracy in modeling, it can suffer from stability compared with MP. This is addressed by providing a stability criterion in this section. It can be said that Volterra series and sub-block structures are a subset of NARX. The other consideration while modeling a nonlinear system is to ensure that all the validation procedures have been performed and not only by measuring the MSE, which is often uninformative and can lead to fitting to the noise [6].

2.2. Nonlinear autoregressive exogenous identification algorithm

In order to identify the pre-distortion coefficients, the indirect learning architecture based on post-distortion as illustrated in Figure 1 is applied. The coefficients are extracted first from the post-distortion structure and applied to the forward path [12]. It can be shown that the extracted coefficients from the reverse path are the same as pre-distortion coefficients in the forward path [4].

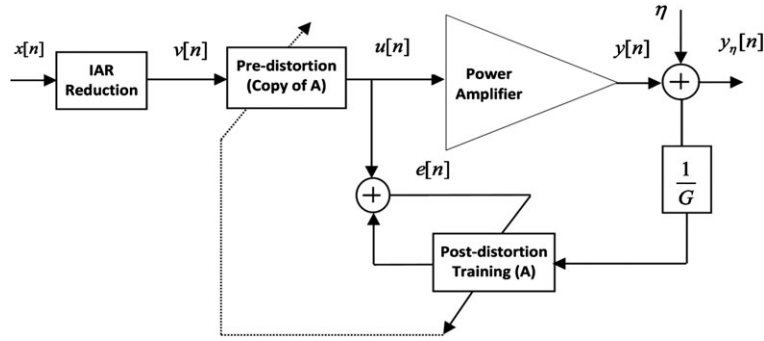


Figure 1. Indirect learning architecture with instantaneous to average ratio (IAR) reduction.

In contrast to nonlinear autoregressive moving average exogenous in which the iterative identification has to be applied to obtain the coefficients [5], the NARX DPD uses a one-off estimation by LS algorithm. The LS algorithm to extract these coefficients by assuming $K'=0$ and writing (1) in matrix form is as follows:

$$u[n] = \sum_{i=1}^I a_i u[n-i] + \sum_{k=1}^K \sum_{m=0}^M b_{km} y[n-d-m] |y[n-d-m]|^{k-1} + e[n]. \quad (2)$$

Let $s = 1 + \max(I, M)$, the input of the power amplifier or output of the pre-distortion vector $u = [u[s], u[s+1], \dots, u[N]]^T$, $e = [e[s], e[s+1], \dots, e[N]]^T$, and $q = [a_1, a_2, \dots, a_I, b_1, b_2, \dots, b_{KM}]^T$. The main feature of (2) is that the nonlinearity order of autoregressive as indicated in (1) by K' is zero. This results in lower number of coefficients in identification algorithm and forward path, which reduces the number of FLOPs.

Both (1) and (2) can be written using delay operator notation, where q^{-1} represents a unit delay as follows:

$$u[n] = A(q)u[n]|u[n]|^{k-1} + B(q)y[n]|y[n]|^{k-1}, \quad (3)$$

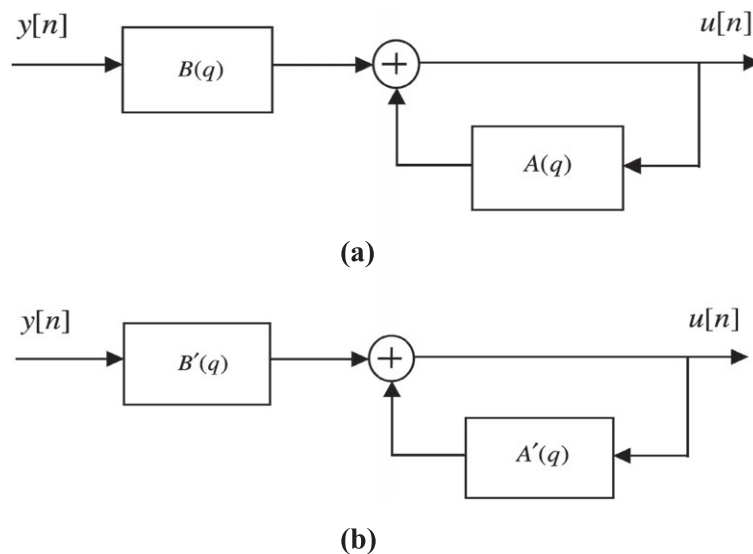


Figure 2. Nonlinear autoregressive exogenous digital pre-distortion (NARX DPD) block diagram a) Original NARX DPD b) Low complexity NARX DPD.

where $A(q) = \sum_{k=1}^K \sum_{i=1}^I a_{ki} q^{-i}$ and $B(q) = \sum_{k=1}^K \sum_{m=0}^M b_{km} q^{-m}$.

$$u[n] = A'(q)u[n] + B'(q)y[n]|y[n]|^{k-1}, \quad (4)$$

where $A'(q) = \sum_{i=1}^I a_i q^{-i}$ and $B'(q) = \sum_{k=1}^K \sum_{m=0}^M b_{km} q^{-m}$.

Figure 2 shows the block diagram of NARX DPD. Figure 2(a) is the block diagram of NARX DPD as expressed in (3) and Figure 2(a) is the block diagram of the low complexity NARX DPD in (4). The NARX DPD in (2) can be written in matrix form by collecting all the coefficients of the nonlinear auto regressive and exogenous input into one $L \times 1$ vector, where L , the total number of coefficients equals $[K(M+1)+I] \times 1$. Each coefficient associated with a sample of signal denotes n .

The post-distortion function for the N data sequence can be written by

$$u = Pq + e. \quad (5)$$

In (5), P is a matrix of $N \times L$, which can be expressed by

$$P = \begin{bmatrix} p_u(s-1) & \cdots & p_u(s-I) & p_y(s) & \cdots & p_y(s-M) \\ p_u(s) & \cdots & p_u(s+1-I) & p_y(s+1) & \cdots & p_y(s+1-M) \\ p_u(s+1) & \cdots & p_u(s+2-I) & p_y(s+2) & \cdots & p_y(s+2-M) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ p_u(N-2) & \cdots & p_u(N-1-I) & p_y(N-1) & \cdots & p_y(N-1-M) \\ p_u(N-1) & \cdots & p_u(N-I) & p_y(N) & \cdots & p_y(N-M) \end{bmatrix}, \quad (6)$$

where $p_u[n] = u[n]$, $p_y[n] = \frac{y[n]}{G} \left| \frac{y[n]}{G} \right|^{k-1}$, and G is the PA gain.

The LS solution to identify the coefficients q is given as

$$\hat{q} = (P^T P)^{-1} P^T u, \quad (7)$$

where $(\cdot)^T$ denotes complex conjugate transpose, and u is a vector of input samples. The coefficients of pre-distortion function (\hat{q}) can be calculated offline from the measured input and output data stream of the power amplifier and then applied to the running path. In actual systems, the algorithm used to obtain the coefficients is performed in the processor, and the running path is implemented on the field programmable gate arrays.

2.3. Stability analysis

To derive the stability criterion of the proposed technique, the frequency domain technique should be used. This is due to the fact that by obtaining the frequency behavior of the pre-distortion function, the stability of the function can be determined. The frequency domain behavior of the pre-distortion in this paper is evaluated by the ACLR parameter. In [13], the bounded input bounded output of the NARX model to ensure the stability is obtained from its frequency response. The MSE can be expressed as

$$MSE\{\hat{u}_n\} = E\left\{|\hat{u}_n - u_n|^2\right\}, \quad (8)$$

where E is the mean value.

The inequality condition can be satisfied by obtaining a series of scalar positive real numbers $c_{k0}, c_{k1}, \dots, c_{kk}$ as follows:

$$|H_k(j\omega_1, \dots, j\omega_k)| \leq c_{k0} + c_{k1}\overline{H}_1 + c_{k2}\overline{H}_1^2 + \dots + c_{kk}\overline{H}_1^k, \quad (9)$$

where the coefficients $c_{k0}, c_{k1}, \dots, c_{kk}$ can be recursively determined from the coefficients of NARX DPD in (2). With the input of the NARX DPD as a generalized input with spectrum $V(j\omega)$, then the output spectrum of the NARX DPD, $U(j\omega)$ is bounded by

$$|U(j\omega)| \leq \bigoplus_{k=1}^K \frac{1}{(2\pi)^{k-1}} \cdot c_k \cdot h_k^T \cdot |V| * \dots * |V(j\omega)| = \left(\bigoplus_{k=1}^K \alpha_k c_k \right) \cdot h_k^T, \quad (10)$$

where $h_K = [1 \ \overline{H}_1 \dots \overline{H}_1^K]$ and the condition that the right side of (10) is convergent is given by the inequality condition

$$\lim_{\substack{K \rightarrow \infty \\ g \rightarrow \infty}} \sqrt[g]{\left(\bigoplus_{k=1}^K \alpha_k c_k \right)}(g) < \frac{1}{H_1}, \quad (11)$$

where the operator \oplus denotes series addition, $c_k = [c_{k0} \ c_{k1} \ \dots \ c_{kk}]$, and $\alpha_k = (2\pi)^{1-k} k \underbrace{|V| * \dots * |V(j\omega)|}_k$.

The proof of the stability criterion in (11) can be found in Appendix. In Section III, the stability analysis is performed by evaluating the normalized mean square error (NMSE) parameter with different coefficients and ACLR performance.

2.4. Instantaneous to average ratio reduction

Instantaneous to average ratio (IAR) is a better comparison metric compared with crest factor reduction. This is because of some nonlinearity that can be tolerated below spectral emission mask. IAR can be defined as

$$IAR[n] = \frac{[|x[n]|^2]}{E[|x[n]|^2]}, \quad (12)$$

where E is the mean value and $x(n)$ is the input signal. Here, a peak cancellation technique [14] is applied to reduce the IAR. The benefit of this approach compared with other techniques such as coding [15] and scrambling techniques [16, 17] is in lower complexity. In this approach, first the magnitude and phase of a signal is determined and then the peak scaling is created as follows:

$$c(n) = \begin{cases} 0 & \text{if } |x(n)| < A \\ (|x(n)| - A) \angle x(n) & \text{if } |x(n)| \geq A \end{cases}, \quad (13)$$

where $x(n)$ is the input signal, and A is the predefined threshold value. The peak scaled signal of (13) passes through a filter to smooth the clipped signal and will be subtracted from the delayed input signal.

$$v(n) = x(n - d) - \sum_{l_p=1}^L [(|x_{l_p}(n)| - A_{th}) \cdot \delta(n - l_p) e^{j\theta_{l_p}}] * w(n), \quad (14)$$

where the sign $*$ is the convolution, $w(n)$ is the FIR filter function, d is the delay, which theoretically equals the group delay of the FIR filter, that is. length of the filter over 2, and L is the number of peaks above the threshold.

The complementary cumulative distribution function (CCDF) of a single carrier long-term evolution (LTE) signal is shown in Figure 3. From this figure, the IAR reduces 3.1 dB at 0.01% CCDF. This value for two-carrier LTE-A signal is almost 4.2 dB.

The output signal is then as follows:

3. COMPARISON OF COMPUTATIONAL EFFICIENCY AND NUMERICAL ACCURACY

3.1. Computational overhead

The MP, which is a subset of NARX DPD, can be expressed by

$$u[n] = \sum_{k=1}^K \sum_{m=0}^M b_{km} v[n-m] |v[n-m]|^{k-1}, \quad (15)$$

where $v[n]$ denotes the pre-distortion input signal as shown in Figure 1.

The computational complexity is assessed based on the number of FLOPs [18]. FLOPs indicate the complexity comprising of the number of additions or subtractions, and multiplications. As explained in [18], the number of FLOPs of a model comprises of basis functions and filtering. Hence, the total number of FLOPs per sample for MP can be derived as follows:

$$C_{total}(MP) = C_{basis} + C_{filter} = [10 + 2(K-1)] + [8K(M+1) - 2], \quad (16)$$

where C_{total} represents the total complexity, C_{basis} represents the complexity of basis function $x(n) \propto (n)^{k-1}$ and C_{filter} represents the complexity of the number of coefficients. It should be noted that K includes both even and odd numbers in order to achieve a more accurate performance.

The generalized memory polynomial, which is also a subset of NARX DPD, can be expressed by

$$u[n] = \sum_{k=1}^K \sum_{m=0}^M b_{km} v[n-m] |v[n-m]|^{k-1} + \sum_{k=1}^{K_a} \sum_{m=0}^{M_a} \sum_{l=1}^{L_a} c_{kML} v[n-m] |v[n-m-l]|^{k-1} + \sum_{k=1}^{K_b} \sum_{m=0}^{M_b} \sum_{l=1}^{L_b} d_{kML} v[n-m] |v[n-m+l]|^{k-1}, \quad (17)$$

where K , K_a , K_b denote the nonlinearity order, M , M_a , M_b denote the memory length, and L , L_a , L_b are the delay taps of each branch, respectively.

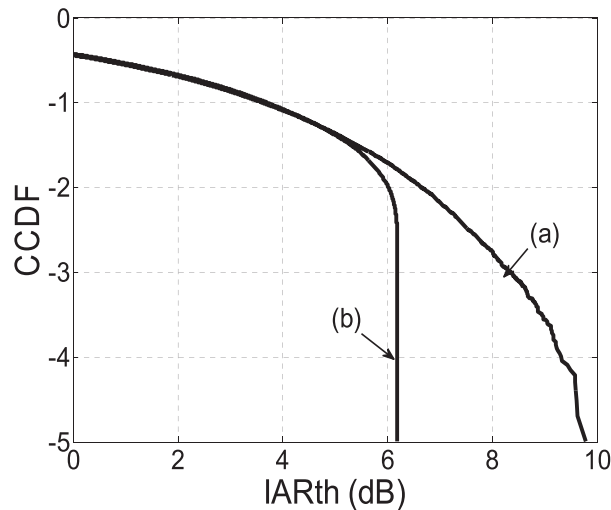


Figure 3. Complementary cumulative distribution function (CCDF) of single carrier long-term evolution signal (a) before instantaneous to average ratio (IAR) reduction (b) after IAR reduction.

In a similar way, the total number of FLOPs per sample for GMP can be derived as follows:

$$C_{total}(GMP) = C_{basis} + C_{filter} = [10 + 2(K - 1) + 2(K_a - 1)L_a + 2(K_b - 1)L_b] + [8K(M + 1) + 8(K_a - 1)(M_a + 1)L_a + 8(K_b - 1)(M_b + 1)L_b - 2], \quad (18)$$

Finally, the total number of FLOPs per sample for the NARX model is given by

$$C_{total}(NARX) = C_{basis} + C_{filter} = [10 + 2(K - 1)] + [8K(M + 1) - 2 + 8I - 2], \quad (19)$$

where I is the output samples memory length.

3.2. Numerical analysis

To evaluate the proposed NARX DPD performance, simulations with different signals and PA models are carried out. Two signals are used in simulations: single carrier LTE and two-carrier LTE both with 20MHz bandwidth and IAR of 6.2dB at 0.01% CCDF. The input back-off is 5 dB. To obtain the NMSE results, 61,440 samples have been captured and used for modeling performance analysis.

In the simulations, two different PA models have been used. The first model is based on the MP [4] given as:

$$y[n] = \sum_{k=1}^K \sum_{m=0}^M b_{km} u[n-m] |u[n-m]|^{k-1}, \quad (20)$$

where K and M are the nonlinearity order and the memory length, respectively, and b_{km} is the complex coefficients. The AM-AM and AM-PM with and without DPD is shown in Figure 4. The other PA model is based on ARCTAN [19] with two cascaded filters, which can be described as follows:

$$y_1[n] = 0.9x[n] + 0.1x[n-1], \quad (21)$$

$$y_2[n] = 20 \arctan(0.7|y_1[n]|) \exp(i\angle y_1[n]), \quad (22)$$

$$y[n] = 0.9y_2[n] + 0.1y_2[n-1], \quad (23)$$

where $x[n]$ and $y[n]$ are the normalized input and output samples, respectively. To model short term memory effects, two filters as given by (21) and (23) are included to the ARCTAN model. The

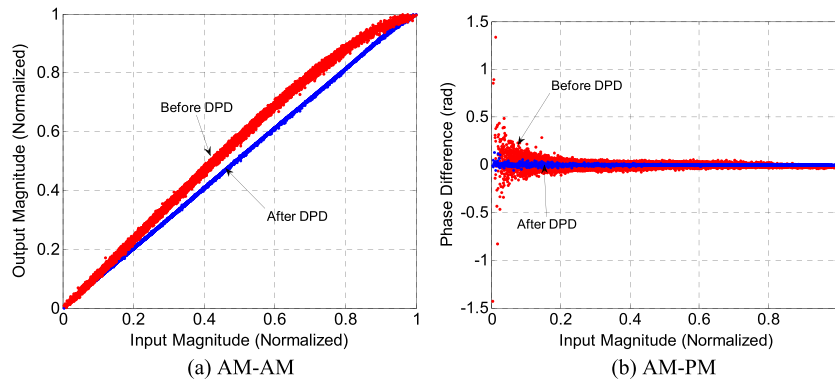


Figure 4. Power amplifier characteristics based on memory polynomial model.

obtained coefficients in (21) to (23) show a good representation of a practical PA. Figure 5 shows the AM-AM and AM-PM with and without DPD. It can be observed that the ARCTAN model exhibits more memory effect than the MP.

Table I shows the NMSE comparison between the proposed DPD and the conventional pre-distortion technique with a different number of coefficients for single carrier LTE signal. From the table, NARX DPD outperforms the other pre-distortion techniques in modeling accuracy and number of FLOPs. It can be seen that the NMSE value of NARX DPD is even lower with less number of coefficients compared with MP and GMP. Therefore, lower complexity can be achieved by using NARX DPD, which is shown by calculating the number of FLOPs per sample. The comparison has been shown for different nonlinearity order $K=5$ and 10 . The number of coefficients is reasonable according to the hardware constraints. The important consideration, however, is the order of nonlinearity (K), which should be the same for all techniques in order to have a fair comparison. Similar results can be obtained for two-carrier LTE signal, which shows comparable improvement by using NARX DPD.

Figure 6 shows the ACLR comparison between the three techniques: MP, GMP, and the proposed DPD technique. From this figure, the ACLR performance comparison shows slightly better

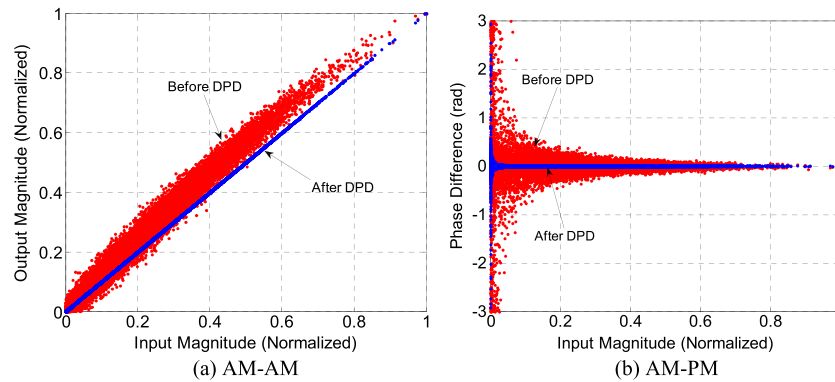


Figure 5. Power amplifier characteristics based on ARCTAN model.

Table I. NMSE comparison of MP, GMP and NARX DPD for single carrier LTE signal.

Technique	No. of coefficients	NMSE (dB)	FLOPs
MP			
($K=5, M=5$)	30	-46.88	256
($K=10, M=9$)	100	-47.65	826
GMP			
$\left(\begin{array}{l} K=5, M=5 \\ K_a=2, M_a=2, L_a=2 \\ K_b=2, M_b=1, L_b=1 \end{array} \right)$	40	-47.04	292
$\left(\begin{array}{l} K=10, M=9 \\ K_a=2, M_a=2, L_a=2 \\ K_b=2, M_b=1, L_b=1 \end{array} \right)$	110	-47.81	898
NARX			
($K=5, M=3, I=1$)	21	-47.73	182
($K=10, M=4, I=1$)	51	-48.3	432

NMSE, normalized mean square error; MP, memory polynomial; GMP, generalized memory polynomial; NARX DPD, nonlinear autoregressive exogenous digital pre-distortion; LTE, long-term evolution.

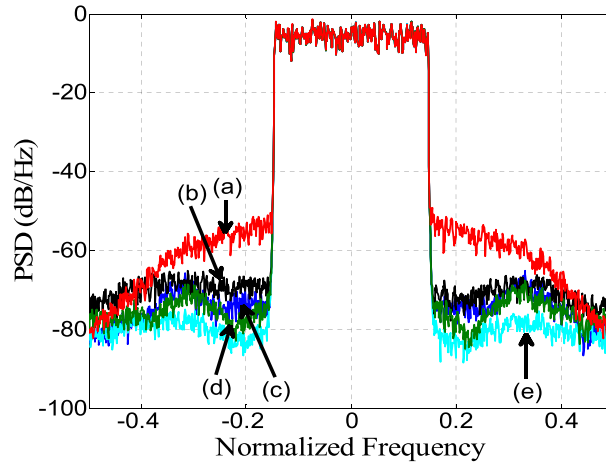


Figure 6. Adjacent channel leakage ratio comparison of a single carrier long-term evolution signal with memory polynomial power amplifier model (a) without digital pre-distortion (b) memory polynomial (c) generalized memory polynomial (d) nonlinear autoregressive exogenous digital pre-distortion ($K=5$, $M=3$, $I=1$) (e) nonlinear autoregressive exogenous digital pre-distortion ($K=5$, $M=4$, $I=1$).

performance with NARX DPD. Further improvement can be achieved when NARX DPD with 26 coefficients is applied; however, the complexity reduction is minor.

Figure 7 shows the ACLR comparison results for two-carrier LTE signal. The result of NARX DPD with 26 coefficients is almost the same with 21 coefficients, hence is not shown in this figure. It can be observed that there is a discontinuity at -0.3 frequency offset. This also exists in the original spectrum of the output signal without DPD. The main reason of this discontinuity is due to the fact that the signal is two carrier and the Third order inter-modulation (IM3) occurs at that point, but because of the asymmetric spectrum as a result of memory effects only in one side is more obvious. The other observation is the different nonlinearity in Figures 6 and 7, which is due to the two different PA used in the simulation.

Table II shows the ACLR comparison for both LTE and LTE signals for different DPD techniques. The results show for both upper band and lower band. It can be observed from this table that NARX DPD has lower ACLR compared with other DPD techniques. The aforementioned analysis of ACLR and NMSE shows that the proposed NARX DPD technique has excellent stability and lower spectral leakage for different degrees of coefficients.

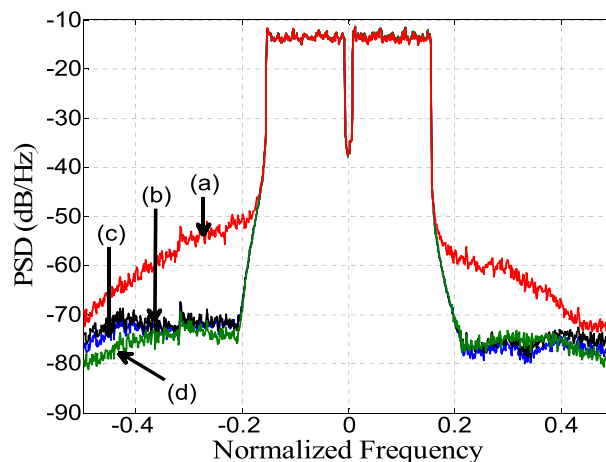


Figure 7. Adjacent channel leakage ratio comparison of a two-carrier long-term evolution signal with ARCTAN power amplifier model (a) without digital pre-distortion (b) memory polynomial (c) generalized memory polynomial (d) nonlinear autoregressive exogenous digital pre-distortion ($K=5$, $M=3$, $I=1$).

Table II. ACLR comparison of different DPD techniques.

Technique	Single carrier LTE ACLR (dBc) (U/L)	Two carrier LTE ACLR (dBc) (U/L)
No DPD	-45/-40	-40/-39
MP	-65/-68	-64/-59
GMP	-69/-70	-64/-59
NARX DPD	-72/-71	-63/-61

ACLR, adjacent channel leakage ratio; DPD, digital pre-distortion; LTE, long-term evolution; MP, memory polynomial; GMP, generalized memory polynomial; NARX DPD, nonlinear autoregressive exogenous digital pre-distortion.

4. CONCLUSIONS

A pre-distortion scheme based on NARX DPD to reduce the computational complexity of the conventional pre-distortion technique is proposed. Indirect learning architecture based on the LS estimation is used to identify the NARX DPD coefficients. The stability criterion of the proposed DPD technique is theoretically derived and verified by evaluating the NMSE and ACLR metrics. Simulation results show 30% reduction in the number of FLOPs with comparable ACLR and NMSE performance with Volterra series techniques. The results with two different wideband signals and two PA models show excellent stability by applying the proposed NARX DPD technique.

ACKNOWLEDGEMENT

The authors would like to thank Science Foundation Ireland for the financial support under grant SFI10/CE/I1853.

APPENDIX

By rewriting the NARX DPD in (5), then

$$u[n] = \sum_{i=1}^I a_i u[n-i] + \sum_{k=1}^K \sum_{m=0}^M b_{km} v[n-m] |v[n-m]|^{k-1} = \sum_{i=1}^I a_i u[n-i] + \sum_{m=0}^M b_m z[n-m], \quad (24)$$

where $z[n]$ is $z[n] = \sum_{k=1}^K b_k v[n] |v[n]|^{k-1}$.

The generalized frequency response function using recursive computation algorithm is

$$H_k(j\omega_1, \dots, j\omega_k) = \frac{\sum_{m=1}^M [b_0 + b_m \exp[-j(\omega_1 + \dots + m\omega_k)]]}{\left\{ 1 - \sum_{i=1}^I a_i \exp[-j(\omega_1 + \dots + \omega_k)i] \right\}}. \quad (25)$$

The frequency transform of (24) can be derived as

$$|U(j\omega)| \leq \sum_{k=1}^K \frac{H_k(j\omega_1^*, \dots, j\omega_k^*)}{\sqrt{k}(2\pi)^{k-1}} \left| \sum_{\omega_1 + \dots + \omega_n = \omega} Z(j\omega) d\sigma_\omega \right| \leq \sum_{k=1}^K \frac{1}{(2\pi)^{k-1}} |H_k(j\omega_1^*, \dots, j\omega_k^*)| \underbrace{|k| |Z| * \dots * |Z(j\omega)|}_k, \quad (26)$$

where $(j\omega_1^*, \dots, j\omega_k^*)$ is a point on the hyperplane $\omega_1 + \dots + \omega_n = \omega$, and σ_ω denotes a small unit on the hyperplane.

The real number coefficients $c_{n0}, c_{n1}, \dots, c_{nm}$ can be expressed by

$$|H_k(j\omega_1, \dots, j\omega_k)| \leq c_{k0} + c_{k1}\bar{H}_1 + c_{k2}\bar{H}_1^2 + \dots + c_{kk}\bar{H}_1^k. \quad (27)$$

Now by using the operator \oplus and by knowing that $Z(j\omega)$ is bounded by $V(j\omega)$ because of its finite impulse response characteristics, the inequality in (26) yields

$$|U(j\omega)| \leq \bigoplus_{k=1}^K \frac{1}{(2\pi)^{k-1}} c_k \cdot h_k^T \cdot |Z| * \dots * |Z(j\omega)| \leq \bigoplus_{k=1}^K \frac{1}{(2\pi)^{k-1}} \cdot c_k \cdot h_k^T \cdot |V| * \dots * |V(j\omega)| = \left(\bigoplus_{k=1}^K \alpha_k c_k \right) \cdot h_k^T, \quad (28)$$

which can be rewritten as

$$|U(j\omega)| \leq \bar{U} = \left(\bigoplus_{k=1}^K \alpha_k c_k \right) (0) + \left(\bigoplus_{k=1}^K \alpha_k c_k \right) (1) \bar{H}_1 + \left(\bigoplus_{k=1}^K \alpha_k c_k \right) (2) \bar{H}_1^2 + \dots + \left(\bigoplus_{k=1}^K \alpha_k c_k \right) (1) \bar{H}_1^k + \dots. \quad (29)$$

The stability of the NARX DPD is guaranteed by the convergence of series in (29). The condition of convergence of the series in (25) can be obtained by Cauchy's criterion as follows:

$$\lim_{\substack{K \rightarrow \infty \\ g \rightarrow \infty}} \sqrt[g]{\left(\bigoplus_{k=1}^K \alpha_k c_k \right) (g) \bar{H}_1^g} < \bar{H}_1 \lim_{g \rightarrow \infty} \sqrt[g]{\left(\bigoplus_{k=1}^{\infty} \alpha_k c_k \right) (g)} < 1. \quad (30)$$

This completes the proof.

REFERENCES

1. Gharaibeh KM, Al-Zayed AS. Performance of feed-forward linearizers of power amplifiers in OFDM systems under complex gain errors. *International Journal of Communication Systems* 2016; **29**(4):734–747.
2. Fehri B, Boumaiza S. Baseband equivalent Volterra series for behavioral modeling and digital predistortion of power amplifiers driven with wideband carrier aggregated signals. *IEEE Transactions on Microwave Theory and Techniques* 2014; **62**(11):2594–2603.
3. Kang HW, Cho YS, Youn DH. On compensating nonlinear distortions of an OFDM system using efficient adaptive predistorter. *IEEE Transactions on Communications* 1999; **47**(4):522–526.
4. Kim J, Konstantinou K. Digital predistortion of wideband signals based on power amplifier model with memory. *Electronics Letters* 2001; **37**(23):1417–1418.
5. Morgan DR, Ma Z, Kim J, Zierdt MG, Pastalan J. A generalized memory polynomial model for digital predistortion of RF power amplifiers. *IEEE Transactions on Signal Processing* 2006; **54**(10):3852–3860.
6. Billings BA. *Nonlinear System Identification, NARMAX Methods in the Time, Frequency, and Spatio-temporal Domains*, Wiley, 2013.
7. Dooley J, Brien BO, Brazil TJ. Behavioral modeling of RF power amplifiers using adaptive recursive polynomial functions 2006. In *Proc. IEEE MTT-S Int. Micro. Symp. Dig.*, 2006; 852–855.
8. Gilbert PL, Montoro G, Cesari A. A recursive digital predistorter for linearizing RF power amplifiers with memory effects. In *Proc. Asia-Pacific Microwave Conference*, 2006; 1040–1043.
9. Aguilar-Lobo L, Garcia-Osorio A, Loo-Yau J, Ortega-Cisneros S, Moreno P, Rayas-Sanchez J, Reynoso-Hernandez J. A digital predistortion technique based on a NARX network to linearize GaN class F power amplifiers. In *IEEE 57th International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2014; 717–720.
10. Abd-Elrady E, Gan L, Kubin G. A nonlinear filtered-x prediction error method algorithm for digital predistortion in digital subscriber line systems. *International Journal of Communication Systems* 2013; **26**(5):651–661.
11. Varahram P, Dooley J, Finnerty K, Farrell R. A digital pre-distortion based on nonlinear autoregressive with exogenous inputs. Accepted in *IEEE Microwave and Wireless Components Letter*, 2016.
12. Kenington PB. *High-Linearity RF Amplifier Design*, MA: Boston Artech House, 2000.
13. Jing X, Lang Z. *Frequency Domain Analysis and Design of Nonlinear Systems based on Volterra Series Expansion, Understanding Complex Systems*, Springer, 2015.
14. Song J, Ochiai H. A low-complexity peak cancellation scheme and its FPGA implementation for peak-to-average power ratio reduction. *EURASIP Journal on Wireless Communications and Networking* 2015; **85**: 1–14.

15. Kim BW, Jung SY. Peak-to-average power ratio reduction of a hidden training sequence-aided precoding scheme for MIMO-OFDM. *International Journal of Communication Systems* 2014. DOI:10.1002/dac.2910.
16. Liang H, Chu HC, Lin CB. Peak-to-average power ratio reduction of orthogonal frequency division multiplexing systems using modified tone reservation techniques. *International Journal of Communication Systems* 2016; **29**(4):748–759.
17. Varahram P, Ali BM. Peak to average power ratio reduction and digital predistortion effects in power amplifiers in OFDM systems. *International Journal of Communication Systems* 2012; **25**(4):543–552.
18. Tehrani AS, Haiying C, Afsardoost S, Eriksson T, Isaksson M, Fager C. A comparative analysis of the complexity/accuracy tradeoff in power amplifier behavioral models. *IEEE Transactions on Microwave Theory and Technique* 2010; **58**(6):1510–1520.
19. Raich R, Qian H, Zhou GT. Orthogonal polynomials for power amplifier modeling and predistorter design. *IEEE Transactions on Vehicular Technology* 2004; **53**(5):1468–1479.