A Direct Learning Structure for Adaptive Polynomial-based Predistortion for Power Amplifier Linearization

Sungho Choi, Eui-Rim Jeong, and Yong H. Lee

School of EECS, Korea Advanced Institute of Science and Technology
373-1 Guseong-dong, Yuseong-gu, Daejeon, 305-701, Republic of Korea
Email: {shchoi, july}@stein.kaist.ac.kr, yohlee@ee.kaist.ac.kr

***Abstract*—A new polynomial predistortion technique for linearizing nonlinear power amplifier is introduced. The proposed predistorter is based on direct learning, whereas most other techniques are based on the indirect learning architecture. The resulting structure of the proposed method is simpler than the existing ones. Computer simulation shows that the proposed method is robust to the initial conditions and attains comparable performance to other methods based on the indirect learning architecture under normal conditions.**

***Index Terms*—Predistortion, polynomial, RLS**

I. Introduction

Due to demand for the high spectral efficiency in mobile communications, the transmitted signal tends to have a high peak-to-average ratio, and linearization of the power amplifier (PA) has become an important issue. A popular linearization technique is the baseband digital predistortion (PD), which is highly cost-effective [1]–[9].

Various digital PD techniques have been proposed, and they can be classified into lookup table (LUT) methods [1]–[4] and polynomial methods [5]–[7]. The LUT methods realize the nonlinear PD characteristic with a set of piecewise linear functions. The main defect of the LUT techniques is the long convergence time in proportion to the LUT size. In contrast, the polynomial methods represent the PD function with a polynomial of several order at most. The polynomial methods can converge faster than the LUT methods while exhibiting good performance. Moreover, they can be easily extended to a wideband PA model, which has the memory effect [8]–[9]. In most polynomial PD techniques, the coefficients of the polynomial are effectively found via the indirect learning architecture [6]–[9], where the coefficient adaptation is done at the *postdistorter* (or training block) in the feedback path and the coefficients are copied to the actual PD in the feed forward path. Therefore, those techniques have somewhat complex structure since the duplicate of the PD is needed in the feedback path. Also, the indirect learning architecture turned out to be sensitive to the initial conditions.

In this paper, we develop a new polynomial-based predistorter based on direct learning. The proposed method directly adjusts the PD polynomial in the feedforward path without the need for the postdistorter. The proposed method is derived for memoryless PA and then extended to the memory PA case. The proposed technique has a simpler structure than the existing indirect learning architecture, and it will be shown through computer simulation that it is more robust to the initial conditions of the tap coefficients than the existing one.

The organization of this paper is as follows. The proposed PD structure is presented in Section II. The proposed PD algorithm is developed in Section III. Section IV presents simulation results and some comparisons between the proposed and indirect learning architecture. Conclusions appear in Section V.

II. Proposed PD Structure

The PD structure based on the indirect learning architecture [7]–[9] and the proposed one are compared in Fig. 1. The indirect learning architecture in Fig. 1(a) adaptively adjusts the postdistorter parameter which linearizes the PA-postdistorter chain rather than the PD-PA chain. Then, this parameter is copied to the PD block at every time samples [7] or once in a while [9]. On the contrary, the proposed structure in Fig. 1(b) has no postdistorter. It directly adjusts the PD parameter. Therefore, the proposed structure is simpler than the existing architecture. The indirect learning architecture minimizes $\xi'(n)$, as shown in Fig. 1(a). If the initial coefficient of the PD is zero vector, which is a common initial condition for an adaptive algorithm, the PD output will be zero and consequently the PA output is also zero. In this case, the adaptive algorithm does not work since $\xi'(n)$ is already zero. In contrast, since the proposed structure uses the PD input signal as the reference, it is expected that the proposed method is more robust to the initial conditions. The sensitivity to the initial conditions will be more apparent in Section IV through computer simulation.

The description of the proposed PD structure is as follows. Transmitted symbols are band-limited by a PSF (pulse shaping filter) to produce $x(n)$, which is predistorted to yield $y(n)$. Then, $y(n)$ is amplified by the power amplifier of desired gain K. The amplifier output divided by the gain K is fed back into adaptive algorithm block. In the baseband-equivalent

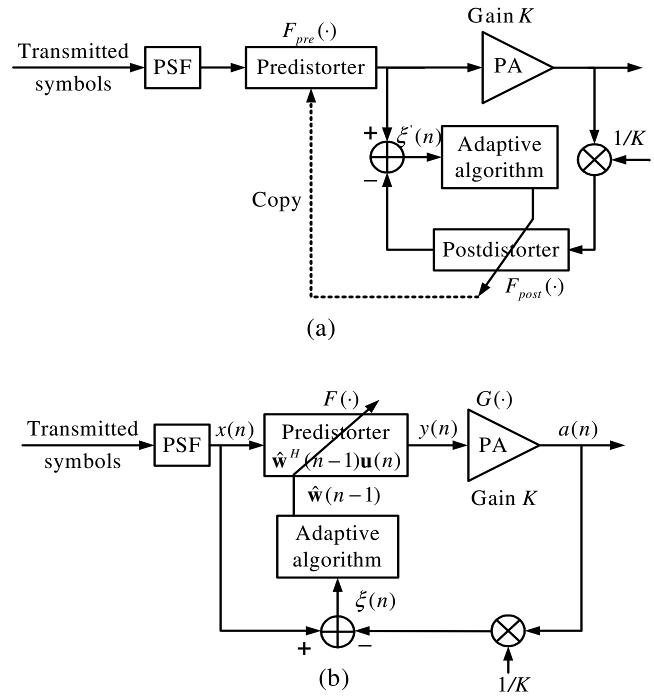

Fig. 1. Transmitter model employing PD. (a) indirect learning architecture [7]–[9] (b) proposed architecture.

system model in Fig.1(b), the RF (radio frequency) up/down conversion is assumed to be ideal. The adaptive algorithm finds the PD parameters to make the PSF output and the feedback signal as close as possible.

If the PA is memoryless, its characteristic is represented by the AM-AM and AM-PM responses [12]. AM-AM and AM-PM responses represent the PA output magnitude and phase as a function of the input magnitude, respectively. If we define the PA characteristic as $G(\cdot)$, the PA output $a(n)$ can be written as

$$a(n) = G(y(n)). \tag{1}$$

Similarly, if we define the PD function as $F(\cdot)$, $y(n)$ can be represented as

$$y(n) = F(x(n)). \tag{2}$$

From (1) and (2), the ideal PD satisfies the following relation:

$$a(n) = G(F(x(n))) = Kx(n). \tag{3}$$

The purpose of this paper, hence, is to find the PD characteristic $F(\cdot)$ that hold (3).

III. Development of The Proposed Algorithm

If the PD is composed of P'th order complex polynomial, F can be expressed as

$$F(x(n)) = \sum_{p=0}^{P-1} w_p^*(n)x(n)|x(n)|^p = \mathbf{w}^H(n)\mathbf{u}(n) \tag{4}$$

where

$$\mathbf{w}(n) = [w_0(n), w_1(n), ..., w_{P-1}(n)]^T$$
$$\mathbf{u}(n) = x(n)[|x(n)|^0, |x(n)|^1, ..., |x(n)|^{P-1}]^T.$$

The PA output, consequently, is written as

$$a(n) = G(\mathbf{w}^H(n)\mathbf{u}(n)). \tag{5}$$

If the PD works ideally, the PA output will be $Kx(n)$ and the corresponding PA input can be represented as $G^{-1}(Kx(n))$. $G^{-1}(Kx(n))$, of course, becomes the ideal PD output for the input $x(n)$. To derive the proposed PD algorithm, we define the error signal $e(n)$ as the difference between the ideal PD output $G^{-1}(Kx(n))$ and the actual output $\mathbf{w}^H(n)\mathbf{u}(n)$:

$$e(n) = G^{-1}(Kx(n)) - \mathbf{w}^H(n)\mathbf{u}(n). \tag{6}$$

The proposed algorithm is developed based on the least squares (LS) criterion. Therefore, the cost function to be minimized is defined by the sum of error squares as follows

$$\mathcal{E}(n) = \sum_{i=1}^{n} \lambda^{n-i}|e(i)|^2 \tag{7}$$

where λ is a forgetting factor taking the values between 0 and 1. Since $\mathcal{E}(n)$ is a quadratic function with respect to $\{w_p(n)|p = 0, ..., P-1\}$, the minimum can be obtained by differentiating (7) and equating it to zero, which yield the following normal equation:

$$\mathbf{\Phi}(n)\hat{\mathbf{w}}(n) = \mathbf{z}(n) \tag{8}$$

where

$$\mathbf{\Phi}(n) = \sum_{i=1}^{n} \lambda^{n-i}\mathbf{u}(n)\mathbf{u}^H(n) \tag{9}$$

$$\mathbf{z}(n) = \sum_{i=1}^{n} \lambda^{n-i}\mathbf{u}(n)G^{-1*}(Kx(n)). \tag{10}$$

$\hat{\mathbf{w}}(n)$ in (8) represents the coefficient of the PD polynomial that minimizes (7). Further, if we employ a recursive least squares (RLS) method to get $\mathbf{\Phi}^{-1}(n)$ instead of direct inversion [11], the RLS PD algorithm can be derived as follows:

$$\boldsymbol{\kappa}(n) = \frac{\lambda^{-1}\mathbf{Q}(n-1)\mathbf{u}(n)}{1 + \lambda^{-1}\mathbf{u}^H(n)\mathbf{Q}(n-1)\mathbf{u}(n)} \tag{11}$$

$$\xi(n) = G^{-1}(Kx(n)) - \hat{\mathbf{w}}^H(n-1)\mathbf{u}(n) \tag{12}$$

$$\hat{\mathbf{w}}^H(n) = \hat{\mathbf{w}}^H(n-1) + \boldsymbol{\kappa}(n)\xi^*(n) \tag{13}$$

$$\mathbf{Q}(n) = \lambda^{-1}\mathbf{Q}(n-1) - \lambda^{-1}\boldsymbol{\kappa}(n)\mathbf{u}^H(n)\mathbf{Q}(n-1). \tag{14}$$

The RLS PD algorithm in (11) – (14), however, is not practicable since the calculation of *a priori error*, $\xi(n)$, requires the inverse characteristic of the PA, which is generally unavailable. To modify the algorithm for practical use, the following approximation is assumed:

Assumption I: PA characteristic $G(y(n))$ can be represented as a piecewise linear function which has M regions, as shown in Fig. 2. When $y(n)$ is in the m'th region, $G(y(n))$ can be represented as follows:

$$G_m(y(n)) \approx K_m(y(n) - b_m) + c_m, \quad 1 \le m \le M \quad (15)$$

and its inverse

$$G_m^{-1}(y(n)) \approx \frac{1}{K_m}(y(n) - c_m) + b_m, \quad 1 \le m \le M \quad (16)$$

where K_m and (b_m, c_m) are is the slope and the crossover coordinate value in the m'th region, respectively.

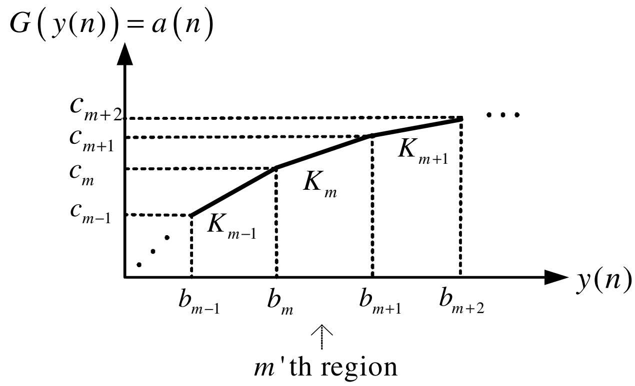

Fig. 2. Piecewise linear model of PA

Using this assumption, when the input of the PA is in the m'th region, *a priori error* can be written by

$$\xi(n) = G^{-1}(Kx(n)) - \hat{\mathbf{w}}^H(n-1)\mathbf{u}(n) \quad (17)$$

$$\approx \frac{1}{K_m}\left(Kx(n) - (K_m(\hat{\mathbf{w}}^H(n-1)\mathbf{u}(n) - b_m) + c_m)\right) \quad (18)$$

$$\approx \frac{1}{K_m}\left(Kx(n) - a(n)\right) \quad (19)$$

where Assumption I is used in (18) and (19). From (19), the only thing we need is the slope K_m. However, since we still do not know the PA characteristic, the slope K_m should be estimated. The approximate grade of the slope K_m at time n is estimated by evaluating the ratio of the PA input difference and output difference of the current and previous samples. However, there is a numerical problem in this approach: estimate $K_m(n)$ will be infinite when the current and previous input are the same. In order to avoid this problem, previous estimate $K_m(n-1)$ is used in this case. Finally, the slope $K_m(n)$ is estimated as follows:

$$K_m(n) = \begin{cases} \frac{a(n)-a(n-1)}{y(n)-y(n-1)}, & \text{for } y(n) \neq y(n-1) \\ K_m(n-1), & \text{for } y(n) = y(n-1). \end{cases} \quad (20)$$

Using (20), $\xi(n)$ in (12) can be replaced by

$$\xi(n) = \frac{K}{K_m(n)}\left(x(n) - \frac{a(n)}{K}\right). \quad (21)$$

TABLE I
SUMMARY OF PROPOSED RLS PREDISTORTION ALGORITHM.

Initialization:
$\mathbf{Q}(n) = c^{-1}\mathbf{I}$ where c is a small positive constant.
$\hat{\mathbf{w}}(0) = [\alpha, 0, ..., 0]^T$ where α is a constant.

For $n = 1, 2, ...$, compute followings:
$\boldsymbol{\kappa}(n) = \frac{\lambda^{-1}\mathbf{Q}(n-1)\mathbf{u}(n)}{1+\lambda^{-1}\mathbf{u}^H(n)\mathbf{Q}(n-1)\mathbf{u}(n)}$
$K_m(n) = \begin{cases} \frac{a(n)-a(n-1)}{\hat{\mathbf{w}}^H(n-1)\mathbf{u}(n)-\hat{\mathbf{w}}^H(n-2)\mathbf{u}(n-1)} \\ \quad \text{for } \hat{\mathbf{w}}^H(n-1)\mathbf{u}(n) \neq \hat{\mathbf{w}}^H(n-2)\mathbf{u}(n-1) \\ K_m(n-1) \\ \quad \text{for } \hat{\mathbf{w}}^H(n-1)\mathbf{u}(n) = \hat{\mathbf{w}}^H(n-2)\mathbf{u}(n-1) \end{cases}$
$\xi(n) = \frac{K}{K_m(n)}\left(x(n) - \frac{a(n)}{K}\right)$
$\hat{\mathbf{w}}^H(n) = \hat{\mathbf{w}}^H(n-1) + \boldsymbol{\kappa}(n)\xi^*(n)$
$\mathbf{Q}(n) = \lambda^{-1}\mathbf{Q}(n-1) - \lambda^{-1}\boldsymbol{\kappa}(n)\mathbf{u}^H(n)\mathbf{Q}(n-1)$

Table I summarizes the proposed RLS predistortion algorithm employing a new representation of the error in (21).

To consider PA memory effects, which cannot be ignored for wideband applications, the PD is designed by a memory polynomial similar to [9]. Let the polynomial order and memory depth be P and Q, respectively. The PD output can be represented as

$$F(x(n)) = \sum_{q=0}^{Q}\sum_{p=0}^{P-1} w_{p,q}^*(n)x(n-q)|x(n-q)|^p = \mathbf{w}^H(n)\mathbf{u}(n) \quad (22)$$

where

$$\mathbf{w}(n) = [\mathbf{w}_0^T(n), \mathbf{w}_1^T(n), ..., \mathbf{w}_Q^T(n)]^T$$
$$\mathbf{w}_q(n) = [w_{0,q}(n), w_{1,q}(n), ..., w_{P-1,q}(n)]^T$$
$$\mathbf{u}(n) = [\mathbf{u}_1^T(n), \mathbf{u}_2^T(n), ..., \mathbf{u}_Q^T(n)]^T$$
$$\mathbf{u}_q(n) = x(n-q)[1, |x(n-q)|^1, ..., |x(n-q)|^{P-1}]^T.$$

The RLS solution for obtaining the optimal $\mathbf{w}(n)$ is identical to the RLS algorithm for the memoryless case except for the increased vector size proportional to the memory depth Q.

IV. SIMULATION RESULTS

In this section, the performance of the proposed predistorter is demonstrated through computer simulation, and the effect of the initial condition is compared between the proposed one and the PD in [7]. The simulation environments is as follows: The transmitted symbols are modulated by 16-QAM, and the PSF is a 10-times oversampled RRC (square root raised cosine) filter with a roll-off factor 0.22. For the memoryless PA model, we use Saleh's model [10], which is given by

$$G(y(n)) = \frac{1.1y(n)}{1+0.3|y(n)|^2}\exp\left(j\frac{|y(n)|^2}{1+|y(n)|^2}\right). \quad (23)$$

The PA gain is assumed to be 1 ($K = 1$). Fig. 3 compares the mean square error (MSE) for the proposed PD and the method in [7]. The MSE is obtained by averaging the results of over 500 trials. The forgetting factor, λ, is 0.95 for both methods.

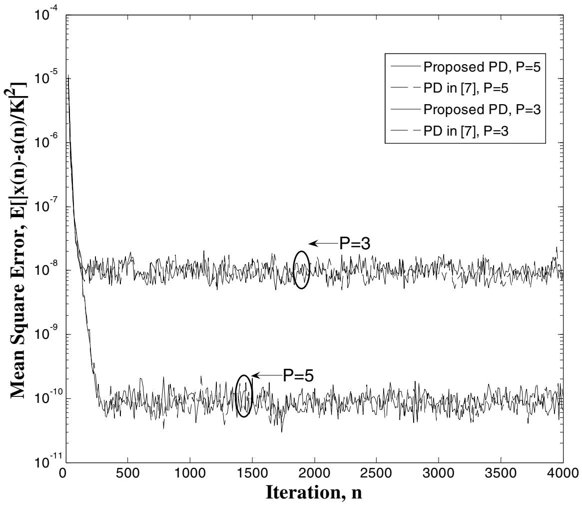

Fig. 3. Learning curves for the proposed PD and PD in [7].

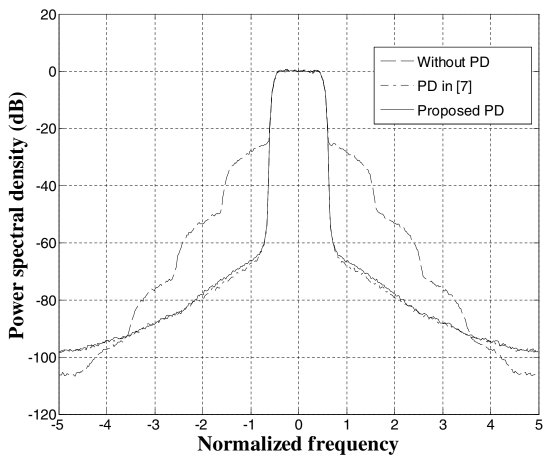

Fig. 4. Comparison of power spectral density at PA output ($P = 3$).

The peak signal is backed-off from the PA's saturation point by 1.5dB, i.e., the peak back-off (PBO) is 1.5dB. The initial coefficient $\hat{\mathbf{w}}(0) = [1, 0, ..., 0]^T$ is used ($\alpha = 1$ in Table I). The larger polynomial order yields the better performance at the expense of longer convergence time. This result shows that the proposed PD has comparable performance with low computational complexity. Fig. 4 shows the power spectral density (PSD) at the PA output. The PSD is captured after 4000 iterations. The PBO is 1.5dB, and the polynomial order is 3 ($P = 3$). As shown in the figure, the proposed PD can improve the out-of-band PSD up to 40dB.

We also compare the effect of the initial condition in Fig. 5. This figure shows MSE after convergence for various initial conditions by varying the first tap α of the $\mathbf{w}(0)$. The PBO is 1.5dB, and the polynomial order is 3 ($P = 3$). The PD in [7] does not converge when the first tap α of $\mathbf{w}(0)$ is near 0 or greater than 3.5. In contrast, the proposed method converges well in the range of 0 to 5. This result indicates that the proposed PD is more robust to the initial condition than the existing PD in [7].

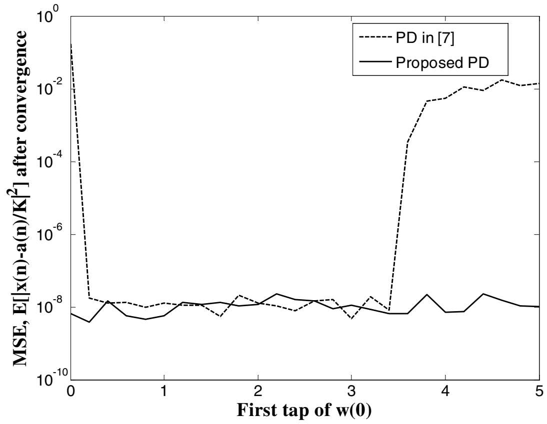

Fig. 5. The effect of the initial condition.

V. Conclusion

A new PA linearization method based on a polynomial PD is developed. The proposed method finds the PD parameter directly from the transmitted and feedback signals while the existing polynomial PD resorts to the indirect learning architecture. The resulting structure is simpler than the existing methods. Simulation results showed that the proposed PD is more robust to the initial coefficient of the PD than the methods based on the indirect learning architecture, and in normal cases, both perform comparably.

Acknowledgement

This work was supported in part by the University Information Technology Research Center Program of the government of Korea.

References

[1] J. K. Cavers, "Amplifier linearization using a digital predistorter with fast adaptation and low memory requirements," *IEEE Trans. Veh. Technol.*, vol. 39, no. 4, pp. 374-382, Nov. 1990.

[2] A. S. Wright and W. G. Durtler, "Experimental performance of an adaptive digital linearized power amplifier," *IEEE Trans. Veh. Technol.*, vol. 41, no. 4, pp. 395-400, Nov. 1992.

[3] K. J. Muhonen, M. Kavehrad and R. Krishnamoorthy, "Look-up table techniques for adaptive digital predistortion: a development and comparison," *IEEE Trans. Veh. Technol.*, vol. 49, no. 5, pp. 1995-2002, Sept. 2000.

[4] A. N. D'Andrea, V. Lottici and R. Reggiannini, "Efficient digital predistortion in radio relay links with nonlinear power amplifiers," *IEE Proc. Commun.*, vol. 147, no. 3, pp. 175-179, June 2000.

[5] M. Ghaderi, S. Kumar and D. E. Dodds, "Fast adaptive polynomial I and Q predistorter with global optimization," *IEE Proc. Commun.*, vol. 143, no. 2, pp. 78-86, Apr. 1996.

[6] Y. Qian and T. Yao, "Structure for adaptive predistortion suitable for efficient adaptive algorithm application," *Electron. Lett.*, vol. 38, no. 21, pp. 1282-1283, Oct. 2002.

[7] R. Marsalek, P. Jardin, and G. Baudoin, "From post-distortion to predistortion for power amplifiers linearization," *IEEE Commun. Letters*, vol. 7, no. 7, pp. 308-310, July 2003.

[8] J. Kim and K. Konstantinou, "Digital predistortion of wideband signals based on power amplifier model with memory," *Electron. Lett.*, vol. 37, no. 23, pp. 1417-1418, Nov. 2001.
[9] L. Ding, G. T. Zhou, D. R. Morgan, Z. Ma, S. Kenney, J. Kim, and C. R. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," *IEEE Trans. Commun.*, vol. 52, no. 1, pp. 159-165, Jan. 2004.
[10] A. A. M. Saleh, "Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers," *IEEE Trans. Commun.*, vol. COM-29, no. 11, pp. 1715-1720, Nov. 1981.
[11] S. Haykin, *Adaptive filter theory*, Prentice Hall, 1996.
[12] N. Pothecary, *Feedforward linear power amplifier*, Artech House, 1999.