Title	Forward modeling assisted 1-Bit data acquisition based model extraction for digital predistortion of RF power amplifiers
Authors(s)	Wang, Haoyu; Li, Gang; Zhang, Yikang; Liu, Falin; Zhu, Anding
Publication date	2017-01-18
Conference details	2017 IEEE Topical Conference on RF/microwave Power Amplifiers for Radio and Wireless Applications (PAWR), Phoenix, Arizona, USA, 15-18 January 2017
Publisher	IEEE
Item record/more information	http://hdl.handle.net/10197/8492
Publisher's statement	© 2017 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
Publisher's version (DOI)	10.1109/PAWR.2017.7875573

Forward Modeling Assisted 1-Bit Data Acquisition Based Model Extraction for Digital Predistortion of RF Power Amplifiers

Haoyu Wang[1,2], Gang Li[1,2], Yikang Zhang[1,2], Falin Liu[1,2], and Anding Zhu[3]

[1]Department of EEIS, University of Science and Technology of China, Hefei 230027, China
[2]Key Laboratory of Electromagnetic Space Information, Chinese Academy of Sciences, Hefei, China
[3]School of Electrical and Electronic Engineering, University College Dublin, Dublin 4, Ireland

***Abstract* — In this paper, a low-cost data acquisition approach with novel forward modeling is proposed for model extraction of digital predistortion (DPD) of RF power amplifiers (PA). Compared to the existing 1-bit DPD solution, the proposed approach employs a new loss function using only 1-bit resolution error signal to extract the forward model and the coefficients of DPD model can be extracted afterwards based on the estimated PA output. The proposed solution avoids the calculation of step size in the existing approach, and the convergence is guaranteed due to the convexity of the loss function. The experimental results show that the proposed algorithm achieves equivalent linearization performance as that using the conventional method.**

***Index Terms* — analog-to-digital converter (ADC), digital predistortion (DPD), model extraction, power amplifier (PA).**

I. INTRODUCTION

Digital predistortion (DPD) has become one of the most popular linearization techniques for RF power amplifiers (PA) in cellular base stations. With further development of future networks, DPD is facing new challenges [1]-[3]. In the forthcoming 5G systems, the signal bandwidth will be well over hundreds megahertz. If the conventional DPD techniques are deployed, mutli-GSPS high resolution analog-to-digital converters (ADCs) are required in the feedback data acquisition path, which is not feasible. It is because designing a high resolution, e.g. 14-bit, ADC with multi-GSPS is very challenging and costly.

In [4], we have previously proposed a 1-bit data acquisition method for model extraction of DPD. In this approach, only signs of the I/Q error signals between input and output of the system are required to be captured. This can be achieved by use 1-bit ADC or simple voltage comparators, which significantly reduce the system complexity and power consumption. However, one of the key challenges in using this approach is that the step size during the iteration process must be correctly chosen in order to guarantee fast convergence. The calculation of step size relies on the prior information of the PA and the input signal. In some situations, if no prior information of the PA or input signal is available, it can be difficult to determine a proper step size at each iteration.

In this paper, we propose an alternative approach, where a forward modeling procedure is carried out first by using the 1-bit data and then direct-learning is used to extract the DPD coefficients. In this approach, the calculation of step size is avoided and the convergence of DPD training is guaranteed because a new objective function is proposed in the modeling process.

II. 1-BIT DPD SYSTEM

A. Notations

Since the proposed DPD training method is different from the conventional one, some basic notations need to be defined in this subsection first.

Denote the bold lower-case complex vectors $\mathbf{x}$, $\mathbf{z}$ and $\mathbf{y} \in \mathbb{C}^{K\times 1}$ for the baseband input, output of DPD and output of PA, respectively, and they all have the length of K. Assume the linear-in-parameters model, e.g. the well-established Volterra series with coefficients $\boldsymbol{\alpha} \in \mathbb{C}^{L\times 1}$, is used for forward modeling, and the input-output relationship can be expressed in a matrix form as

$$\hat{\mathbf{y}} = \boldsymbol{\psi}(\mathbf{z}) \cdot \boldsymbol{\alpha}, \tag{1}$$

where $\boldsymbol{\psi}(\mathbf{z})$ is a matrix that each column of which consists of a function of $\mathbf{z}$, i.e. a basis function, describing the nonlinearity and memory effect of PA. $\hat{\mathbf{y}}$ is the output of the forward model, as shown in Fig. 1. For the DPD (inverse) model, the corresponding coefficients is denoted as $\boldsymbol{\beta} \in \mathbb{C}^{L\times 1}$. The signs of the difference between the PA output $\mathbf{y}$ and the output of forward model $\hat{\mathbf{y}}$ are $\mathbf{s} = sign(\mathbf{y} - \hat{\mathbf{y}}) = [\ldots, \pm 1 \pm \mathrm{j}, \ldots]^T$, where the operator $sign(\bullet)$ calculates the signs of real and imaginary parts of a given complex value separately.

To better understand the model identification procedure proposed in II.C, the existing complex values notations and formula needs to be converted into real values version. The real values vectors are composed of the real and imaginary parts their complex versions, e.g.

$$\tilde{\mathbf{x}} = \begin{bmatrix} \mathrm{Re}(\mathbf{x}) \\ \mathrm{Im}(\mathbf{x}) \end{bmatrix} \in \mathbb{R}^{2K\times 1}, \tilde{\boldsymbol{\alpha}} = \begin{bmatrix} \mathrm{Re}(\boldsymbol{\alpha}) \\ \mathrm{Im}(\boldsymbol{\alpha}) \end{bmatrix} \in \mathbb{R}^{2L\times 1}. \tag{2}$$

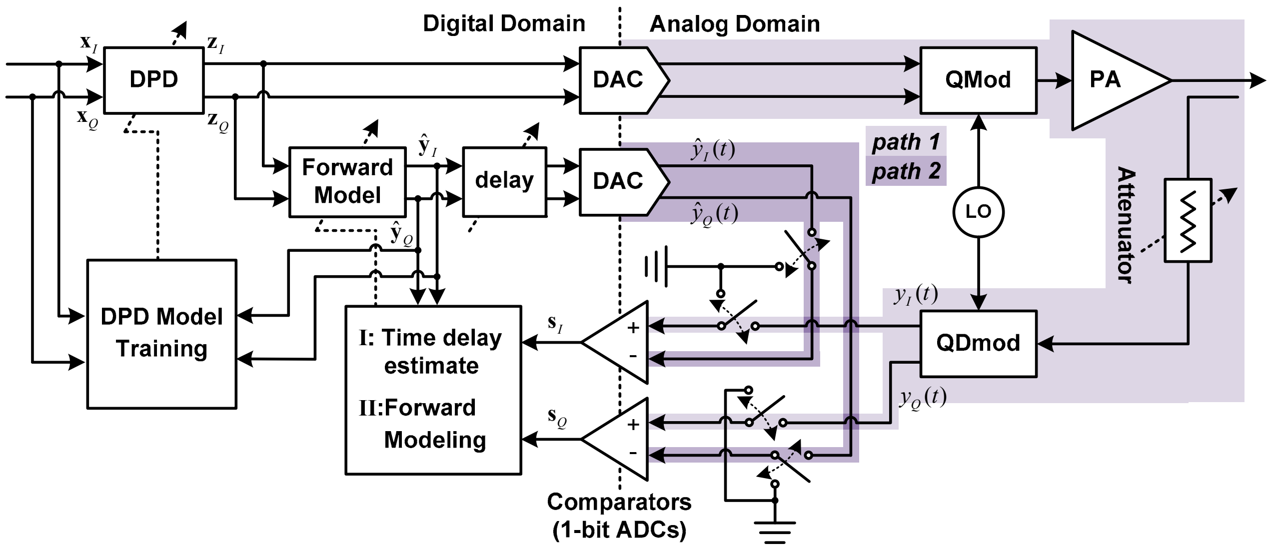

Fig. 1. Proposed 1-bit data acquisition DPD system.

And the real values version of matrix $\boldsymbol{\psi}(\mathbf{z})$ is expressed as

$$\tilde{\boldsymbol{\psi}}(\mathbf{z}) = \begin{bmatrix} \operatorname{Re}(\boldsymbol{\psi}(\mathbf{z})) & -\operatorname{Im}(\boldsymbol{\psi}(\mathbf{z})) \\ \operatorname{Im}(\boldsymbol{\psi}(\mathbf{z})) & \operatorname{Re}(\boldsymbol{\psi}(\mathbf{z})) \end{bmatrix} \in \mathbb{R}^{2K \times 2L}. \quad (3)$$

B. System Architecture

The proposed architecture of 1-bit DPD system is illustrated in Fig. 1. The DPD model identification procedure consists of two parts: (i) train the forward model and (ii) update the coefficients of DPD. Different from the conventional architecture, the coefficients of the forward model are extracted by using 1-bit data acquisition process. In this configuration, an additional digital to analog conversion path, *path 2*, as highlighted in Fig. 1, is added to convert the digital I/Q output of the forward model to the analog domain to compare with the PA output in *path 1*. The signs of the error signal are then obtained from the output of the comparators and can be used for forward model extraction. Once the forward model is identified, the high resolution output of the PA can be estimated, and the DPD model then can be extracted by using the conventional direct learning process.

Before forward modeling, time alignment needs to be conducted to compensate the time mismatch between *path 1* and *path 2*. Firstly the two switches in *path 1* are connected to comparators and the other two in *path 2* are connected to ground so that the comparators directly measure the signs of the output signal from the PA to estimate the delay in *path 1*, t_1, by using the frequency domain estimation approach proposed in [4]. We then switch *path 1* to ground and connect *path 2* to comparators to estimate the time delay in *path 2, t_2*. Once both t_1 and t_2 in the two paths are estimated, the delay block upstream of the DAC in *path 2* is switched on to delay the input signal by $t_1 - t_2$ (assuming $t_1 > t_2$), and the switches in both *path 1* and *path 2* are connected to comparators, thus the proposed system works in Mode II.

C. Loss Function of Forward Modeling

In the conventional DPD, an l_2-norm loss function, e.g., least squares error (LSE), can be used to extract the PA model coefficients. In this system, however, the LSE cannot be obtained because only the signs of the PA output can be captured. A new loss function and related model extraction procedure must be developed.

As shown in Fig. 1, the model identification includes two loops: forward model training (inner loop) and DPD model extraction (outer loop), and two loops are run alternatively in multiple iterations. Let's assume the coefficients vector of the forward model after *i*-th iteration is denoted as $\tilde{\boldsymbol{\alpha}}^{(i)}$. The signs vector of the error signal measured between the real PA output and the output of *i*-th forward model is given as $\tilde{\mathbf{s}}^{(i)}$. To estimate the new PA model coefficients, denoted by $\tilde{\boldsymbol{\alpha}}_{PA}$, at next iteration, the following two conditions must be satisfied: 1) the sign of the difference between the newly estimated PA output and the output of the *i*-th forward model is close enough to the measured sign; 2) the magnitude and the characteristics of the newly estimated PA output waveform are consistent with those of the actual one.

For the first condition mentioned above, a logarithmic function is defined as

$$f(x) = \ln\left(1 + e^{-x}\right), \quad (4)$$

where x is an arbitrary real value. With this logarithmic function, a positive number results in small function value, whereas a negative number leads to large penalty.

The input of (4) can be defined as the product of the estimated error signal, i.e., $\Delta_n^{(i)}$ as given in (5b) and the measured sign $\tilde{\mathbf{s}}_n^{(i)}$, at instance n. If the sign of this product is positive, the value of (4) will be close to zero, otherwise it will be large. Thus the loss function can be expressed as the sum of such logarithmic functions for all training samples:

$$L_1^{(i)} = \sum_{n=1}^{2K} f\left[\Delta_n^{(i)} \cdot \tilde{\mathbf{s}}_n^{(i)}\right], \tag{5a}$$

$$\Delta_n^{(i)} = \tilde{\boldsymbol{\psi}}_n(\mathbf{z}) \cdot \tilde{\boldsymbol{\alpha}}_{PA} - \tilde{\boldsymbol{\psi}}_n(\mathbf{z}) \cdot \tilde{\boldsymbol{\alpha}}^{(i)}, \tag{5b}$$

where the subscript n denotes the n-th row of the corresponding column vector and matrix. By minimizing the loss function in (5a), the coefficients of the PA model can be obtained such that the signs of the estimated error signal are close enough to the measured ones.

Only using (5) is not enough because the magnitude and characteristics of the estimated PA output could be quite different from those of the actual output sequence. To prevent this issue, a penalty function needs to be incorporated [5]. In this paper, the penalty function is considered to be the l_2-norm of the difference between the estimated PA output and the output of forward model, which is given as

$$L_2^{(i)} = \left\| \tilde{\boldsymbol{\psi}}(\mathbf{z}) \cdot \tilde{\boldsymbol{\alpha}}_{PA} - \tilde{\boldsymbol{\psi}}(\mathbf{z}) \cdot \tilde{\boldsymbol{\alpha}}^{(i)} \right\|_2^2. \tag{6}$$

Combine the two functions given in (5) and (6), and the coefficients vector $\tilde{\boldsymbol{\alpha}}^{(i+1)}$ for the $(i+1)$-th outer loop can be obtained by solving the following objective:

$$\begin{aligned} \tilde{\boldsymbol{\alpha}}^{(i+1)} &= \arg\min_{\tilde{\boldsymbol{\alpha}}_{PA}} L^{(i)} \\ &= \arg\min_{\tilde{\boldsymbol{\alpha}}_{PA}} \left(L_1^{(i)} + \lambda \cdot L_2^{(i)}\right) \\ &= \arg\min_{\tilde{\boldsymbol{\alpha}}_{PA}} \begin{pmatrix} \sum_{n=1}^{2K} f\left\{\left[\tilde{\boldsymbol{\psi}}_n(\mathbf{z}) \cdot \left(\tilde{\boldsymbol{\alpha}}_{PA} - \tilde{\boldsymbol{\alpha}}^{(i)}\right)\right] \cdot \tilde{\mathbf{s}}_n^{(i)}\right\} \\ +\lambda \left\| \tilde{\boldsymbol{\psi}}(\mathbf{z}) \cdot \tilde{\boldsymbol{\alpha}}_{PA} - \tilde{\boldsymbol{\psi}}(\mathbf{z}) \cdot \tilde{\boldsymbol{\alpha}}^{(i)} \right\|_2^2 \end{pmatrix}, \end{aligned} \tag{7}$$

where λ is a weight to give different priorities to the two loss functions during the iteration. When solving (7), the forward model $\tilde{\boldsymbol{\alpha}}^{(i)}$ is fixed, and $\tilde{\boldsymbol{\alpha}}_{PA}$, which represents the estimated PA model, is to be optimized. After the optimal model $\tilde{\boldsymbol{\alpha}}_{PA}$ is obtained by conducting some optimizing algorithms, it is assigned to $\tilde{\boldsymbol{\alpha}}^{(i+1)}$ for the next outer loop. The objective function $L^{(i)}$ has the advantages of both $L_1^{(i)}$ and $L_2^{(i)}$. On one hand, the signs of the estimated error signal can be close enough to the measured ones. On the other hand, the magnitude and characteristics of the estimated PA output waveform are consistent with those of the actual PA output.

D. DPD Model Extraction

The problem by now is solving (7) to obtain the forward model and then extract the DPD coefficients. Here we propose a joint optimization algorithm to identify the DPD model iteratively.

In the proposed algorithm, the first-order (Jacobian) and second-order (Hessian) derivatives of $L^{(i)}$, i.e. $\mathbf{J}^{(i)} = \partial L^{(i)} / \partial \tilde{\boldsymbol{\alpha}}_{PA}$ and $\mathbf{H}^{(i)} = \partial^2 L^{(i)} / \partial \tilde{\boldsymbol{\alpha}}_{PA}^2$, are calculated first. Then the coefficients of forward model for the $(i+1)$-th outer loop can be calculated as

$$\tilde{\boldsymbol{\alpha}}_{PA}^{(k+1)} = \tilde{\boldsymbol{\alpha}}_{PA}^{(k)} - \left(\mathbf{H}^{(i)}\right)^{-1} \mathbf{J}^{(i)}, \tag{8a}$$

$$\tilde{\boldsymbol{\alpha}}^{(i+1)} := \tilde{\boldsymbol{\alpha}}_{PA}^{(k+1)}, \tag{8b}$$

where $\tilde{\boldsymbol{\alpha}}_{PA}^{(0)} = [1, 0, \ldots, 0]^T$, and $\tilde{\boldsymbol{\alpha}}_{PA}^{(k)}$ is the estimated PA model after k iterations of (8a) and (8a) may only be calculated once in each forward modeling procedure.

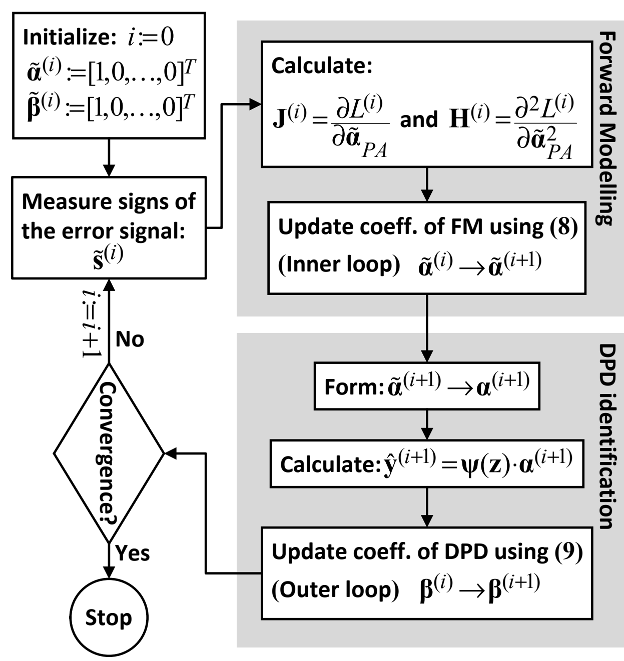

Fig. 2. Procedure of the proposed model extraction.

In the second step, i.e., updating the coefficients of DPD after the procedure of forward modeling, the output of PA, $\hat{\mathbf{y}}^{(i+1)}$, is estimated by conducting (1) using the newly identified forward model coefficients $\tilde{\boldsymbol{\alpha}}^{(i+1)}$. Since the difference between $\hat{\mathbf{y}}^{(i+1)}$ and the actual output $\mathbf{y}^{(i+1)}$ can be large in the first few iterations, calculating the Hessian matrix using indirect learning architecture may result in large error, thus the algorithm based on direct learning architecture is preferred here. The DPD coefficients $\boldsymbol{\beta}^{(i+1)}$ is updated as

$$\boldsymbol{\beta}^{(i+1)} = \boldsymbol{\beta}^{(i)} - \mu\left(\boldsymbol{\psi}^{H}(\mathbf{x})\cdot\boldsymbol{\psi}(\mathbf{x})\right)^{-1}\boldsymbol{\psi}^{H}(\mathbf{x})\left(\hat{\mathbf{y}}^{(i+1)} - \mathbf{x}\right), \quad (9)$$

where μ is the damping factor. The overall procedure of model identification, shown in Fig. 2, stops until the expected linearization performance is achieved.

III. Experimental Results

The test bed setup for experimental validation consists of a signal generator (PSG E8267D), a signal analyzer (PXA N9030A) and a 10 W GaN HEMT inverse class-F PA operated at 940 MHz. Both PSG and PXA are controlled by a PC running MATLAB through LAN ports. A single carrier LTE signal with 20 MHz bandwidth and 7.6 dB PAPR is used to excite the PA. Both the baseband I/Q data sent to the signal generator and the feedback signal are sampled at 92.16MSPS. A decomposed vector rotation (DVR) model [6] is used as the DPD model, with the partition number of 8 and memory depth equaling 2. Since the test bench was designed for the conventional DPD system and the extra analog path in the proposed method cannot be added in short time period, the high resolution output signals were captured in our test, and then the low resolution samples were obtained in MATLAB afterwards to emulate the scenario employing the 1-bit observation approach.

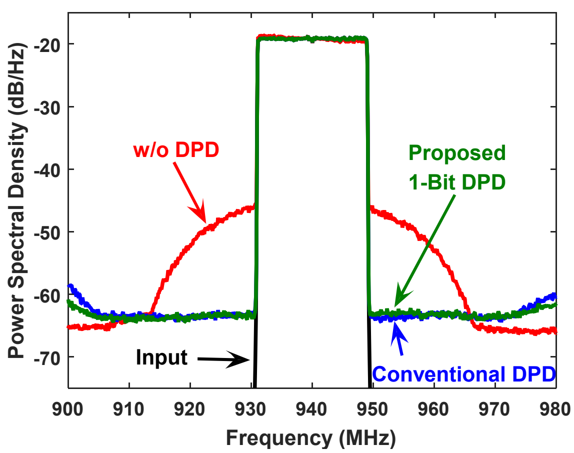

Fig. 3. Power spectral density comparison with different DPDs.

The power spectral density (PSD) comparison is demonstrated in Fig. 3. The conventional method is based on direct learning architecture and converges after 10 iterations. The proposed algorithm takes about 30 iterations for convergence, which is slower than that of the conventional method. However, as can be seen from the PSD plot, the proposed 1-bit DPD shows comparative linearization performance with the conventional method.

IV. Conclusion

This paper proposes a low-complexity 1-bit data acquisition approach for estimation of DPD coefficients. The proposed method uses only 1-bit resolution error signal to extract the forward model and then update the DPD coefficients using direct learning. This approach avoids the calculation of step size in the previous solution and guarantees the convergence of the algorithm. This novel 1-bit observation solution eases the requirement of ADC in DPD system, and thus reduces both the power consumption and the cost of the feedback path, compared to those methods using high resolution data. The proposed method has big potential in future broadband systems, such as 5G, since it can achieve ultra-high sampling speed by using comparators instead of high resolution ADCs.

Acknowledgement

This work is supported in part by the National Natural Science Foundation of China under Grant Number 61471333 and by the Science Foundation Ireland under Grant Numbers 13/RC/2077 and 12/IA/1267.

References

[1] L. Guan and A. Zhu, "Green communications: digital predistortion for wideband RF power amplifiers," *IEEE Microw. Mag.*, vol. 15, no. 7, pp. 84-99, Nov.-Dec. 2014.

[2] J. Wood, "Digital pre-distortion of RF power amplifiers: progress to date and future challenges," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Phoenix, AZ, 2015, pp. 1-3.

[3] F. M. Ghannouchi, O. Hammi and M. Helaoui, "Behavioral modeling and predistortion of wideband wireless transmitters," *John Wiley & Sons*, 2015.

[4] H. Wang, G. Li, C. Zhou, W. Tao, F. Liu and A. Zhu, "1-bit observation for direct learning based digital predistortion of RF power amplifiers," *IEEE Trans. Microw. Theory Techn.*, under review.

[5] J. Fang, Y. Shen, H. Li and Z. Ren, "Sparse signal recovery from one-bit quantized data: an iterative reweighted algorithm," *Signal Process.*, vol. 102, pp. 201-206, 2014.

[6] A. Zhu, "Decomposed vector rotation-based behavioral modeling for digital predistortion of RF power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 2, pp. 737-744, Feb. 2015.