Power Adaptive Digital Predistortion for Wideband RF Power Amplifiers with Dynamic Power Transmission

Yan Guo, *Student Member, IEEE*, Chao Yu, *Member, IEEE*, and Anding Zhu, *Senior Member IEEE*

***Abstract*— To reduce power consumption of wireless transmitters, the transmission power level of RF power amplifiers (PAs) may dynamically change according to real-time data traffic. This leads that the existing digital predistortion (DPD) techniques cannot be directly employed because they are mainly suitable for eliminating distortion induced by the PAs operated at a relatively stable condition, e.g., at a constant average power level. To resolve this problem, a power adaptive DPD is proposed in this paper. By accurately modeling the behavior change pattern of the PA with the input power adjustments and embedding it into the DPD model, the proposed DPD system is able to adjust its coefficients to adapt to the behavior variation of the PA induced by the power adjustments without real-time recalibration. A low complexity online coefficients updating method is also proposed to track the behavior change of the PA caused by other factors, such as bias shifting or temperature variation, during real-time operation. Measurements with a high power LDMOS Doherty PA have been used to validate the proposed approach. Results show that the proposed DPD and its coefficients updating approach can produce excellent performance with very low complexity compared to the conventional approaches.**

***Index Terms*—Average power, behavioral model, digital predistortion, dynamic power transmission, RF power amplifier.**

I. Introduction

To simultaneously achieve high efficiency and high linearity, digital predistortion (DPD) techniques have been widely employed to compensate for the nonlinear distortion induced by radio frequency (RF) power amplifiers (PAs) in high power cellular base stations [1]. In the past years, many advanced DPD models [2]-[4] and related model extraction structures [5] [6] have been proposed. The existing DPD systems, however, are mainly suitable for the scenarios where the PAs are operated at relatively stable conditions. For instance, the average input power level of the PA does not change dramatically within a short time period.

In the next generation of communication systems, the transmission power may frequently change with real-time traffic in order to minimize power consumption of the wireless system. For example, in the long term evolution advanced (LTE-A) system, resource blocks (RBs) of the data frame to be transmitted may be assigned according to the incoming data traffic. This can lead to the dynamic changes of the PA input power level in a short time period, e.g., milliseconds [7]. Since the PA is a nonlinear device, power changes at the input can significantly affect its nonlinear behavior [8] [9]. To properly compensate for the distortion caused by the power change in a short time period, almost real-time recalibration of the DPD must be conducted if the existing DPD techniques are employed, which is often not feasible in practice.

In [10], the average power within a long time window is included as long term memory terms in the model to eliminate the distortion induced by the PA in the transition process between two input power levels. In [11], the input power is employed as an index for the DPD look up table (LUT). Although these methods can eliminate distortion caused by power changes within a certain range, linearization of a PA under large dynamic power ranges, especially at steady-state stages, has not been addressed adequately. In [12], a power adaptive model was proposed to eliminate distortion at different power levels within a large power range without recalibration. This was achieved by embedding the PA behavior change with the input power adjustments into a conventional behavioral model structure. Due to the limited space, only the outline of the idea was presented in [12]. In this paper, a complete description of the approach, including problem statements, theoretical reasons, mathematical derivations and detailed descriptions for each single step of implementation, is provided to show how this approach can be applied in a real system when a PA is operated under various conditions. An efficient online coefficient updating technique is also presented to resolve the problem when the PA behavior changes in real-time operation, which further enhances the application of the proposed approach in practice. Extensive experimental results and comparisons with the technique proposed in [10] are also given.

The rest of the paper is organized as follows. After giving the problem statement in Section II, the details of the proposed DPD model together with the coefficients extraction and interpolation are presented in Section III. Online coefficients updating is presented in Section IV. The experimental validation results are given in Section V with a conclusion in Section VI.

II. Problem Statement

A. Digital Predistortion

Digital predistortion is a technique based on the principle

This work was supported in part by the Science Foundation Ireland under Grant Number 12/IA/1267.

Y. Guo and A. Zhu are with the School of Electrical and Electronic Engineering, University College Dublin, Dublin 4, Ireland (e-mail: yan.guo.1@ucdconnect.ie; anding.zhu@ucd.ie).

C. Yu is with the State Key Laboratory of Millimeter Waves, School of Information Science and Engineering, Southeast University, Nanjing, China (e-mail: chao.yu@seu.edu.cn).

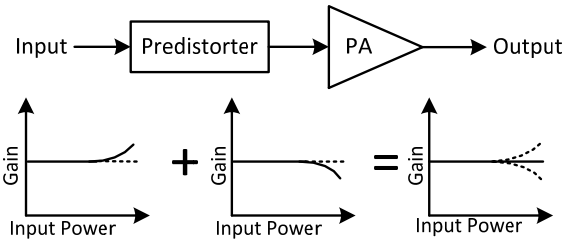


Fig. 1. Predistortion concept.

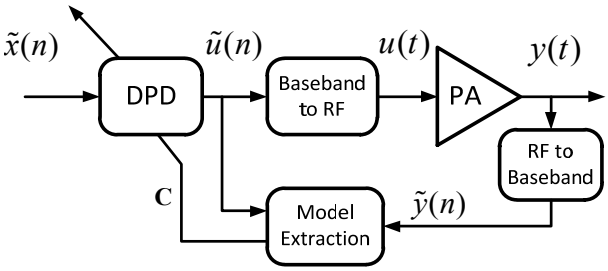


Fig. 2. Block diagram of a DPD system.

of nonlinear inversion. Let's take a simple example as shown in Fig. 1. The gain of a PA is ideally flat in the small signal region and tends to decrease when the input power level increases. To compensate for the distortion caused by the gain compression, a linearization block, called the predistorter, can be inserted into the signal path to "predistort" the signal before it enters the PA. If the transfer function of the predistorter is the exact inverse of the PA, a linear amplification can be achieved at the final PA output.

To make the predistortion system work, one of the most critical conditions is that the nonlinear transfer function of the predistorter must be correctly constructed. Although the input signal passes through the predistorter block before entering the PA, the input and the output of the PA must be captured and used to characterize the predistorter. In other words, the transfer function of the predistorter is decided by the PA characteristics, and it must be a one-to-one mapping. In a real application, the predistortion system is normally implemented in digital baseband and it normally includes two parts: a predistorter unit (DPD block) in the transmitter chain that predistorts the input signal in real-time and a model extraction unit in the feedback path that compares the captured PA output with the original input to extract the coefficients for the DPD model, as illustrated in Fig. 2. Since model extraction often takes some time to process, we must make sure that the behavior of the PA does not change during the model extraction. Otherwise the extracted DPD function will not be able to compensate for the nonlinear distortion induced by the PA and thus the linearization performance will deteriorate. This is not a vital issue in most existing communication systems, where the PAs are usually operated in a relatively stable condition and their behavior does not change significantly within a short time period.

B. Dynamic Change of Power

In the next generation of communication systems, e.g., LTE-A, the input power level of the PA may be adjusted according to real-time traffic. For instance, as shown in the Fig. 3 (a), the average power levels of the transmitted signal may change on a frame by frame basis. In this case, the PA is no longer operated under a time-invariant condition.

From the modeling point of view, we can divide the

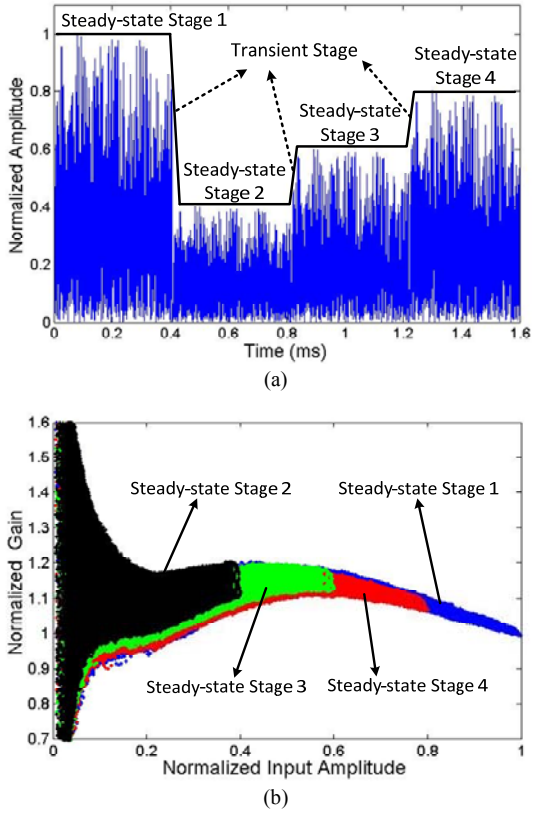


Fig. 3. A Doherty PA with dynamic power transmission: (a) truncated LTE-A signal sub-frames with power changes; (b) gain plots of the PA.

transmission into two stages: (1) the *transient* stage: the period when the input power of the PA jumps from one level to another; (2) the *steady-state* stage: the period when the average input power level is fixed. Due to long term memory effects, especially thermal effects that often occur in the PA, power changes in the transition will degrade the linearity performance if DPD does not take into account these effects properly [10], while at the steady-state stage, even if the average transmission power level is fixed, the nonlinear behaviors of the PA at different power levels are also different. As shown in Fig. 3 (b), there are clear differences between the gain plots for a Doherty PA operated at different power levels. This is because multiple amplifiers/transistors are used inside the Doherty PA and each amplifier exhibits different behavior with different power levels. The overall transfer functions that map the total input to output are not the same when the PA is excited with different power levels of signals. This leads that, when the input power level changes, a different DPD function must be employed or the coefficients must be recalibrated if the same DPD function is used. Since the input power level can change very quickly, e.g., with the units of sub-frame (1.0 millisecond) in LTE-A, the nonlinear behavior of the PA can change very frequently. As mentioned earlier, model extraction consumes time and it is a feedback process. It is very difficult to extract the coefficients and apply to DPD in a frame by frame basis, even if the state-of-the-art digital processor is employed. Another solution to this problem would be to store the coefficients or DPD functions in advance using a look-up-table (LUT), but it will require very large storage on-chip if the power changes are in a wide dynamic range.

C. Transient vs Steady-state Stages

The problems at both the transient and the steady-state

stages must be resolved for DPD when a PA is operated with dynamic power transmission. In [10], the average input power over a finite time window is built into the DPD model for modeling the long term memory effects during the transition, which can effectively prevent the performance degradation during this stage, but the problem at the steady-state stage has not been addressed adequately in the literature.

If we consider the linearity requirements of the communication standards, the problem at the steady-state stage needs to be resolved more urgently comparing to that at the transient stage. The reasons are as follows. In practical applications, the change of the PA input power level only occurs in the duration of time units of data sub-frames. In other words, the power only changes within a certain length of time intervals. Because of uncertainties in the transition, in communication standards, guard periods or preambles are usually inserted in the data frame. In the beginning or the end of the data frame, the transmission power level is usually set to very low and that part of the signal does not carry user data information. This means that, although we have to guarantee the distortion does not cause significant surge in the output spectrum, reasonably low distortion generated in this transient period would not have significant impact on the overall system performance. This is why communication standards usually do not put stringent requirements for this period [13]. On the other hand, significant time constant mainly occurs in GaN PAs. There have been considerable amount of efforts made to resolve this long term memory issue at the device level in the last couple of years and this issue has been gradually improved. The problem at the transient stage thus may be eased in the near future. Furthermore, the approach proposed in [10] has already largely addressed this issue.

At the steady-state stage, the user information data are carried by the transmitted signal. Any dynamic change of the PA behavior during this period directly affects the quality of the information data and thus stringent linearity requirements are often imposed during this period [13]. These dynamic changes are mainly resulted from long term memory effects in the PA and the impact of these memory effects becomes more severe in Doherty PAs. While the Doherty PAs will continue dominant the cellular base station market, this dynamic power transmission issue cannot be simply avoided. Although the method in [10] can also be employed at the steady-state stage to compensate for these nonlinearity changes, its performance is limited because only a first-order approximation is used in that model. More accurate models or approaches must be developed to resolve this issue in order to satisfy the stringent linearity requirements in future communication systems. In this paper, we mainly focus on finding a solution for characterizing and compensating the nonlinear behavior changes at the steady-state stage when a PA is operated with dynamic power transmission.

III. Power Adaptive DPD

To avoid the recalibration and reduce the system complexity, we propose a power adaptive approach as discussed below.

A. Model Structure of Power Adaptive DPD

In existing systems, most DPD models are simplified from the Volterra series. Without loss of generality, we take the 1st-order truncated dynamic deviation deduction (DDR) based Volterra model [4] as example. The DPD function can be expressed as:

$$\begin{aligned}\tilde{u}(n) = &\sum_{p=0}^{\frac{P-1}{2}}\sum_{m=0}^{M} c_{2p+1,1}(m)\left|\tilde{x}(n)\right|^{2p}\tilde{x}(n-m) \\ &+\sum_{p=1}^{\frac{P-1}{2}}\sum_{m=1}^{M} c_{2p+1,2}(m)\left|\tilde{x}(n)\right|^{2(p-1)}\tilde{x}^2(n)\tilde{x}^*(n-m),\end{aligned} \tag{1}$$

where $\tilde{x}(n)$ and $\tilde{u}(n)$ are the input and output of the DPD, respectively. $c_{2p+1,j}$ (j=1, 2) is the DPD complex coefficient. $(\cdot)^*$ represents the complex conjugate operation and $|\cdot|$ returns the magnitude. P is the order of nonlinearity and M represents the memory length. In a real system, the data are normally processed in blocks and thus we can express the DPD model in the following matrix form:

$$\begin{bmatrix}\tilde{u}(1)\\ \tilde{u}(2)\\ \vdots\\ \tilde{u}(N)\end{bmatrix} = \begin{bmatrix}\tilde{x}(1),\cdots, & \left|\tilde{x}(1)\right|^{P-3}\tilde{x}^2(1)\tilde{x}^*(1-M)\\ \tilde{x}(2),\cdots, & \left|\tilde{x}(2)\right|^{P-3}\tilde{x}^2(2)\tilde{x}^*(2-M)\\ \vdots\quad\ddots & \vdots\\ \tilde{x}(N),\cdots, & \left|\tilde{x}(N)\right|^{P-3}\tilde{x}^2(N)\tilde{x}^*(N-M)\end{bmatrix}\begin{bmatrix}c_{1,1}(0)\\ \vdots\\ c_{P,2}(M)\end{bmatrix}. \tag{2}$$

To simplify the expression, (2) can be rewritten in the following compact form:

$$\mathbf{U}_{N\times 1} = \mathbf{X}_{N\times L}\mathbf{C}_{L\times 1}, \tag{3}$$

where $\mathbf{U}$ is the output vector, $\mathbf{X}$ is the regression matrix containing all of the linear and nonlinear terms constructed from $\tilde{x}(n)$. The coefficients vector $\mathbf{C}$ includes all of the required coefficients $c_{2p+1,j}$ (j=1, 2). N and L are the length of input signal and the number of the coefficients, respectively.

As discussed earlier, the change of input power levels can lead to the PA behavior variation, which requires changes of the corresponding DPD. To facilitate the following derivation, we assume the input power level of the PA is arranged in a descend order and represented by PL_r, where r ($r = 1, \ldots, R$) is the index of the power level, and $PL_1 > PL_2 > \ldots > PL_R$. To identify the changes of the DPD at different power levels, we can use the DPD at the highest power level PL_1 as the reference. One way to represent the difference from the reference is to add a delta term. For instance, the output of the DPD at PL_r can be expressed as,

$$\mathbf{U}^{(r)} = \hat{\mathbf{U}} + \Delta\mathbf{U}^{(r)}, \tag{4}$$

where $\hat{\mathbf{U}}$ is the reference output, namely, the output at PL_1, while $\Delta\mathbf{U}^{(r)}$ is the delta value that can compensate for the behavior change of the PA at particular power levels. $\Delta\mathbf{U}^{(r)}$ can be constructed by using another nonlinear function. In this work, to avoid changing the model structure, we propose to use the same function that is used for the reference to model changes of the system. In other words, $\hat{\mathbf{U}}$ and $\Delta\mathbf{U}^{(r)}$ use the exact same model but the corresponding coefficient values are different. In this case, the different outputs at

different power levels can be represented by different sets of coefficients. For instance, the DPD coefficients at PL_r, which is labeled as $\mathbf{C}^{(r)}$, can be expressed as following:

$$\mathbf{C}^{(r)} = \hat{\mathbf{C}} + \Delta\mathbf{C}^{(r)}, \tag{5}$$

where $\hat{\mathbf{C}}$ is the coefficients vector at the reference power level, while $\Delta\mathbf{C}^{(r)}$ represents the difference from the reference. $\mathbf{C}^{(r)}$ and $\hat{\mathbf{C}}$ contains the exact same number of coefficient terms. If we extend this to multiple power levels, we can treat $\hat{\mathbf{C}}$ as the "static" coefficients vector that does not change during dynamic power transmission while $\Delta\mathbf{C}^{(r)}$ is the "dynamic" part that changes with the behavior variations of the PA. To identify the dynamic changes at different power levels, the coefficient values in $\Delta\mathbf{C}^{(r)}$ need to be linked to the signal changes, e.g., input power adjustments, to track the behavior variations of the PA. Ideally, $\Delta\mathbf{C}^{(r)}$ should be extracted at each different power level separately. This however, will significantly increase the system complexity.

From Fig. 3, we can see that the gain of the PA changes when the PA is operated at different power levels but the overall nonlinear behavior of the PA still follows a similar pattern. This means that if we model the PA behavior at one power level as the reference, the changes from the reference are relatively small. It indicates that the coefficient values in $\Delta\mathbf{C}^{(r)}$ are relatively small compared to those in $\hat{\mathbf{C}}$. Therefore, to reduce the system complexity, it is reasonable to extract a common set of dynamic coefficients, $\Delta\mathbf{C}$, and then use some tuning factors to scale the values to different power levels. This results that $\Delta\mathbf{C}^{(r)}$ can be expressed as:

$$\Delta\mathbf{C}^{(r)} = \begin{bmatrix} \Delta c_{1,1}^{(r)}(0) \\ \vdots \\ \Delta c_{1,1}^{(r)}(M) \\ \Delta c_{3,1}^{(r)}(0) \\ \vdots \\ \Delta c_{P,2}^{(r)}(M) \end{bmatrix} \approx \begin{bmatrix} \alpha_{1,1}^{(r)}(0)\times\Delta c_{1,1}(0) \\ \vdots \\ \alpha_{1,1}^{(r)}(M)\times\Delta c_{1,1}(M) \\ \alpha_{3,1}^{(r)}(0)\times\Delta c_{3,1}(0) \\ \vdots \\ \alpha_{P,2}^{(r)}(M)\times\Delta c_{P,2}(M) \end{bmatrix}, \tag{6}$$

where $\Delta c_{p,i}(m)$ (p=1, …, P, m=0, …, M and i=1, 2) is the common dynamic coefficient while $\alpha_{p,i}^{(r)}(m)$ is the scaling factor. Ideally, $\alpha_{p,i}^{(r)}(m)$ should be linked to each power level and extracted for each individual coefficient but that again will require significant efforts and increase the system complexity. As mentioned earlier, in a real system, the coefficients changes are relatively small and these changes can be summarized into two categories: average gain variation and nonlinear memory effects. In this work, we

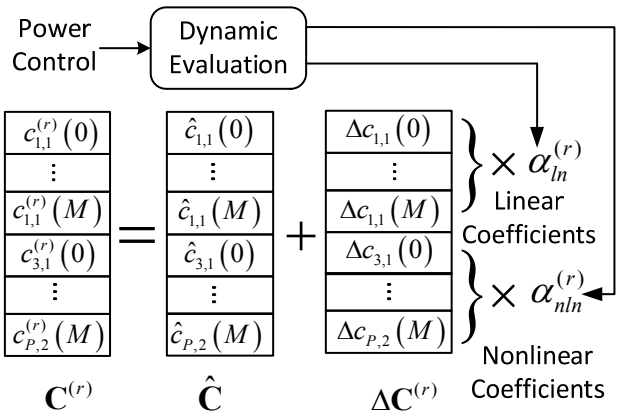


Fig. 4. Model coefficients structure.

introduce two scaling factors: linear scaling and nonlinear scaling that are directly related to the input power levels of the PA. The linear scaling factor $\alpha_{ln}^{(r)}$ can be employed to adjust all the linear coefficients. For the nonlinear coefficients, we assume that all the nonlinear coefficients can be tuned by the same nonlinear scaling factor $\alpha_{nln}^{(r)}$. It means that we make the following assumptions:

$$\begin{cases} \alpha_{1,1}^{(r)}(m) \approx \alpha_{ln}^{(r)}, \ m = 0,1,...,M \\ \alpha_{p,i}^{(r)}(m) \approx \alpha_{nln}^{(r)}, \ p = 3,...P, \ m = 0,1,...,M, \text{and } i = 1,2. \end{cases} \tag{7}$$

Thus, (6) can be simplified to

$$\Delta\mathbf{C}^{(r)} \approx \begin{bmatrix} \alpha_{ln}^{(r)}\times\Delta c_{1,1}(0) \\ \vdots \\ \alpha_{ln}^{(r)}\times\Delta c_{1,1}(M) \\ \hline \alpha_{nln}^{(r)}\times\Delta c_{3,1}(0) \\ \vdots \\ \alpha_{nln}^{(r)}\times\Delta c_{P,2}^{(r)}(M) \end{bmatrix} = \begin{bmatrix} \alpha_{ln}^{(r)}\times\Delta\mathbf{C}_{ln} \\ \hline \alpha_{nln}^{(r)}\times\Delta\mathbf{C}_{nln} \end{bmatrix}, \tag{8}$$

where $\Delta\mathbf{C}_{ln}$ and $\Delta\mathbf{C}_{nln}$ stand for the linear and nonlinear coefficients vector, respectively. Substituting (8) into (5), the coefficients equation can be rewritten as:

$$\mathbf{C}^{(r)} = \hat{\mathbf{C}} + \begin{bmatrix} \alpha_{ln}^{(r)}\times\Delta\mathbf{C}_{ln} \\ \alpha_{nln}^{(r)}\times\Delta\mathbf{C}_{nln} \end{bmatrix}. \tag{9}$$

The coefficients structure is illustrated in Fig. 4, where we can see that the final DPD coefficients include two parts: the

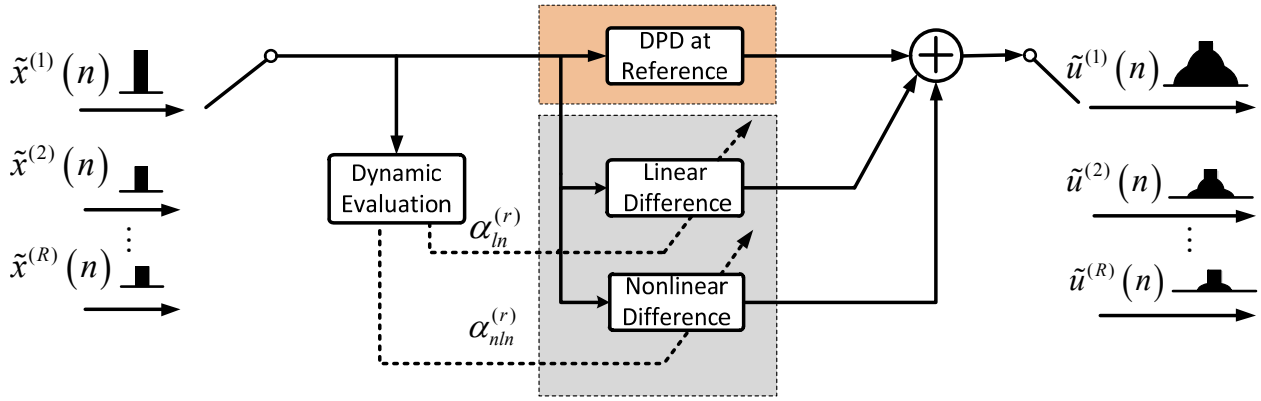


Fig. 5. Model operation principle.

reference $\hat{\mathbf{C}}$ and the dynamic part $\Delta\mathbf{C}^{(r)}$. The linear part of $\Delta\mathbf{C}^{(r)}$ is scaled by the linear scaling factor $\alpha_{ln}^{(r)}$ while the nonlinear part is scaled by the nonlinear scaling factor $\alpha_{nln}^{(r)}$. Combining (3) and (9), we can obtain the complete power adaptive DPD model as

$$\mathbf{U}^{(r)} = \mathbf{X}^{(r)}\hat{\mathbf{C}} + \mathbf{X}^{(r)}\begin{bmatrix}\alpha_{ln}^{(r)} \times \Delta\mathbf{C}_{ln} \\ \alpha_{nln}^{(r)} \times \Delta\mathbf{C}_{nln}\end{bmatrix}. \tag{10}$$

Since the model is linear-in-parameters, the scaling operation to the dynamic coefficients can be factored into the model terms, resulting:

$$\begin{aligned}\mathbf{U}^{(r)} &= \mathbf{X}^{(r)}\hat{\mathbf{C}} + \begin{bmatrix}\alpha_{ln}^{(r)}\mathbf{X}_{ln}^{(r)} & \alpha_{nln}^{(r)}\mathbf{X}_{nln}^{(r)}\end{bmatrix}\begin{bmatrix}\Delta\mathbf{C}_{ln} \\ \Delta\mathbf{C}_{nln}\end{bmatrix} \\ &= \mathbf{X}^{(r)}\hat{\mathbf{C}} + \begin{bmatrix}\alpha_{ln}^{(r)}\mathbf{X}_{ln}^{(r)} & \alpha_{nln}^{(r)}\mathbf{X}_{nln}^{(r)}\end{bmatrix}\Delta\mathbf{C},\end{aligned} \tag{11}$$

where $\mathbf{X}_{ln}^{(r)}$ and $\mathbf{X}_{nln}^{(r)}$ are the linear and nonlinear model terms corresponding to $\Delta\mathbf{C}_{ln}$ and $\Delta\mathbf{C}_{nln}$. In this model, the DPD function at PL_r is characterized as the additive combination of the reference, the linear and nonlinear differences from it, as illustrated in Fig. 5.

This model brings three advantages: (i) PA distortion at different power levels is characterized as a dynamic nonlinear function instead of a set of independent ones, which reduces the system complexity. (ii) The corresponding scaling factors $\alpha_{ln}^{(r)}$ and $\alpha_{nln}^{(r)}$ can be evaluated in advance, and only two sets of coefficients, $\hat{\mathbf{C}}$ and $\Delta\mathbf{C}$, need to be extracted in model extraction. (iii) Both the reference and the dynamic changing part are characterized by using the same model. There are no model structure changes in the system, which simplifies the system operation.

To accurately characterize the behavior change patterns of the PA, $\alpha_{ln}^{(r)}$ and $\alpha_{nln}^{(r)}$ must be chosen correctly. In this work, we propose the following approach to define them:

$$\begin{cases}\alpha_{ln}^{(r)} = \theta_r - \theta_1 \\ \alpha_{nln}^{(r)} = \varphi_r - \varphi_1\end{cases}, \tag{12}$$

where θ_r and φ_r are employed to evaluate the linearity and nonlinearity of the PA, and θ_1 and φ_1 are the references. To quantify the linear changes, the linear level evaluation θ_r can be obtained by using the best linear approximation [14]. In a system with memory, such as a PA, the transfer function of its best linear approximation in the time domain can be expressed by using a finite impulse response. This leads that θ_r is not a single value but a vector. Since the memory part of the coefficient values are relatively small compared to the static linear gain, to simplify the process, in this work, we use a time domain average gain to calculate θ_r instead of the best linear approximation. For the quantification of nonlinearity, a method called Measurements of Nonlinearity (MoN) [14]-[18] can be deployed. The principle of MoN is to quantify the nonlinear degree of a system by calculating the distance from its best linear approximation, therefore the numerical value of MoN only contains the "pure" nonlinearity. The linear approximation of the PA inverse transfer function is obtained through the calculation of θ_r while its nonlinearity evaluation based on MoN can be obtained by calculating the normalized l_2-based distance from its linear approximation, which can be expressed as:

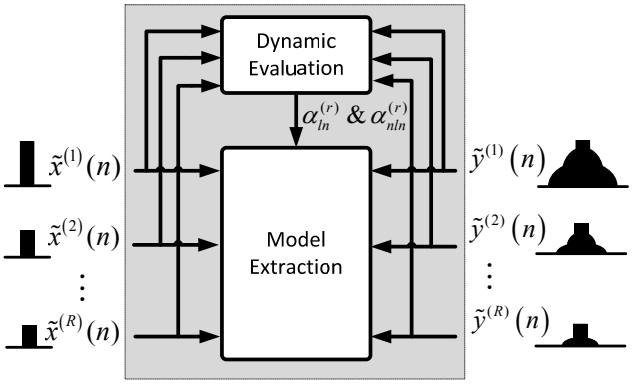


Fig. 6. Model extraction.

$$\varphi_r = \sqrt{\frac{\sum_{n=1}^{N}\left|\tilde{x}^{(r)}(n) - \theta_r\tilde{y}^{(r)}(n)\right|^2}{\sum_{n=1}^{N}\left|\theta_r\tilde{y}^{(r)}(n)\right|^2}}. \tag{13}$$

After obtaining θ_r and φ_r, the linear and nonlinear levels are quantified and the required scaling factors $\alpha_{ln}^{(r)}$ and $\alpha_{nln}^{(r)}$ can be generated from (12).

B. Model Extraction

To accurately extract the coefficients vector [$\hat{\mathbf{C}}$ $\Delta\mathbf{C}$], the input and output data measured from the PA at multiple power levels are required. To cover a large dynamic range, ideally, all possible power levels should be included. In practice, it is not possible because the complexity of model extraction will increase dramatically if a large number of datasets are included. In this work, we propose to select only limited sets of power levels for the model extraction and a coefficients interpolation approach will be introduced in the next sub-section to obtain the DPD coefficients for all the other required power levels.

The block diagram of model extraction is shown in Fig. 6. The input and output data at different power levels are captured first. The scaling factors $\alpha_{ln}^{(r)}$ and $\alpha_{nln}^{(r)}$ are then calculated independently for different power levels. Since the same model is used for all the power levels, we can simply put all data into one matrix. Using the p^{th}-order inverse or indirect leaning, we can swap the input and output data to construct the DPD model, namely, the output of the PA is used at the input of the DPD function while the original PA input is used as the expected output. The DPD matrix can then be constructed as,

$$\begin{bmatrix}\mathbf{X}^{(1)} \\ \mathbf{X}^{(2)} \\ \vdots \\ \mathbf{X}^{(R)}\end{bmatrix} = \begin{bmatrix}\mathbf{Y}^{(1)} \\ \mathbf{Y}^{(2)} \\ \vdots \\ \mathbf{Y}^{(R)}\end{bmatrix}\hat{\mathbf{C}} + \begin{bmatrix}\alpha_{ln}^{(1)}\mathbf{Y}_{ln}^{(1)} & \alpha_{nln}^{(1)}\mathbf{Y}_{nln}^{(1)} \\ \alpha_{ln}^{(2)}\mathbf{Y}_{ln}^{(2)} & \alpha_{nln}^{(2)}\mathbf{Y}_{nln}^{(2)} \\ & \vdots \\ \alpha_{ln}^{(R)}\mathbf{Y}_{ln}^{(R)} & \alpha_{nln}^{(R)}\mathbf{Y}_{nln}^{(R)}\end{bmatrix}\Delta\mathbf{C}, \tag{14}$$

where $\mathbf{X}^{(r)}$ is the original PA input while the regression matrix $\mathbf{Y}^{(r)}$ can be formed by using the output data, which include two parts: (i) the normal terms as those in the conventional DPD model; (ii) the same set of terms, but the linear terms are scaled by $\alpha_{ln}^{(r)}$ and the nonlinear terms are scaled by $\alpha_{nln}^{(r)}$, where $\mathbf{Y}_{ln}^{(r)}$ and $\mathbf{Y}_{nln}^{(r)}$ are the linear and nonlinear model terms in $\mathbf{Y}^{(r)}$, respectively.

Since the output is in linear relationship with the

coefficients, the least squares (LS) algorithm can be employed:

$$\begin{bmatrix}\hat{\mathbf{C}}\\ \Delta\mathbf{C}\end{bmatrix}=\left[\mathbf{Y}^{\mathrm{H}}\mathbf{Y}\right]^{-1}\mathbf{Y}^{\mathrm{H}}\begin{bmatrix}\mathbf{X}^{(1)}\\ \mathbf{X}^{(2)}\\ \vdots\\ \mathbf{X}^{(R)}\end{bmatrix}, \tag{15}$$

where

$$\mathbf{Y}=\begin{bmatrix}\mathbf{Y}^{(1)} & \alpha_{ln}^{(1)}\mathbf{Y}_{ln}^{(1)} & \alpha_{nln}^{(1)}\mathbf{Y}_{nln}^{(1)}\\ \mathbf{Y}^{(2)} & \alpha_{ln}^{(2)}\mathbf{Y}_{ln}^{(2)} & \alpha_{nln}^{(2)}\mathbf{Y}_{nln}^{(2)}\\ \vdots & \vdots & \vdots\\ \mathbf{Y}^{(R)} & \alpha_{ln}^{(R)}\mathbf{Y}_{ln}^{(R)} & \alpha_{nln}^{(R)}\mathbf{Y}_{nln}^{(R)}\end{bmatrix}. \tag{16}$$

Since the regression matrix includes all data at different driven levels, the condition number of the regression matrix can be large which may cause difficulties in the matrix inversion. To avoid this issue, the 1-bit ridge regression method in [19] can be employed. Usually, the model extraction process needs to be conducted in several iterations to achieve the best linearization performance.

C. Coefficients Interpolation

Since the model extraction can only cover a limited number of operating power levels, an interpolation approach is required to obtain the DPD coefficients for the power levels that are not included in the model extraction process. To accomplish this task, we set PL_1 and PL_R to the maximum and the minimum power level, respectively, and divide the entire input power range uniformly as following:

$$\left|PL_{r-1}-PL_r\right|=\frac{\left|PL_1-PL_R\right|}{R-1}. \tag{17}$$

For the power levels which are selected in the model extraction process, the corresponding scaling factors can be calculated and the coefficients at these power levels can be made available immediately. For instance, the DPD coefficients for PL_r can be obtained by using (9). The task now is to estimate the coefficients vector $\mathbf{C}^{(r,k)}$ for the power level $PL_{r,k}$, which is not in the model extraction. Let's assume $PL_{r,k}$ is located between PL_r and PL_{r-1} ($PL_r<PL_{r,k}<PL_{r-1}$) and the PA behavior change is consistent with the input power levels. Since $\mathbf{C}^{(r)}$ is constructed from linear and nonlinear tuning factors as in (9), the interpolation formula could be conducted by using the same tuning factors with a weighting

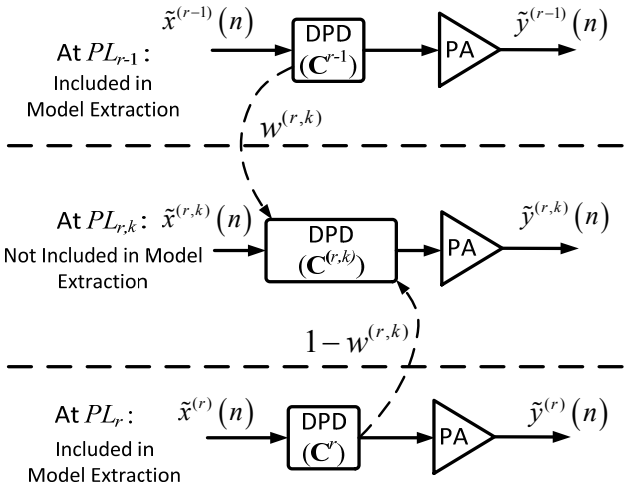


Fig. 7. DPD coefficients interpolation.

function, which is expressed as

$$\begin{cases}\alpha_{ln}^{(r,k)}=w^{(r,k)}\alpha_{ln}^{(r-1)}+\left(1-w^{(r,k)}\right)\alpha_{ln}^{(r)}\\ \alpha_{nln}^{(r,k)}=w^{(r,k)}\alpha_{nln}^{(r-1)}+\left(1-w^{(r,k)}\right)\alpha_{nln}^{(r)}\end{cases},r=2,3,...R \tag{18}$$

In practical systems, the changes of the PA behavior often follow a nonlinear pattern. For instance, higher input power levels may lead to more severe nonlinearities in a Class-AB or a Doherty PA. Thus, in the high power region, the coefficients values at the middle power level should be made closer to the ones at the higher power level rather than that at the lower side. While, in the low power region, PA

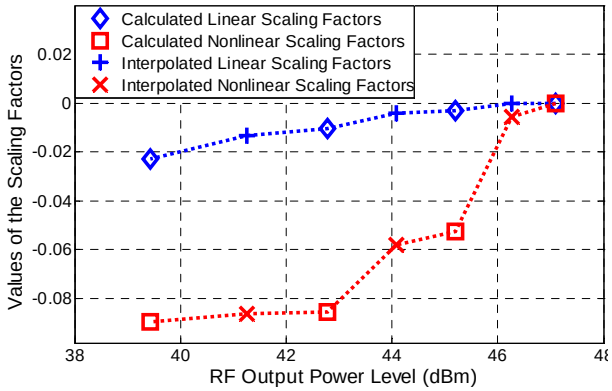


Fig. 8. Example values of scaling factors.

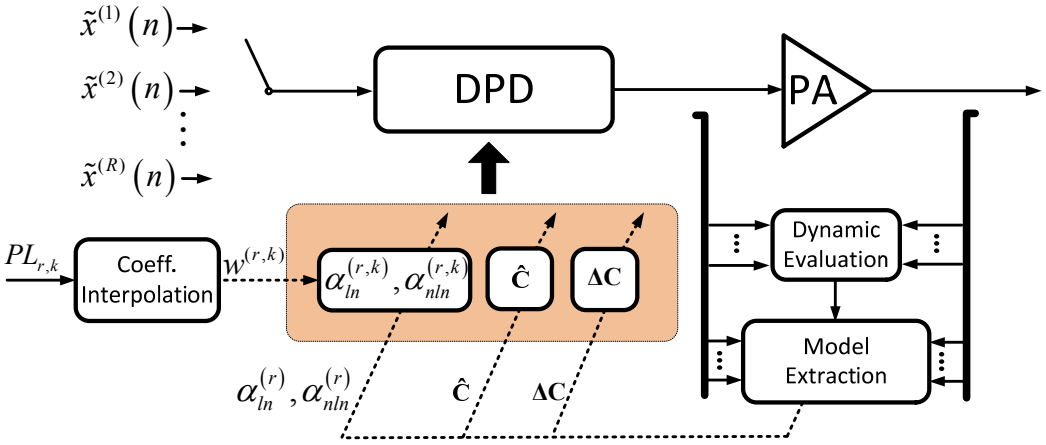


Fig. 9. Block diagram of the complete power adaptive DPD system.

nonlinearity becomes milder, the coefficients at the middle power level can be interpolated more evenly. The weighting factor $w^{(r,k)}$ is thus designed to reflect this pattern. Here we take the following function as an example:

$$w^{(r,k)}=\left(1-\left(\frac{PL_{r,k}-PL_{r-1}}{PL_{r-1}}\right)^2\right)^2 . \quad (19)$$

The interpolation process is illustrated in Fig. 7. The weighting factor $w^{(r,k)}$ is calculated first, the tuning factors can be then interpolated and finally the corresponding coefficients can be obtained. Fig. 8 shows an example of the coefficients scaling factor values used for a Doherty PA excited with a 20 MHz LTE signal. The power levels of the baseband signal were normalized with the peak values from 1.0 to 0.4 with 0.1 as the step size, leading to the corresponding changes of the PA output power from 47.1 dBm to 39.2 dBm. The scaling factors at the power levels of 1.0, 0.8, 0.6 and 0.4 were calculated from the measured data while the ones at other power levels were interpolated with the proposed approach. From the plot, we can clearly see that the interpolated scaling factors at high power levels, e.g., at 46.3 dBm, are closer to the ones at 47.1 dBm while at the lower power levels, e.g., at 41.3 dBm, the scaling factors are interpolated more evenly. One may note that (19) is only one example. Different types of nonlinear functions can be used and they should properly reflect the characteristics of the PA in the system.

D. The Complete System Structure

From the analysis above, the structure of the complete power adaptive DPD system can be constructed as the block diagram shown in Fig. 9. Before the system operation, the input and output of the PA operated at selected power levels are captured and used to calculate the corresponding tuning factors and extract the model coefficients [$\hat{\mathbf{C}}$ $\Delta\mathbf{C}$]. During the online system operation, the power level of the input signal, e.g., $PL_{r,k}$, is first employed to calculate the weighting factor according to (19). Then the required tuning factor $\alpha_{ln}^{(r,k)}$ and $\alpha_{nln}^{(r,k)}$ are generated and fed into the DPD block to multiply with the elements of the dynamic coefficients $\Delta\mathbf{C}$. The adjusted $\Delta\mathbf{C}$ is then combined with $\hat{\mathbf{C}}$ to generate the coefficients vector $\mathbf{C}^{(r,k)}$ to be used in the DPD block.

In this new system, although the number of coefficients that needs to be extracted is doubled as that of the conventional DPD, the two sets of DPD coefficients can be combined together after the model extraction to form the DPD vector as shown in (9). Therefore, there is no change required for implementing the DPD block. In other words, the same hardware resource is used and the only difference is that the coefficients values are updated according to the input power levels on operation. This leads to a significant reduction on system implementation complexity.

IV. Online Coefficients Updating

In the previous section, we mainly focus on the PA behavior variations caused by the input power changes. In those cases, we assume the PA is a time invariant system apart from the excitation signals. In other words, the PA behavior keeps the same if it is operated at the same input power level. Therefore once the coefficients are extracted, there is no further update required. In a real operation, the PA behavior may change due to other factors, e.g., temperature variation, bias drifting or aging. Unlike the input power adjustments, this type of changes usually cannot be predicted in advance. To maintain the linearization level, the DPD coefficients must be occasionally updated online.

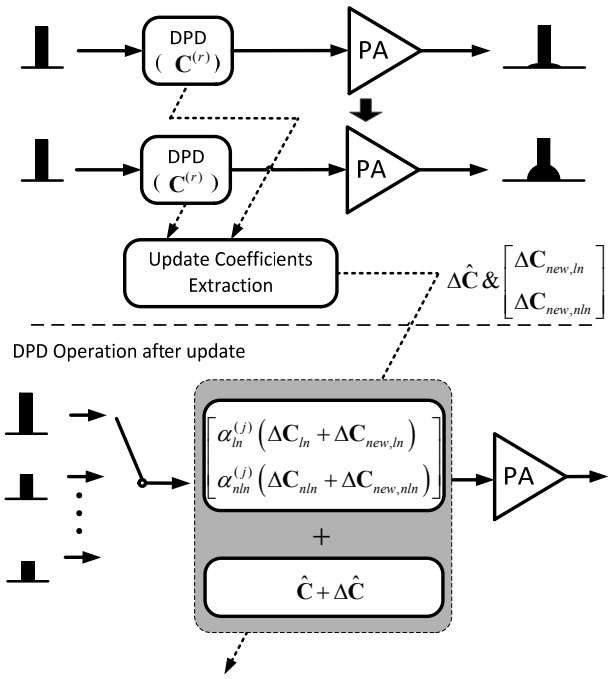


Fig. 10. Online coefficients updating.

In conventional DPD systems, recalibration can be conducted by re-running the model extraction based on the current transmitted signals. In the power adaptive system, however, re-running the model extraction requires that the signals at certain fixed power levels must be captured in order to accurately extract the coefficients, as described in Section III. This is usually not feasible during the real-time transmission since it may take a long period to gather all the signals at the required power levels. The transmission quality during the waiting period may be thus significantly affected. To mitigate this problem, we propose to characterize the PA behavior change from the signal gathered at the current transmission level and then apply it to other power levels so that the performance deterioration during the waiting period can be avoided or reduced.

Let's assume the PA transfer function is changed at the power level PL_r during the real-time operation. To recalibrate the DPD coefficients at this power level, the data sets can be simply captured to construct the following PA inverse model:

$$\mathbf{X}_{new}^{(r)}=\mathbf{Y}_{new}^{(r)}\mathbf{C}_{new}^{(r)}, \quad (20)$$

where $\mathbf{X}_{new}^{(r)}$ and $\mathbf{Y}_{new}^{(r)}$ are the new PA input vector and the model regression matrix, respectively. A normal LS algorithm can be operated to obtain the new coefficients at PL_r, labeled as $\mathbf{C}_{new}^{(r)}$. By comparing $\mathbf{C}_{new}^{(r)}$ and the original coefficients, the PA behavior change at this power level, can then be obtained by

$$\Delta\mathbf{C}_{new}^{(r)}=\mathbf{C}_{new}^{(r)}-\mathbf{C}^{(r)} . \quad (21)$$

In the power adaptive DPD, the PA behavior variation is characterized by multiplying the dynamic coefficients $\Delta\mathbf{C}$

with the scaling factors $[\alpha_{ln}^{(r)}\ \alpha_{nln}^{(r)}]$. Theoretically speaking, if the PA behavior is affected by other factors, the $\hat{\mathbf{C}}$, $\Delta\mathbf{C}$ and $[\alpha_{ln}^{(r)}\ \alpha_{nln}^{(r)}]$ should all be different from the original values. However, in practical systems, the PA change level during the online calibration is relatively small compared to the original nonlinearity. Hence, we can assume the linear and nonlinear change patterns $[\alpha_{ln}^{(r)}\ \alpha_{nln}^{(r)}]$ remain the same and the variation can be characterized by changing the coefficients only. $\Delta\mathbf{C}_{new}^{(r)}$ can thus be decomposed as

$$\Delta\mathbf{C}_{new}^{(r)} = \Delta\hat{\mathbf{C}} + \begin{bmatrix} \alpha_{ln}^{(r)}\Delta\mathbf{C}_{new,ln} \\ \alpha_{nln}^{(r)}\Delta\mathbf{C}_{new,nln} \end{bmatrix}, \tag{22}$$

where $\Delta\hat{\mathbf{C}}$ stands for the change of the reference coefficients $\hat{\mathbf{C}}$. $\Delta\mathbf{C}_{new,ln}$ and $\Delta\mathbf{C}_{new,nln}$ represent the change of the linear and nonlinear dynamic coefficients, respectively. To obtain the values for those decomposed coefficients, we can use the modified model extraction. First, according to (20) and (21), the change of the PA inverse model can be expressed as:

$$\Delta\mathbf{X}_{new}^{(r)} = \mathbf{X}_{new}^{(r)} - \mathbf{Y}_{new}^{(r)}\mathbf{C}^{(r)} = \mathbf{Y}_{new}^{(r)}\Delta\mathbf{C}_{new}^{(r)}. \tag{23}$$

Then, by substituting (22) to (23), we have:

$$\Delta\mathbf{X}_{new}^{(r)} = \mathbf{Y}_{new}^{(r)}\Delta\hat{\mathbf{C}} + \begin{bmatrix} \alpha_{ln}^{(r)}\mathbf{Y}_{ln,new}^{(r)} & \alpha_{nln}^{(r)}\mathbf{Y}_{nln,new}^{(r)} \end{bmatrix}\begin{bmatrix} \Delta\mathbf{C}_{new,ln} \\ \Delta\mathbf{C}_{new,nln} \end{bmatrix}. \tag{24}$$

The equation above is linear-in-parameters and the LS method can then be employed to extract $\Delta\hat{\mathbf{C}}$, $\Delta\mathbf{C}_{new,ln}$ and $\Delta\mathbf{C}_{new,nln}$.

Since the scaling factors $[\alpha_{ln}^{(r)}\ \alpha_{nln}^{(r)}]$ are assumed the same, $\Delta\hat{\mathbf{C}}$, $\Delta\mathbf{C}_{new,ln}$ and $\Delta\mathbf{C}_{new,nln}$ can be immediately applied to the other power levels, PL_j, to conduct the coefficients updating, i.e.,

$$\mathbf{C}_{new}^{(j)} = \hat{\mathbf{C}} + \Delta\hat{\mathbf{C}} + \begin{bmatrix} \alpha_{ln}^{(j)}\Delta\mathbf{C}_{ln} + \alpha_{ln}^{(j)}\Delta\mathbf{C}_{new,ln} \\ \alpha_{nln}^{(j)}\Delta\mathbf{C}_{nln} + \alpha_{nln}^{(j)}\Delta\mathbf{C}_{new,nln} \end{bmatrix}, \quad j = 1,2,...,R. \tag{25}$$

The block diagram of the coefficients updating process is shown in Fig. 10. Since there is only one set of data required during this model extraction, the coefficients updating process can be very fast. This is very important because there is often no waiting time allowed during the real-time operation and updating the coefficients in time is essential in order to avoid distortion surge and ensure smooth transition.

In real-time operation, it is not necessary to wait for signals at the highest power level, but to achieve good performance, it prefers to capture the data at relatively higher power levels for model extraction because generally nonlinearity changes at those power levels are more severe in the PA. Overall, it is worth mentioning that this online updating is only an interim approach for avoiding sudden surge of distortion during the transmission when the behavior of the PA changes. In order to obtain the optimized linearization level, the model should be fully re-calibrated by using the approach proposed in Section III when the data sets of all required power levels are available.

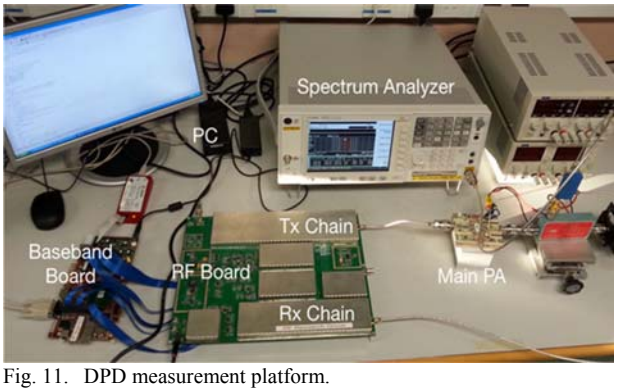


Fig. 11. DPD measurement platform.

V. Experimental Measurements

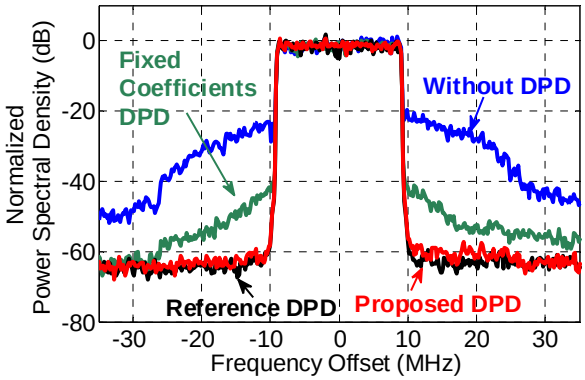
To validate the proposed DPD model and its coefficients updating approach, we tested a high power LDMOS Doherty PA operated at 2.14 GHz based on the test bench designed in [20], as shown in Fig. 11. On the transmitter side, the baseband signal is generated in a PC and sent to a baseband board for digital signal processing. An RF board then converts the baseband signal into the analog domain and up-converts it to the operated RF frequency. The signal is then sent to a driver and finally fed into the main PA. On the receiver side, the feedback loop is used to capture the PA output signal and down-convert it to baseband and then send it to PC for model extraction and DPD generation.

A. Power Adaptive DPD Validation

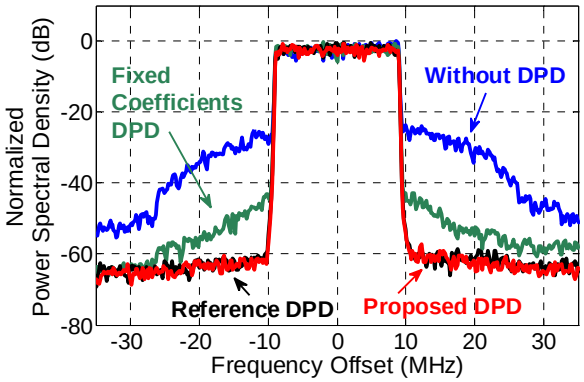
In this part, the performance of the proposed DPD model is tested. To reflect the power adjustments during the dynamic power transmission, a normalized scaling factor from 1 to 0.4 with step size 0.1 was applied to the baseband signal before transmission, corresponding to the changes of the average RF output power from 47.1 dBm to 39.2 dBm. The datasets at scaled levels 1, 0.8, 0.6 and 0.4 were captured for the model extraction, while data at all other power levels were employed for the DPD validation. All the datasets for model extraction are normalized by the highest amplitude at 47.1 dBm. During the test, 5,000 signal points were used for model extraction while the entire 32,000 samples were tested for cross validation at each power level.

The tests were conducted in four scenarios: 1) without DPD; 2) the fixed coefficients DPD that linearizes the PA at different driven levels by using a common set of coefficients which was extracted from putting signals at different levels into one matrix; 3) the proposed power adaptive DPD; and 4) the reference DPD that can be considered as an ideal case, which employed a set of coefficients extracted at the same power level as that of the validation. The 2nd order DDR model [21] combined with piecewise decomposed method [22] and power adaptive approach was employed in this work. The nonlinear order and the memory length were set to [7, 7] and [3, 3], respectively. The piecewise decomposition threshold was set to 0.5.

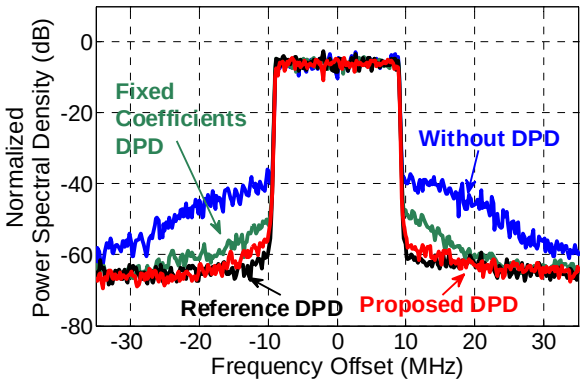
The normalized power spectral density (PSD) of the PA output at four power levels are plotted in Fig. 12. Two power levels are selected from the ones for model extraction, i.e., 47.1 dBm and 42.9 dBm, while another two power levels are obtained from the interpolation for validation purpose, i.e., 46.3 dBm and 41.5 dBm. The normalized mean square errors (NMSE) values and adjacent channel power ratio (ACPR) values are given in Fig. 13 and Fig. 14 for all the power levels, respectively. Due to the inherent nonlinear nature, the PA under test is sensitive to the variation of its input power level.



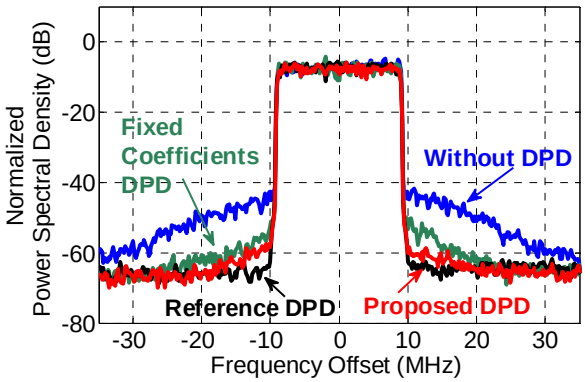
(a)



(b)



(c)



(d)

Fig. 12. The normalized PSD at different RF output power levels: (a). 47.1 dBm (in model extraction), (b) 46.3 dBm (with coefficients interpolation), (c) 42.9 dBm (in model extraction), (d) 41.5 dBm (with coefficients interpolation).

Therefore, if we compensate the distortion induced by the PA at different driven levels by using only one set of DPD coefficients, the performance will deteriorate and

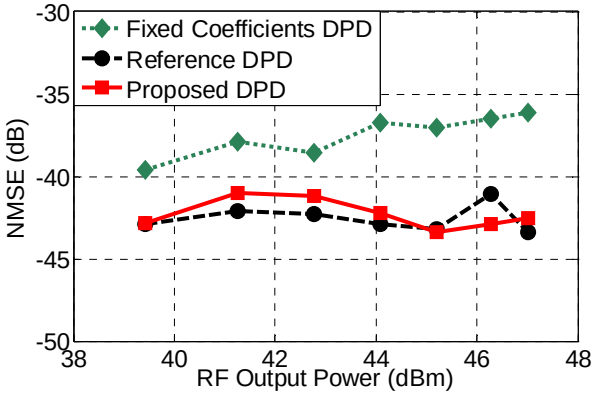
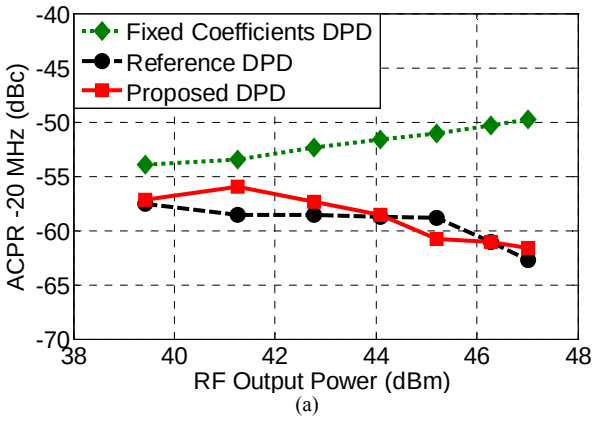
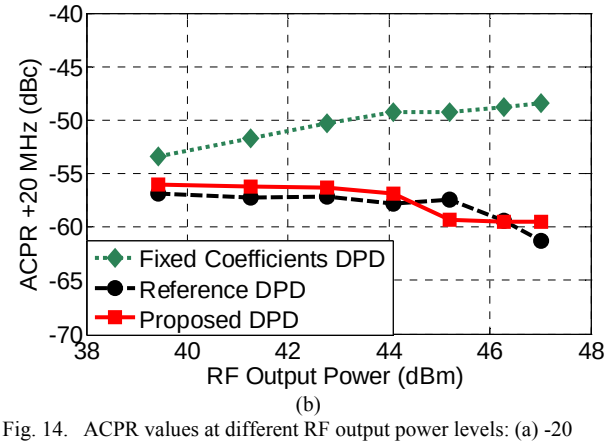


Fig. 13. NMSE values at different RF output power levels.



(a)



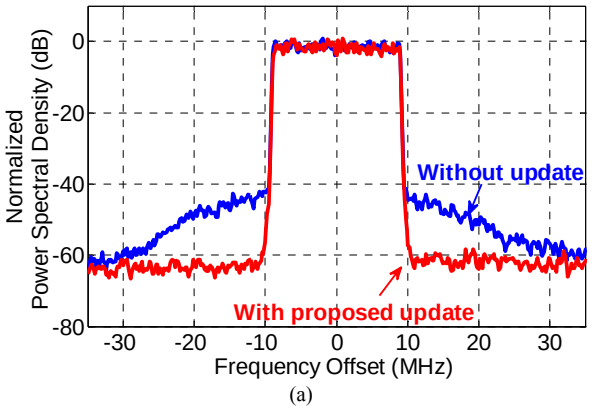
(b)

Fig. 14. ACPR values at different RF output power levels: (a) -20 MHz; (b) +20 MHz.

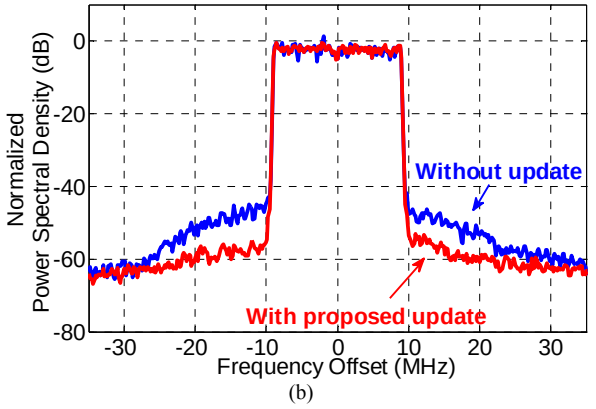
TABLE I
COMPARISONS OF THE COEFFICIENT NUMBERS

Scenario	No. of Coeff. Model Extraction	No. of Coeff. DPD Block
Reference DPD	602=86×7	86
Proposed DPD	172=86×2	86
Fixed Coefficients DPD	86	86

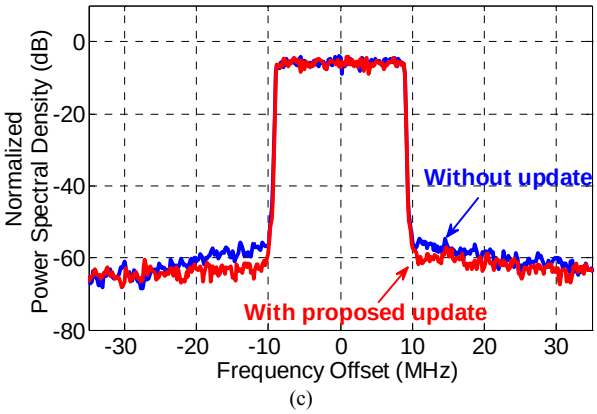
considerable nonlinear residuals will remain, especially at the high driven levels. This can be easily identified by checking the NMSE and ACPR values of the fixed coefficients DPD results in Fig. 13 and Fig. 14, respectively. At high power levels, e.g., 47.1 dBm, the ACPR is only about -48 dBc. However, by employing the power adaptive DPD, the distortion of the PA at each power level can be effectively compensated, achieving almost the same ACPR and NMSE performance as that achieved by employing the reference DPD. This concludes that the power adaptive DPD is capable of accurately characterizing the behavior change pattern of



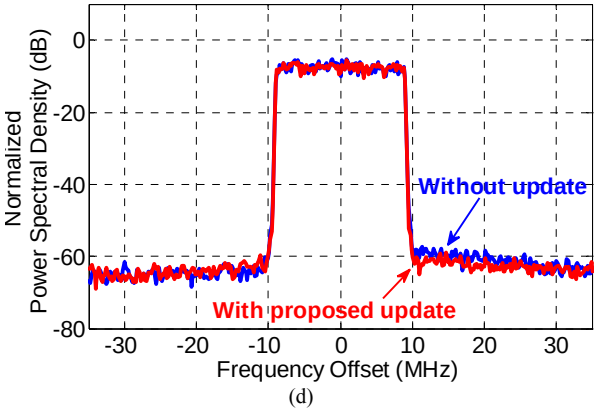
(a)



(b)



(c)



(d)

Fig. 15. The normalized PSD with and without coefficients update: (a). 47.1 dBm, (b) 46.3 dBm, (c) 42.9 dBm, (d) 41.5 dBm,

the PA with the input power adjustments.

Since the coefficients of the conventional DPD are obtained by the linear combination of $[\hat{\mathbf{C}}\ \Delta\mathbf{C}]$ and the scaling factors $\alpha_{ln}^{(r,k)}$ and $\alpha_{nln}^{(r,k)}$, the total number of coefficients does

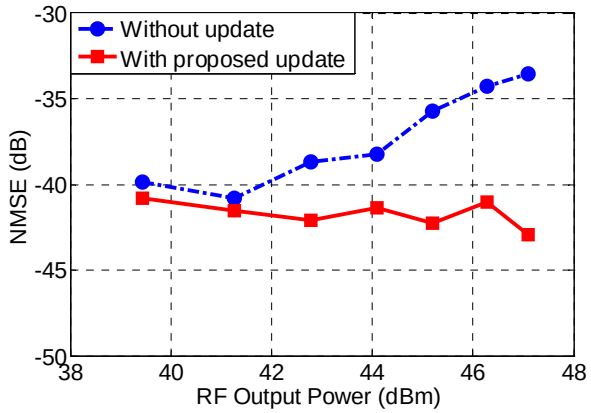
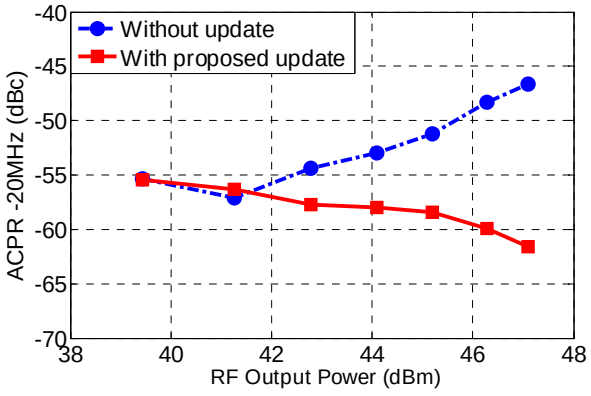
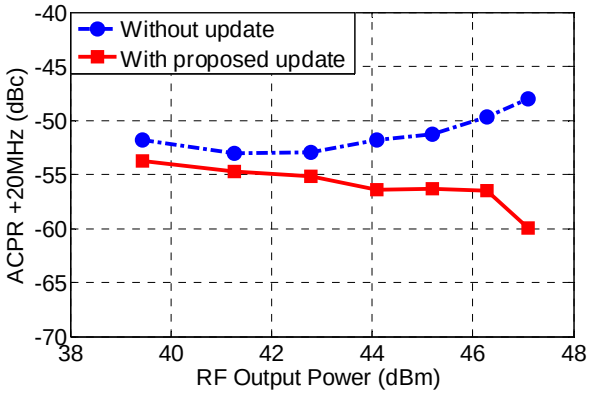


Fig. 16. NMSE values with and without update.



(a)



(b)

Fig. 17. ACPR values with and without update at different RF output power levels: (a) -20 MHz; (b) +20 MHz

not increase with the number of the operation power levels which can be seen from TABLE I. In the reference DPD, each power level is treated independently, so that the number of coefficients is therefore increased to 602. While in the power adaptive DPD, the total number of coefficients is only 172, yielding significantly reduction of the system complexity. Meanwhile, because there is no new DPD model term added in the power adaptive model, the DPD signal generation complexity remains the same. In other words, the same number of coefficients, e.g., 86, is used in the predistortion block as that is used in the conventional DPD.

B. Coefficients Updating Validation

In this part, the performance of the coefficients updating approach of the power adaptive DPD was tested. To illustrate the PA behavior changes caused by other factors rather than power level changes during transmission, we shifted the drain

supply bias from 28 V to 26 V which causes the change of the PA behavior. The dataset at the highest power level was employed to extract the updated coefficients. The PSDs with and without update are shown in Fig. 15. The NMSE and ACPR values are presented in Fig. 16 and Fig. 17, respectively. The power levels used for these figures are the same as the ones in Part A.

Due to the change of bias, the transfer function of the PA is seriously affected, leading to the deterioration of the DPD performance which can be seen from the spectral regrowth at the PSDs and the increases of NMSE values. At higher driven levels, the PA transfer functions are more sensitive to the bias change, so the DPD performances drop more significantly. With the employment of the proposed approach, the regrowth of the nonlinear distortion can be effectively compensated and the ACPR and NMSE values can be significantly improved. At lower driven levels, e.g., 41.5 dBm, the PA is operated at almost linear region therefore it is not sensitive to the change of bias. Under this condition, the DPD performance with and without coefficients update both maintains at fine levels.

C. Model Comparison

As mentioned before, the proposed approach is developed to eliminate different nonlinearity for each steady-state stage so that the accuracy of the transmitted information can be guaranteed. In [10], a long term memory DPD is designed in order to compensate the nonlinear effects in the transient stage. This method can also be used at the steady-state stage. In order to compare this approach with the proposed one, a 10 MHz LTE signal with two different power levels, as shown in Fig. 18 (a), was used to conduct the DPD measurements. The output power level at high and low power states were 46 dBm and 38 dBm, respectively. This test scenario is very similar to the one in [10].

The basic DPD model employed here was the 2nd order DDR model (P=9, M=3). Three test cases were considered for this comparison: 1) the fixed coefficients method which used the coefficients extracted at high power level to linearize the entire signal sequence; 2) the method in [10], the coefficients were extracted from 30,000 points with a slide window length of 10,000 sampling points; 3) the proposed method, the coefficients were extracted from 15,000 points (7,500 points at the high level and 7,500 points at the low level). The final DPD measurements results are shown in Fig. 18.

For comparison, two slices of the error signals between the linearized output and the original baseband signals for the steady-state stage and the transient stage are provided in Fig. 18 (a), where we can see that the model in [10] produces smaller errors in transition while larger errors at the steady-state stage compared to that with the proposed approach. To fully quantify the accuracy, the entire input signal sequence is decomposed into 99 sub-sections with 3,000 points each and the NMSE for each section is calculated and given in Fig. 18 (b). Due to the PA behavior changes with the input power levels, using the fixed coefficients extracted at the high power level at the low power level cannot achieve fine linearization performance, which leads high NMSE values at the low power level as shown in the plot. The method in [10] is able to effectively

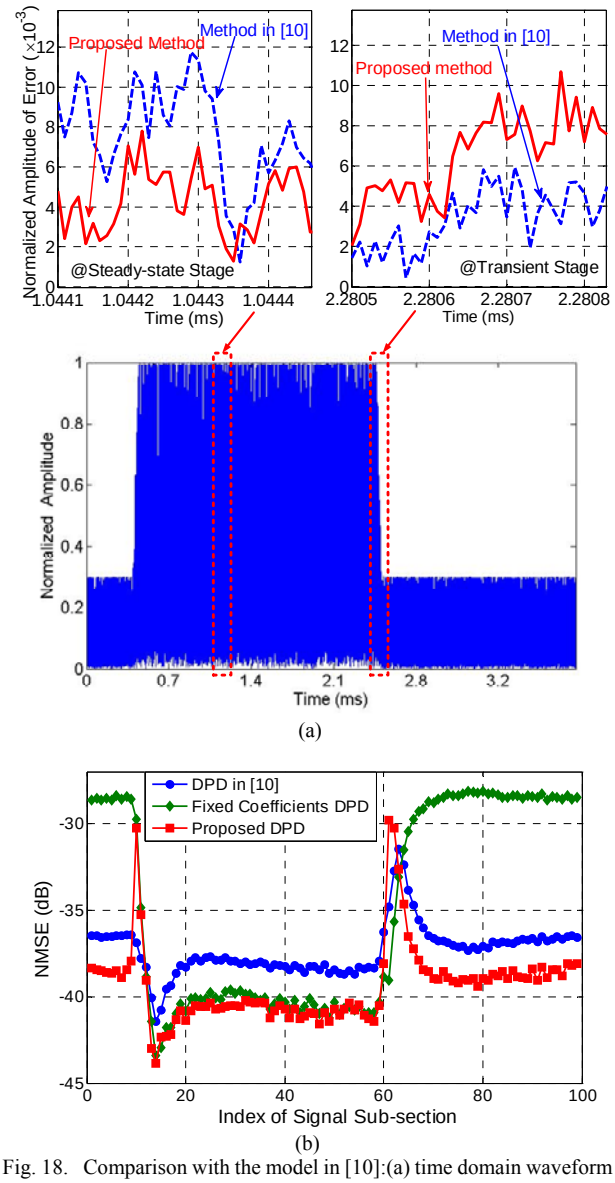


Fig. 18. Comparison with the model in [10]:(a) time domain waveform and error signal amplitudes; (b) NMSE performance

compensate the distortion during the transient stage but a certain level of distortion still remains at the steady-state stage. With the proposed method, we can see that the performance can be significantly improved at the steady-state stage, with more than 3 dB NMSE improvements over the method in [10] during both the high and low power levels and the performance is close to the ideal case, as shown in Fig. 18 (b). However, at the transient stage, there is a spur in the NMSE plot. Through this comparison, we can conclude that the method in [10] is better for compensating the nonlinearity at the transient stage while the proposed method has higher performance at the steady-state stage. Since these two methods aim at different types of PA nonlinearity during the dynamic power transmission, it is possible to combine them together to provide a complete solution.

VI. Conclusion

In this paper, a novel power adaptive digital predistortion technique has been proposed. By quantifying the linear and nonlinear changes of the PA behavior varying with the input power adjustments and deploying them to scale the DPD

coefficients, the dynamic distortion of the PA in the steady-state stage with multiple power levels can be effectively compensated without real-time recalibration. A simple online coefficient updating approach is also proposed to provide a fast and effective solution for occasional coefficients updating during real-time operation.

This proposed approach does not require any changes on the model structures; only two sets of coefficients need to be extracted during model extraction while the number of coefficients that are used in the real-time DPD operation remains the same. Measurement results confirm that the nonlinear distortions of the PA at multiple operation power levels can be almost fully compensated by employing the power adaptive method. This approach provides a very promising solution for future transmitters with dynamic power transmissions.

REFERENCES

[1] P. B. Kennington, *High Linearity RF Amplifer Design*. Norwood. MA: Artech House, 2000, pp. 397-408.

[2] L. Ding, G. T. Zhou, D. R. Morgan, Z. Ma, J. S. Kenney, J. Kim, and C. R. Giardina, "A robust digital baseband prediostorter constructed using memory polynomials," *IEEE Trans. Commun.,* vol. 52, no. 1, pp. 159-165, Jan. 2004.

[3] D. R, Morgan, Z. Ma, J. Kim, M. G. Zierdt, and J. Pastalan, "A generalized memory polynomial model for digital predistortion of RF power amplifiers," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3852-3860, Oct. 2006.

[4] A. Zhu, J. C. Pedro, and T. J. Brazil, "Dynamic deviation reduction based Volterra behavioral modeling of RF power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol.54, no. 12, pp. 4323-4332, Dec. 2006.

[5] A. Zhu, P. J. Draxler, J. J. Yan, T. J. Brazil, D. F. Kimball, and P. M. Asbeck, "Open-loop digital predistorter for RF power amplifiers using dynamic deviation reduction-based Volterra series," *IEEE Trans. Microw. Theory Techn.*, vol. 56, no. 7, pp. 1524-1534, Jul. 2008.

[6] L. Guan and A. Zhu, "Dual-loop model extraction for digital predistortion of wideband RF power amplifiers," *IEEE Microw. Wireless Compon. Lett.*, vol. 21, no. 9, pp. 501-503, Sept. 2011.

[7] T. Ali-Yahiya, *Understanding LTE and its Performance.* New: Springer, 2011, pp. 63-64.

[8] F. M. Ghannouchi and O. Hammi, "Behavioral modeling and predistortion," *IEEE Microw. Mag.,* vol. 10, no. 7, pp.52-64. Dec. 2009.

[9] O. Hammi, S. Boumaiza, and F. M. Ghannouchi, "On the robustness of digital predistortion function synthesis and average power tracking for highly nonlinear power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 55, no. 6, pp. 1382-1389, Mar. 2007.

[10] A. S. Tehrani, T. Eriksson, and C. Fager, "Modeling of long term memory effects in RF power amplifiers with dynamic parameters," in *Proc. IEEE MTT-S Int. Microw. Symp. Dig.*, Montreal, QC, Canada, Jun. 2012, pp. 1-3.

[11] O. Hammi, A. Kwan, and F. M. Ghannouchi, "Bandwidth and power scalable digital predistorter for compensating dynamic distortions in RF power amplifiers," *IEEE Trans. Broadcast.*, vol. 59, no. 3, pp. 520-527, Sept. 2013.

[12] Y. Guo, C. Yu, and A. Zhu, "Power adaptive digital predistortion for RF power amplifers," in *Proc. IEEE MTT-S Int. Microw. Symp. Dig.*, Tempa Bay, FL, USA, Jun, 2014, pp. 1-3.

[13] "3rd generation partnership project; technical specification group radio access network; user equipment (UE) radio transmission and reception (FDD) (Release 10)," 3GPP, Valbonne, France, Tech. Spec. 3GPP TS 25.101 V10.2.0 (2011-06), Jun. 2011.

[14] T. Schweickhardt and F. Allgower, "On systems gains, nonlinearity measures, and linear models for nonlinear systems," *IEEE Trans. Automat. Control,* vol. 54, no. 1, pp. 62-78, Jan. 2009.

[15] K. Emancipator and M. H. Kroll. "A quantitative measure of nonlinearity," *Clin. Chem.,*vol. 39, no. 5, pp. 766–772, 1993.

[16] A. Helbig, W. Marquardt, and F. Allgower. "Nonlinearity measures: definition, computation and applications," *J. Proc. Control*, vol. 10, no. 2, pp. 113–123, 2000.

[17] E. M. L. Beale. "Confidence regions in non-linear estimation," *J. R. Stat. Soc., Ser. B. Soc.*, vol. 10, pp.41–88, 1960.

[18] D. Kihas, H. J. Marquez. "Computing the distance between a nonlinear model and its linear approximation: an L_2 approach," *Comput. Chem. Eng.* vol. 28, pp. 2659-2666, 2004

[19] L. Guan and A. Zhu, "Optimized low-complexity implementation of least squares-based model extraction for digital predistortion of RF power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 60, no. 3, pp. 594-603, Mar. 2012.

[20] L. Guan, R. Kearney, C. Yu, and A. Zhu, "High performance digital predistortion test platform development for wideband RF power amplifiers," *Int. J. Microw. Wireless Tech.*, vol. 5, no. 2, pp. 149-162, Apr. 2013.

[21] L. Guan and A. Zhu, "Simplified dynamic deviation reduction-based Volterra model for Doherty power amplifiers," in *Proc. IEEE Int. Integer. Nonlinear Microw. Millimeter-Wave Circuits Workshop*, Vienna, Austria, Apr. 2011, pp.1-4.

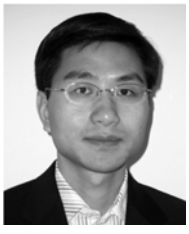
[22] A. Zhu, P. J. Draxler, C. Hsia, T. J. Brazil, D. F. Kimball, and P. M. Asbeck, "Digital predistortion for envelope-tracking power amplifiers using decomposed piecewise Volterra series," *IEEE Trans. Microw Theory Techn.*, vol. 56, no. 10, pp. 2237-2247, Oct. 2008.

Yan Guo (S'13) received the B.E. degree in information science and engineering from East China Jiaotong University, Nanchang, Jiangxi Province, China, in 2007, the M.E. degree in communication and information systems from Southeast University, Nanjing, China, in 2011, and is currently working toward the Ph.D. degree with University College Dublin (UCD), Dublin, Ireland.

He is currently with the RF and Microwave Research Group, UCD. His research interests include spectrum sensing for cognitive radio, digital predistortion for RF power amplifiers, and field-programmable gate-array (FPGA) hardware implementations.

Chao Yu (S'09- M'15) received the B.E. degree in information engineering and M.E. degree in electromagnetic fields and microwave technology from Southeast University (SEU), Nanjing, China, in 2007 and 2010, respectively, and the Ph.D. degree in electronic engineering from University College Dublin (UCD), Dublin, Ireland, in 2014.

He is currently an Associate Professor with the State Key Laboratory of Millimeter Waves, School of Information Science and Engineering, SEU. His research interests include RF power amplifiers modeling and linearization, high-speed ADC digital correction. He is also interested in antenna design, FPGA hardware implementation and RF wireless system design.



Anding Zhu (S'00–M'04–SM'12) received the B.E. degree in telecommunication engineering from North China Electric Power University, Baoding, China, in 1997, the M.E. degree in computer applications from the Beijing University of Posts and Telecommunications, Beijing, China, in 2000, and the Ph.D. degree in electronic engineering from University College Dublin (UCD), Dublin, Ireland, in 2004.

He is currently a Senior Lecturer with the School of Electrical, Electronic and Communications Engineering, UCD. His research interests include high frequency nonlinear system modeling and device characterization techniques with a particular emphasis on Volterra-series-based behavioral modeling and linearization for RF power amplifiers (PAs). He is also interested in wireless and RF system design, digital signal processing, and nonlinear system identification algorithms.