CHALMERS
UNIVERSITY OF TECHNOLOGY

Waveform optimization for high efficiency 5G operation

Master's thesis in Communication Engineering

Yibo Wu

Department of Electrical Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2019

Master's thesis 2019:EENX30

Waveform optimization for high efficiency 5G operation

Yibo Wu

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Electrical Engineering
Chalmers University of Technology
Gothenburg, Sweden 2019

Master Thesis for the Master Programme "Communication Engineering"
Waveform optimization for high efficiency 5G operation
Yibo Wu

Supervisor: Ulf Gustavsson, Ericsson Research
Examiner: Thomas Eriksson, Department of Electrical Engineering

Master's Thesis 2019:EENX30
Department of Electrical Engineering
Chalmers University of Technology
SE-412 96 Gothenburg

Abstract

In the upcoming 5G wireless communication system, the radio frequency (RF) power amplifier (PA) is a critical component. The signal with low power is amplified by the PA to the high power signal with as less signal quality loss as possible. The performance of the PA is limited by a trade-off between linearity and efficiency. Some techniques such as digital predistortion (DPD) and crest factor reduction (CFR) have been studied to address this trade-off, where the waveform of the PA's input signal is optimized such that the output signal of PA is linearly amplified. One technique is to combine the CFR with DPD, called CFR-DPD, while we argue that CFR and DPD perform nearly inverse operations in this technique. In this thesis, we proposed a CFR-free DPD technique where the CFR is removed but a simple limiter is placed after the DPD for PA safety protection. Simulations on the simulated and experimental power amplifiers (PAs) are implemented, respectively. The comparison of CFR-DPD and the proposed CFR-free DPD is analyzed by the distortion level at the output of a PA, which is evaluated by the normalized mean square error (NMSE) between the input and output. Also, the peak-to-average power ratio (PAPR) of the PA's input signal and the power spectral density (PSD) of the PA's output signal are evaluated. Results show that the proposed CFR-free DPD has lower signal distortion level as well as better control of signal's PAPR than CFR-DPD.

Keywords: Power amplifier (PA), linearization, digital predistortion (DPD), crest factor reduction (CFR).

Acknowledgements

I would like to thank my two supervisors, namely Prof. Thomas Eriksson at Chalmers University of Technology and Dr. Ulf Gustavsson at Ericsson Research. Thank you Thomas, for the fruitful discussions and great suggestions, and thank you Ulf, you are always kind to help me from the first day at Ericsson, and patient in answering all my questions in research and life.

I would also like to thank Adrián Lahuerta Lavieja at Catholic University of Leuven for the valuable discussions and comments. Finally, I would like to thank Dr. Milad Fozooni and Dr. Mingquan Bao at Ericsson Research for encouraging and supporting me.

This thesis has been carried out at Chalmers University of Technology and Ericsson Research in Gothenburg.

Yibo Wu
Gothenburg, September 5, 2019

Contents

Acronyms

5G fifth generation
ACPR adjacent channel power ratio
ADC analog-to-digital converter
AM/AM amplitude-to-amplitude
AM/PM amplitude-to-phase
CF crest-factor
CFR crest factor reduction
DAC digital-to-analog converter
DPD digital predistortion
EVM error vector magnitude
GMP generalized memory polynomial
ICF iterative clipping and filtering
ILA indirect learning architecture
ILC iterative learning control
ILC-DPD ILC-based DPD
MP memory polynomial
NMSE normalized mean square error
OFDM orthogonal frequency division multiplexion
PA power amplifier
PAPR peak-to-average power ratio
PAs power amplifiers
PSD power spectral density
RF radio frequency
RMS root mean squares
VST vector signal transceiver

List of Figures

List of Tables

1

Introduction

1.1 Background

In the upcoming fifth generation (5G) wireless communication system, multicarrier transmission, such as orthogonal frequency division multiplexion (OFDM), is of great importance. OFDM presents attractive advantages such as high spectral efficiency, immunity to multipath fading, and simple channel equalization [1]. However, one major drawback of OFDM signal is the high PAPR, which challenges the operation range of the RF PAs.

The PA aims to convert low power signals to a suitable high power signal that can be transmitted by antennas. The ideal PA have constant gain before the saturation point However, in practice, PA design is limited by linearity and efficiency constraints[2]. High signal peaks are strongly compressed by the PA, which leads to in-band errors and out-of-band emissions of the output signal. To smooth the nonlinear behaviour and relieve the compression effect, one solution is to back off the PA operation region to a lower power level which is far from the saturation power point, so the amplified signal experiences less nonlinear distortion. However, the amplified signal by the PA obtains smaller power gain which may not be enough for transmission. Besides, the power conversion of the PA is also inefficient in this region. So it is no surprise that techniques dealing with the PA trade-off between linearity and efficiency have attracted many research interests.

For many years, different techniques have been developed to address the linearity-efficiency trade-off. Some of them are dedicated to extending the linearity behavior to the high power region, so-called linearization techniques. DPD is one of the most popular PA linearization techniques in the modern communication system. It adds a complementary operation before the PA to compensate the PA nonlinear compression effects. By this compensation, the amplitude and phase of the signal are distorted in a manner that the output signal of the PA is linearly amplified to the input signal of DPD. To achieve this compensation, DPD finds the inverse operation of the PA. The block diagram of this system is shown in Figure. 1.1 and the behaviors of PA and DPD are shown in Figure 1.2. In practice, DPD techniques are challenged by the PA identification problem and DPD modeling limitation. DPD-avalanche [3] is another problem. To compensate the high compression near the saturation point of PA, DPD distorts the signal to extremely high peaks which may be physically unachievable, and these peaks suffer more severe compression. This

problem leads to a strong performance deterioration.

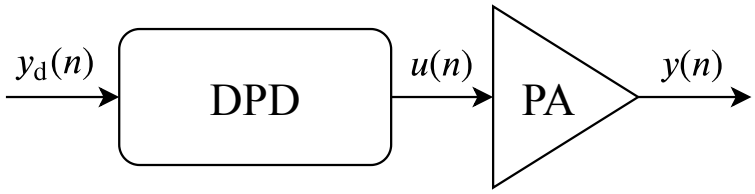


Figure 1.1: System model. DPD distorts the input $y_d(n)$ such that the PA output $y(n)$ is linearly amplified.

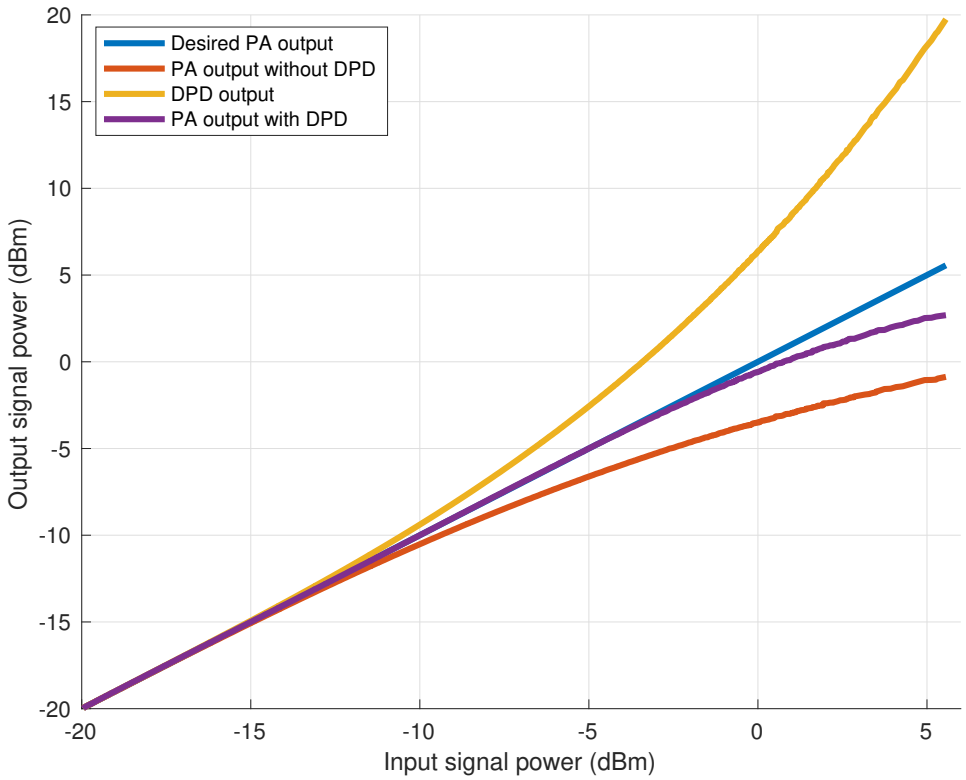


Figure 1.2: The behaviors of the PA and DPD. For the sake of analysis convenience, the power of each signal are normalized to the same level.

Another approach to address the PA trade-off is to reduce the high PAPR (also known as high crest factor) of the input signal, namely CFR [1] techniques. Because signals with high crest factor have some high peaks,these peaks may suffer from the PA compression even if the average output power level is low. To reduce the signal quality loss, the PA has to work at a low power level, which is not power efficient. So if the crest factor of the signal can be reduced, PAs can operate at a higher power level. A number of CFR techniques have been proposed for multicarrier transmission such as clipping and filtering [4, 5], coding [6], interleaving [7] and selected mapping [8], etc. Iterative clipping and filtering (ICF) [5] is one of the

most used techniques because of simple implementation and hardware friendly. The block diagram and the behavior of ICF are shown in Figure 1.3, and both time- and frequency-domain signals at each blocks are illustrated in Figure 1.4. ICF first clips the signal amplitude to the desired level (y to y_c), and then filters the signal out-of-band distortion out (y_c to y_{cf}). The filtering process generates amplitude peaks again in y_{cf}, so these two steps are iteratively performed to restrain the amplitude regrowth. At the cost of signal quality loss, ICF achieves crest factor reduction as well as keeps a small adjacent channel interference level. Hence, the PA can perform at a higher power level when the signal has a small crest factor.

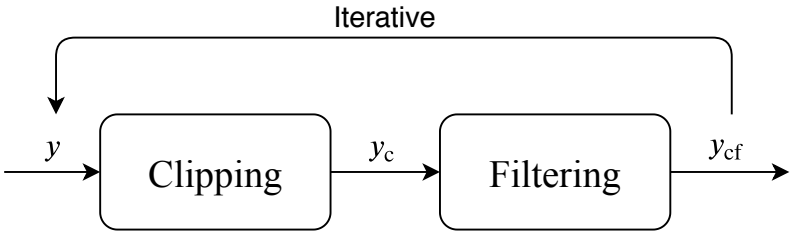


Figure 1.3: The block diagram of iterative clipping and filtering.

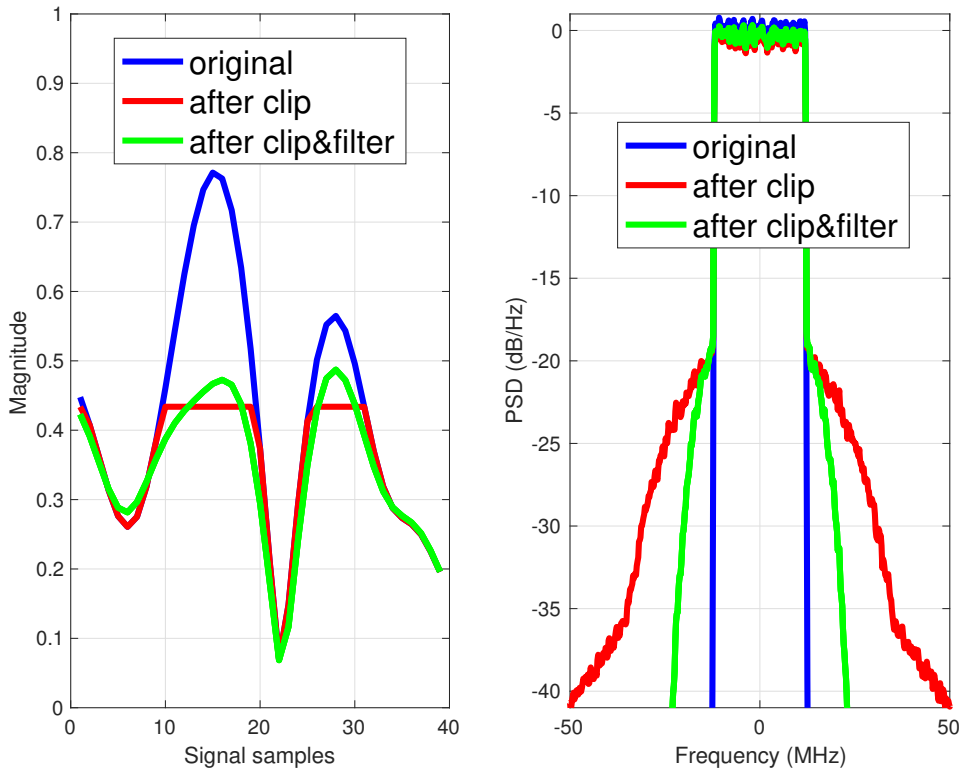
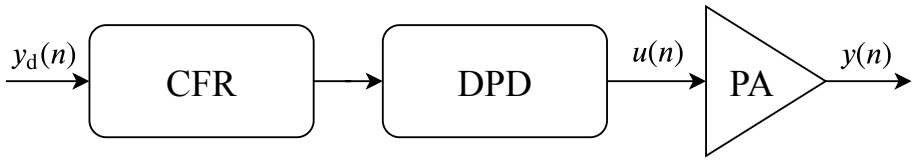
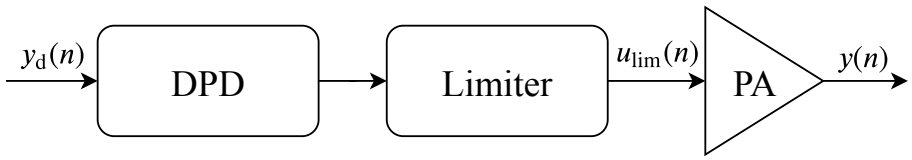


Figure 1.4: Time and frequency domain signals before clipping, after clipping, and after filtering. Clipping introduces out-of-band distortion. Although filtering restricts these out-of-band leakage, it also reproduces peaks in time domain.

Recently, some studies have presented techniques of combining CFR and DPD [9, 10], where CFR is applied before DPD. We denote CFR-DPD in this thesis. It aims to push the PA operation region to a higher power level so that the PA is more efficient. However, we argue that CFR and DPD perform inverse operations on the signal, and putting CFR before DPD seriously affect the quality of PA's output. The signal distortion at the output will not benefit from using CFR. Although the combined system offers advantages on reducing out-of-band distortion, this benefit can be achieved by some other operations, such as weighted DPD method. We proposed a CFR-free DPD structure. The block diagrams of CFR-DPD and the proposed CFR-free DPD are shown in Figure 1.5.



(a) Block diagram of CFR-DPD structure.



(b) Block diagram of the proposed CFR-free DPD structure.

Figure 1.5: System models

1.2 Thesis contribution

In this thesis, we proposed a PA linearization system without CFR, namely CFR-free DPD, where a simple limiter is placed after the DPD. The performance of the proposed system is compared with the CFR-DPD system in terms of multiple metrics, such as NMSE, adjacent channel power ratio (ACPR), etc. Results show that the proposed CFR-free DPD structure has better PA linearization performance than CFR-DPD structure with respect to output signal distortion. The output signal of the proposed structure has advantage on signal distortion level. This advantage is much greater at high output power region, which is also the interest region in practice. Meanwhile, the limiter of the proposed method can not only protect the PA safety but also reduce the signal's crest-factor (CF) greatly.

More specifically, two predistortion identification methods, indirect learning architecture (ILA) [11] and iterative learning control (ILC) [12] are applied. Results on two memoryless PA models verified the theoretical NMSE lower bound and proved the advantage of the proposed method. It also showed that ILC has the best noise robustness and linearity performance. What's more, experimental results are also obtained by RF WebLab [13], where a class-AB PA is employed. And the experimental results offered further proof for our statement that the proposed system has

lower signal distortion level compared to CFR-DPD system.

1.3 Thesis outline

The rest of the thesis chapters will be organized as follows. Chapter 2 introduces the theory of PA behavior modeling. Chapter 3 presents the concept of different DPD identification algorithms and a CFR technique. The proposed CFR-free DPD is described in Chapter 4 which also contains the evaluation strategies. Finally, the next Chapter presents numerical results, concludes the thesis, and discusses future work.

2

Power Amplifier Modeling

2.1 Nonlinearity of Power Amplifiers

The nonlinearity of PAs consists of two main parts, static nonlinearity, and memory effects. Static nonlinearity causes the PA to compress the signal with high power, mainly exhibiting amplitude-to-amplitude (AM/AM) nonlinearity. Memory effects mean that the nonlinearity of PA is time-dependent, so output signals are affected by not only the simultaneous input signal but also the previous input signal. Memory effects can be observed in the AM/AM and amplitude-to-phase (AM/PM) conversions of PAs. In order to have a reliable modeling performance of the PA, both static nonlinearity and memory effects should be considered properly.

Static nonlinearity

In the narrow-band case, the PA is dominated by static nonlinearity, which happens in the PA compression region. The higher output signal power goes, the more severe compression effect it suffers. This phenomenon is usually evaluated by the AM/AM conversion, which represents the amplitude relationship between input signals and output signals. The AM/AM characteristic of a PA is shown in Figure 2.1, where measurements are obtained from RF WebLab [13]. The input/output relation is almost stable at the lower power region, namely the linear region, and it becomes very nonlinear when the input power is higher than 25 dBm, namely the compression region.

Memory effects

In the wide-band case, memory effects cannot be ignored. Because of PA frequency selectivity, signal with large bandwidth experiences different gain. This selectivity becomes to be time-dependence in the time domain, where the output signal depends on both the current input and the previous input. This phenomenon is known as memory effects of PAs. It may result in dynamic distortions such as asymmetries in inter-modulation and dynamic AM/PM characteristic [14]. The memory effect on AM/PM is shown in Figure 2.2. The attributions of memory effects come from different aspects, such as thermal effects and electrical effects [15]. Thermal effects happen when the PA temperature is increased by the dissipated power. The temperature change process is often slow, so thermal effects are also called long-term memory effects [16]. Meanwhile, electrical effects appear due to the varying bias-

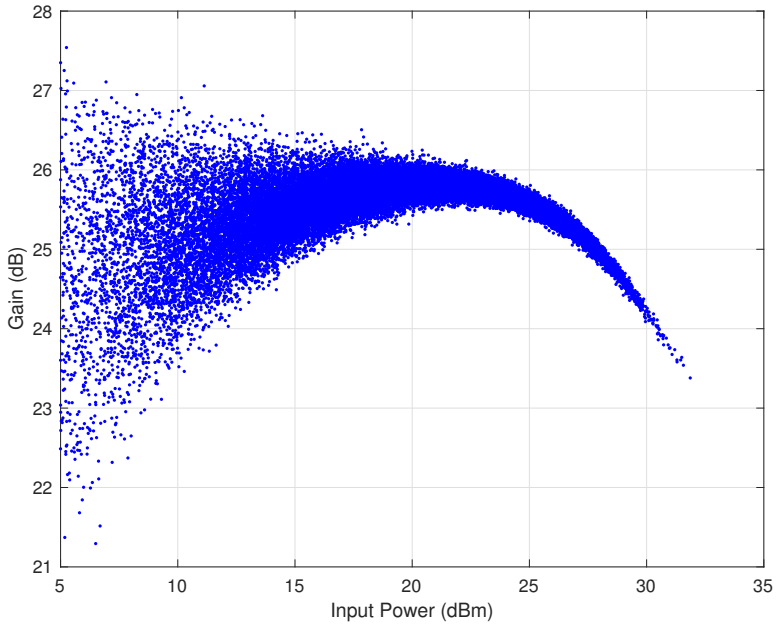


Figure 2.1: AM/AM characteristic of the class-AB PA used in RF WebLab.

ing and impedance in PAs, and it is usually called short-term memory effects [16] because of short time duration.

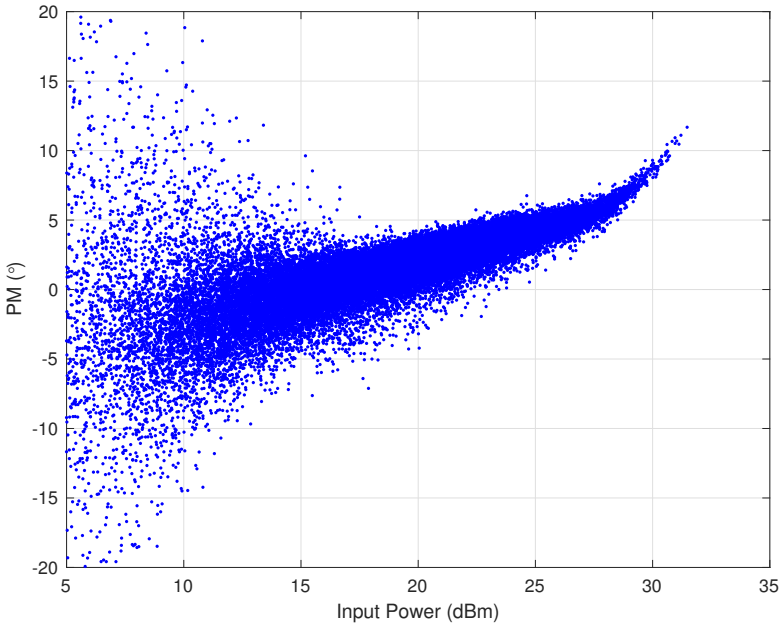


Figure 2.2: AM/PM characteristic of the class-AB PA used in the RF WebLab.

2.2 Behavioral Modeling

To represent these PA characteristics, a PA model is essential. The PA modeling process is often called behavioral modeling, also known as black-box modeling. It refers to modeling the nonlinear distortion and memory effects of the PA. There are plenty of modeling methods. One intuitive way is to use the magnitude and phase relationship between the input and output signal, which is well-known as the AM/PM model [17]. However, this model is computationally expensive, and it has a limited performance on dealing with memory effects, which is a serious problem in the wide bandwidth system, e.g., 5G wireless communication system. A good PA model should take into account both the linearity performance and computation complexity.

In this section, we describe several types of PA memory models, which are mainly the Volterra series based memory models. We first introduce the concept of Volterra series and then describe some variant Volterra models such as memory polynomial and generalized memory polynomial. These models simplify the complexity of Volterra series and have great modeling performance.

Volterra series based memory models

To have a good modeling performance of the PA, both nonlinearity and memory effects are essential to be considered. One most common modeling approach is the Volterra series [18], which describes the nonlinear relation between the output and input of a system by considering multiple input instances. Volterra series is widely used in modeling PA, and many papers have been published.

The block diagram of the PA modeling is shown in Figure 2.3. The relation between the PA input signal $u(n)$ and the PA output signal y_n can be represented by discrete Volterra series, which is a sum of multidimensional convolutions:

$$y(n) = \sum_{k=1}^{K} y_k(n) \tag{2.1}$$

where $y_k(n)$ is the k-dimensional convolution

$$y_k(n) = \sum_{m_1=0}^{M-1} \cdots \sum_{m_k=0}^{M-1} h_k(m_1, \cdots, m_k) \prod_{j=1}^{k} u(n - m_j) \tag{2.2}$$

where $h_k(m_1, \cdots m_k)$ is the k-th Volterra kernel with a nonlinear order K and a memory length M. Due to the kernel symmetry of (2.2), redundant summation indices can be removed [19], so (2.2) can be rewritten to

$$y_k(n) = \sum_{m_1=0}^{M-1} \sum_{m_2=m_1}^{M-1} \cdots \sum_{m_k=m_{k-1}}^{M-1} h_k(m_1, \cdots m_k) u(n) \prod_{j=1}^{k-1} u(n - m_j) \tag{2.3}$$

Although Volterra series have good performance on modeling the PA, it is challenged

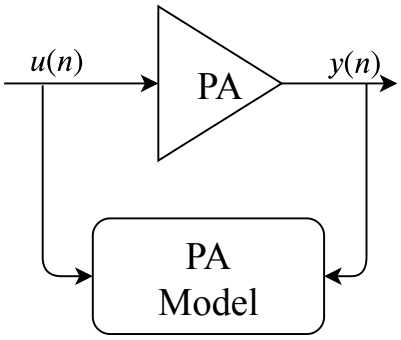


Figure 2.3: Block diagram of the PA modeling process.

by the high computational complexity because the number of Volterra parameters increase rapidly as the memory depth M and nonlinear order K grow up. To address the computational complexity issue, several simplified Volterra series algorithms have been studied. Because of simplicity and good performance, they are more generally used. We choose two simplified Volterra series models to introduce, the memory polynomial (MP) and the generalized memory polynomial (GMP) model.

Memory Polynomial

One of the most popular variant Volterra series is the MP model [20]. It simplifies the process (2.3) by only keeping the diagonal Volterra kernels $h_k(m, m, m, \cdots, m)$ and set other kernels to zero. The kernels are changed to

$$h_k(m_1, m_2, \cdots, m_k) = \begin{cases} a_{km}, & \text{if } m_1 = m_2 = \cdots = m_k \\ 0, & \text{otherwise} \end{cases} \tag{2.4}$$

where a_{km} is the coefficient at nonlinear order k and memory depth m. Eventually, the modified modeling step can be rewritten to [21]

$$y(n) = \sum_{k=1}^{K} \sum_{m=0}^{M} a_{km} u(n-m)|u(n-m)|^{k-1} \tag{2.5}$$

where K and M are the maximal nonlinear order and memory depth, respectively. In general, the even-order nonlinear terms $u(n-m)|u(n-m)|^{2(k-1)}$ need more concern because they are usually inside of the signal bandwidth. Thus MP can be further simplified by removing the odd-order terms, and the modeling equation becomes

$$y(n) = \sum_{k=1}^{K} \sum_{m=0}^{M} a_{2k-1,m} u(n-m)|u(n-m)|^{2(k-1)} \tag{2.6}$$

Compared to the general Volterra series, the computational complexity of the MP model is dramatically reduced because the number of coefficients drops from $(M+1)^K$ to $K(M+1)$ [21]. Meanwhile, MP is claimed to have great ability of modeling.

Generalized memory polynomial

Based on the MP model, GMP model [22] take into account the relation between signals with different time instants. Because the product of input signals, namely $u(n-m)|u(n-m)|^{k-1}$ in (2.5), only considers signals at the same time $(n-m)$, while the general Volterra series has products of signals at different time, namely $\prod_{j=1}^{k} u(n-m_j)$ in (2.2). These products are often called cross terms. Based on MP model, GMP model adds cross terms, which can be expressed as

$$\sum_{k=1}^{K}\sum_{m=0}^{M} b_{km}u(n)|u(n-m)|^{k-1} \tag{2.7}$$

where the memory polynomial term is formed by the product of signal $u(n)$ and its time delay signal $u(n-m)$. Combining (2.7) with 2.5, the modeling equation of GMP is formed by [23]

$$\begin{aligned} y(n) =& \sum_{k=1}^{K}\sum_{m=0}^{M} a_{km}u(n-m)|u(n-m)|^{k-1} \\ &+ \sum_{k=2}^{K}\sum_{m=0}^{M}\sum_{\substack{l=-m \\ l\neq 0}}^{L} b_{kml}u(n-m)|u(n-m-l)|^{k-1} \end{aligned} \tag{2.8}$$

where K, M, L are the nonlinear order, memory depth, and cross-term length, respectively. Note the cross-term length is limited by m because the GMP is causal which means the future signal $u(n+1)$ should not affect the current situation.

By adding cross terms to the MP model, GMP builds more possible relations between the input/output signal. GMP has advantages on reaching a better modeling performance, and the computation complexity increase is gentle.

2.3 Parameter estimation method

Volterra series based models are popular, and one reason is that they are linear in parameters. Estimation of those model parameters becomes easy, and one simple method is the least squares [24], which is introduced in the following section.

Least Squares

First, the PA modeling process can be written into a matrix expression form:

$$\mathbf{y} = \mathbf{U}\mathbf{w} \tag{2.9}$$

where $\mathbf{w}$ denotes the vector of model parameters such as a_{km} and b_{km}, $\mathbf{U}$ represents the matrix of PA inputs and $\mathbf{y}$ denotes the matrix of PA outputs. To find the parameters $\mathbf{w}$ that match the input $\mathbf{U}$ and output $\mathbf{y}$ best, least squares method aims to seek a $\mathbf{w}$ that minimize the square errors between the observation $\mathbf{y}$ and estimation $\hat{\mathbf{y}}$, written as

$$\arg\min_{\mathbf{w}} ||\mathbf{e}||^2 = \arg\min_{\mathbf{w}} ||\hat{\mathbf{y}} - \mathbf{y}||^2 \tag{2.10}$$

a standard solution for (2.10) is given as

$$\mathbf{w} = (\mathbf{U}^H\mathbf{U})^{-1}\mathbf{U}^H\mathbf{y} \tag{2.11}$$

where $(\cdot)^{-1}$ is the Moore-Penrose pseudo-inverse [25] and $(\cdot)^H$ denotes the Hermitian transpose.

3

Digital Predistortion and Crest Factor Reduction Techniques

3.1 Learning architectures for DPD

DPD aims to provide a compensating operation before the PA to cancel out some of the nonlinear effects caused by compression, phase noise, etc. It is designed to find an inverse model of PA such that the output signals of PA are linear amplified. The block diagram of this process is shown in Figure 3.1.

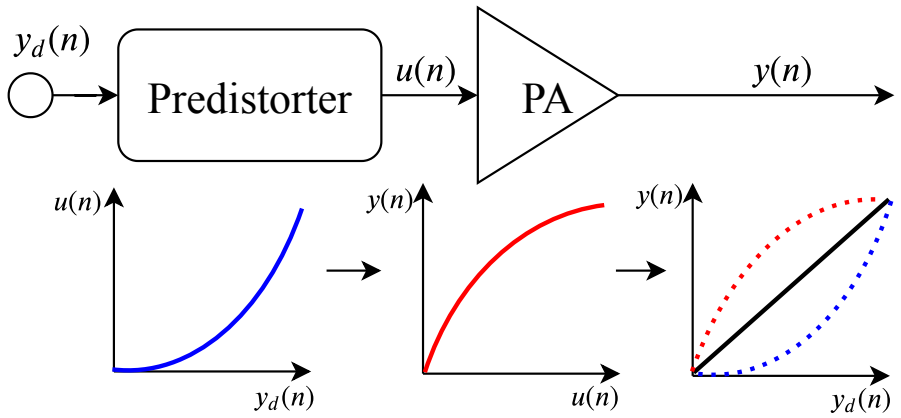


Figure 3.1: The process of pre-distortion in block diagrams. An ideal DPD seeks a inverse operation of the PA so that the output from PA is linearly amplified. To facilitate the understanding of each block, the gain of PA is removed.

Recalling the PA behavioral models presented in Chapter 2, once we have the input $u(n)$ and the output $y(n)$ of a PA, it is easy to calculate parameters of a Volterra series based model using least squares estimation. However, we can not use this way to compute the parameters of DPD because the optimal output of pre-distorter, namely $u(n)$, is unknown [26]. To address this estimation problem, different learning architectures have been proposed, such as ILA, ILC, etc., which are introduced in the following sections.

3.1.1 Indirect learning architecture

One common structure to estimate the pre-distorter parameters is the ILA, which attracts many research interests [11, 27, 28] due to the simple implementation. The

block diagram of ILA is shown in Figure 3.2. The indirect learning means the parameters of pre-distorter are estimated indirectly. Specifically, ILA puts a distorter after the PA which is usually called post-distorter. The coefficients of post-distorter can be estimated easily by using a behavioral model such as MP model. Applying the MP model equation (2.5) into the post-distorter :

$$\hat{u}(n) = \sum_{k=1}^{K} \sum_{m=0}^{M} a_{km} y(n-m)|y(n-m)|^{k-1} \tag{3.1}$$

where the input of the post-distorter is the output sequence $y(n)$ of PA. Similar to (2.9), we can express the matrix form of (3.1) by considering all the N samples

$$\hat{\mathbf{u}} = \mathbf{Y}\mathbf{w} \tag{3.2}$$

where $\hat{\mathbf{u}} = [\hat{u}(1), \cdots, \hat{u}(N)]^T$, coefficient vector $\mathbf{w} = [a_{10}, \cdots, a_{km}, \cdots, a_{KM}]^T$, input matrix $\mathbf{Y} = [Y_{10}, \cdots, Y_{km}, \cdots, Y_{KM}]$ and

$$Y_{km} = \begin{bmatrix} y(1-m)|y(1-m)|^{k-1} & \cdots & y(N-m)|y(N-m)|^{k-1} \end{bmatrix}$$

The coefficients $\mathbf{w}$ can be estimated by least squares, which minimizes the square error between $\hat{\mathbf{u}}$ and $\mathbf{u}$,

$$\arg\min_{\mathbf{w}} ||\mathbf{u} - \hat{\mathbf{u}}||^2 \tag{3.3}$$

and the coefficients $\mathbf{w}$ is estimated as

$$\mathbf{w} = (\mathbf{Y}^H\mathbf{Y})^{-1}\mathbf{Y}^H\mathbf{u} \tag{3.4}$$

Finally, the parameters of pre-distorter can use a copy of post-distorter parameters due to the p-th order inverse theory [29]. In general, this procedure could iterate more than one time to have a stable estimation of pre-distorter coefficients [27].

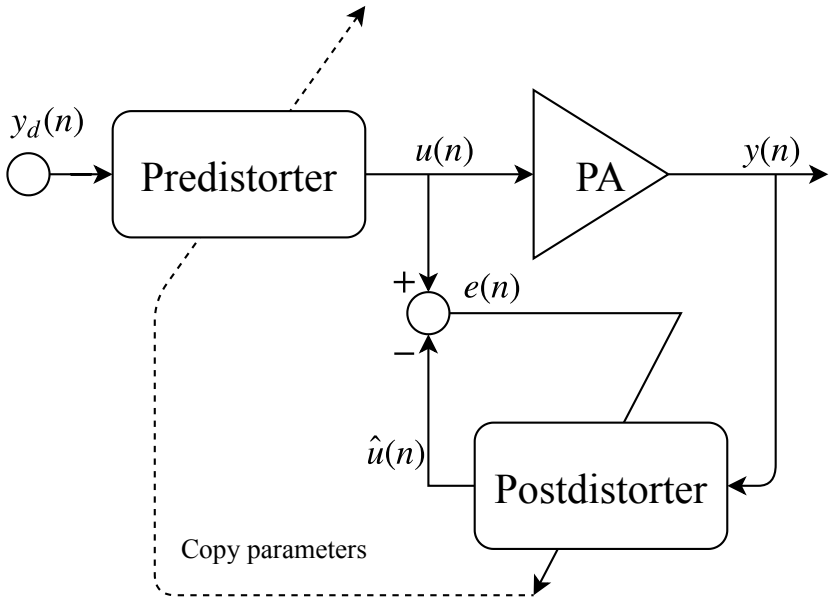


Figure 3.2: The structure of ILA.

In a traditional ILA, the PA output $y(n)$ is normalized by a given gain G such that the power level of $y(n)$ is the same as the DPD input signal. The convergence situation is $y(n) \approx G y_{\mathrm{d}}(n)$. However, the author in [28] argued that the choice of gain G directly affected the output power level and proposed a variant ILA structure where the gain normalization step is removed. The variant ILA is shown in Figure 3.2. The convergence situation is changed to $y(n) \approx y_{\mathrm{d}}(n)$, so there is no need for normalization. It was shown in [28] that the output power is no longer depended on the gain G.

There are many drawbacks for ILA, but it is still the most popular DPD identification structure due to simplicity and good performance. The first drawback is that the noise of PA's output measurement $y(n)$ affects the estimation accuracy of post-distorter seriously, which could lead to a biased convergence. This problem can be even worse in practical applications because additional noise is produced when the analog signal $y(n)$ is converted to a digital signal by a analog-to-digital converter (ADC) with large noise. Besides, post-distorter represents the post-inverse model of the PA, but a copy of that distorter cannot perfectly represent the pre-inverse model because the nonlinear filter cannot be permuted [30, 31]. This problem is particularly serious for highly nonlinear PAs.

3.1.2 Iterative learning control

In the PA modeling process, it is easy to identify model parameters once the input and the output of PA are known. However, the parameter identification in DPD is not as easy because the optimal output of the DPD is unknown, which constrain the performance of ILA. One of the motivations to choose another method is to find the optimal DPD output, which is also the input to the PA. ILC [12] is such a method to find the optimal input of a system. The input is iteratively updated until the system output equal to the desired output. Specifically, at each iteration, the system input updates itself by learning from the current output. This iterative procedure stops until the input converges. Instead of estimation some model parameters, ILC directly control the input, which is the main difference from ILA. A block diagram of ILC is shown in Figure 3.3.

The input of PA at iteration k is denoted by $u_k(n)$, and the current output of PA is denoted by $y_k(n)$, which can be represented by $y_k(n) = F_n[u_k(n), u_k(n-1)\cdots]$, where F_n is defined as the function of PA at the nth sample instance. By considering all N samples of signal, the vector form of the PA process can be rewritten as

$$\mathbf{y}_k = \mathbf{F}(\mathbf{u}_k) \tag{3.5}$$

where $\mathbf{u}_k = [\mathbf{u}_k^0, \cdots, \mathbf{u}_k^n, \cdots \mathbf{u}_k^{N-1}]^T$, $\mathbf{y}_k = [y_k(0), y_k(1), \cdots y_k(N-1)]^T$, $\mathbf{F}$ denotes the vector-valued function of PA and the element $\mathbf{u}_k^n$ in $\mathbf{u}_k$ denotes $[u_k(0), \cdots, u_k(n)]$. The error vector $\mathbf{e}_k$ at iteration k between the desired output vector $\mathbf{y}_{\mathrm{d}}$ and PA output vector $\mathbf{y}_k$ can be written as

$$\mathbf{e}_k = \mathbf{y}_{\mathrm{d}} - \mathbf{y}_k \tag{3.6}$$

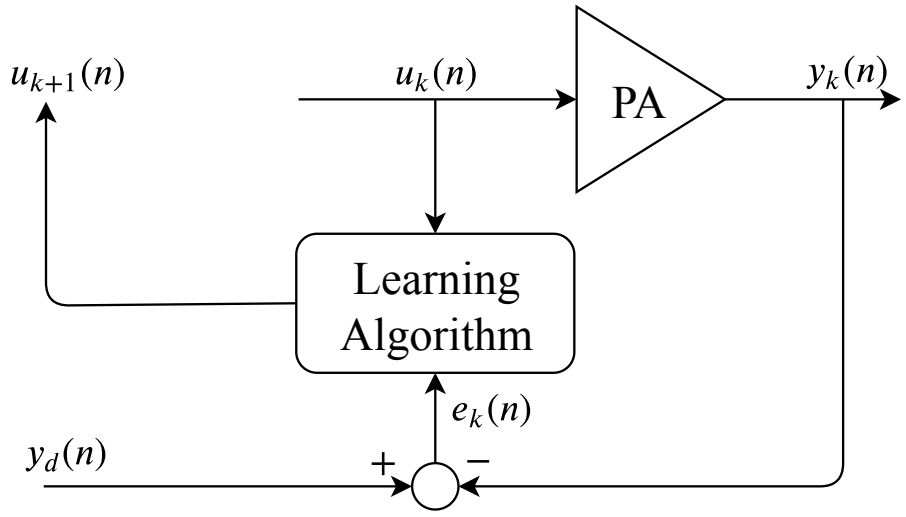


Figure 3.3: The structrue of ILC.

where $\mathbf{e}_k = [e_k(0), \cdots, e_k(N-1)]^T$, and $\mathbf{y}_\mathrm{d} = [y_\mathrm{d}(0), y_\mathrm{d}(1), \cdots y_\mathrm{d}(N-1)]^T$. Then, based on the error vector $\mathbf{e}_k$, the current input $\mathbf{u}_k$ can be updated to the next iteration input $\mathbf{u}_{k+1}$ by a learning algorithm. The updating step can be expressed by

$$\mathbf{u}_{k+1} = \mathbf{u}_k + \mathbf{\Gamma}\mathbf{e}_k \tag{3.7}$$

where $\mathbf{\Gamma}$ denotes the learning matrix. Eventually, this process is iterated until the error converges to the desired value. In that situation $\mathbf{y}_k \approx \mathbf{y}_\mathrm{d}$. There are several ILC learning algorithms controlling the update step (3.7). A proper learning matrix $\mathbf{\Gamma}$ should guarantee the convergence condition and also provide a fast convergence speed. (3.6) can be rewritten to

$$\mathbf{e}_k = \mathbf{y}_\mathrm{d} - \mathbf{F}(\mathbf{u}_k) \tag{3.8}$$

where $\mathbf{e}_k$ can be view as a vector-valued function of the vector input $\mathbf{u}_k$. The purpose is to find an input $\mathbf{u}_k$ such that the function output $\mathbf{e}_k$ is a zero vector, in other words, to find the roots of the vector-valued function (3.8). One root-finding method is the Newton's method that the roots of functions are approached iteratively by

$$\mathbf{u}_{k+1} = \mathbf{u}_k - J_{\mathbf{e}_k}(\mathbf{u}_k)^{-1}\left(\mathbf{y}_\mathrm{d} - \mathbf{F}(\mathbf{u}_k)\right) \tag{3.9}$$

$$= \mathbf{u}_k + J_{\mathbf{F}}(\mathbf{u}_k)^{-1}\mathbf{e}_k \tag{3.10}$$

where $\mathbf{J}_{\mathbf{e}_k}(\mathbf{u}_k)$ is the Jacobian matrix of $\mathbf{e}_k$ with respect to $\mathbf{u}_k$ and $\mathbf{J}_{\mathbf{e}_k}(\mathbf{u}_k) = -\mathbf{J}_{\mathbf{F}}(\mathbf{u}_k)$. $\mathbf{J}_{\mathbf{F}}(\mathbf{u}_k)$ is defined as

$$\mathbf{J}_{\mathbf{F}}(\mathbf{u}_k) = \left[\frac{d\mathbf{F}}{d\mathbf{u}_k^0}, \cdots, \frac{d\mathbf{F}}{d\mathbf{u}_k^{N-1}}\right] = \begin{bmatrix} \frac{dF_0}{u_k(0)} & \cdots & \frac{dF_0}{u_k(N-1)} \\ \vdots & \ddots & \vdots \\ \frac{dF_{N-1}}{u_k(0)} & \cdots & \frac{dF_{N-1}}{u_k(N-1)} \end{bmatrix} \tag{3.11}$$

Thus, the learning matrix $\mathbf{\Gamma}$ can be selected as the inverse of the first order partial derivative matrix, namely $\mathbf{J}_{\mathbf{F}}(\mathbf{u}_k)^{-1}$. Although the Newton's method provides fast

convergence, it requires the prior knowledge of function $\mathbf{F}$, which is difficult to obtain and to guarantee accuracy. Besides, the computation of calculating the inverse of Jacobian matrix $\mathbf{J_F}(\mathbf{u}_k)$ is high.

Considering the drawbacks of Newton's method, some alternatives of learning matrix $\mathbf{\Gamma}$ are proposed in [12]. One method is to use the gain of PA as $\mathbf{\Gamma}$, called gain-based learning type. The update step (3.10) changes to

$$\mathbf{u}_{k+1} = \mathbf{u}_k + \mathbf{G}(\mathbf{u}_k)^{-1}\mathbf{e}_k \tag{3.12}$$

where $\mathbf{G}(\mathbf{u}_k)$ is defined as a diagonal gain matrix

$$\mathbf{G}(\mathbf{u}_k) = \mathrm{diag}\left[G[u_k(0)], \cdots, G[u_k(N-1)]\right] = \mathrm{diag}\left[\frac{y_k(0)}{u_k(0)}, \cdots, \frac{y_k(N-1)}{u_k(N-1)}\right] \tag{3.13}$$

Instead of calculating the Jacobian matrix $\mathbf{J_F}(\mathbf{u}_k)$, the gain matrix is more straightforward to compute. Meanwhile, this method does not require the prior knowledge of the PA function. It also has an adequate convergence rate compared with the Newton's method.

Another method that further simplifies the computation complexity is using a linear learning matrix, called linear type. The update step can be rewritten to

$$\mathbf{u}_{k+1} = \mathbf{u}_k + \lambda\mathbf{e}_k \tag{3.14}$$

where λ is the linear gain. This linear-based method greatly simplify the learning process, and it is shown in [12, 32, 33] that the convergence condition can be guaranteed if $0 < \lambda < \frac{2}{\mathbf{J_{max}}}$, where $\mathbf{J_{max}}$ is the maximum of the diagonal entries of $\mathbf{J_F}(\mathbf{u}_k)$.

3.1.3 Iterative learning control based digital predistortion

Given a desired PA output $y_\mathrm{d}(n)$, ILC iteratively find the optimal PA input $u_\mathrm{opt}(n)$, which can be viewed as the optimal DPD output. Then, a DPD identification algorithm can be used to build the predistortion model whose input is $\mathbf{y}_\mathrm{d}$ and output is $\mathbf{u}_\mathrm{opt}$. This special DPD algorithm is called ILC-based DPD (ILC-DPD) [12], and the block diagram is shown in Figure 3.4.

ILC-DPD first find the optimal input of PA with respect to the desired output, which is unknown in other DPD structures such as ILA. Once the input and output of a DPD are known, it is easy to identify its parameters by some identification models such as GMP (2.8). With the model input $\mathbf{y}_\mathrm{d}$ and the model output $\mathbf{u}_\mathrm{opt}$, the GMP forms the DPD relation as

$$\begin{aligned} u_\mathrm{opt}(n) = &\sum_{k=1}^{K}\sum_{m=0}^{M} a_{km} y_\mathrm{d}(n-m)|y_\mathrm{d}(n-m)|^{k-1} \\ &+ \sum_{k=2}^{K}\sum_{m=0}^{M}\sum_{\substack{l=-m \\ l\neq=0}}^{L} b_{kml} y_\mathrm{d}(n-m)|y_\mathrm{d}(n-m-l)|^{k-1} \end{aligned} \tag{3.15}$$

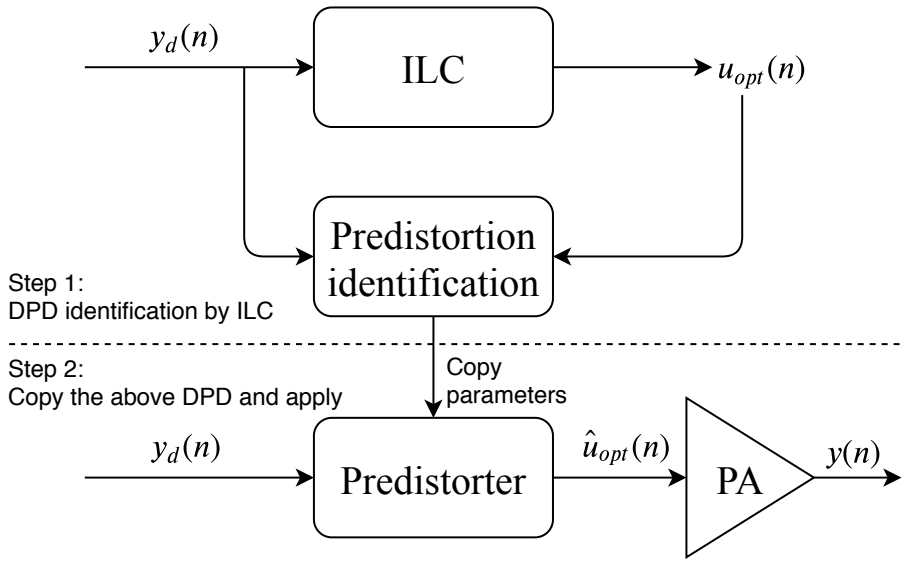


Figure 3.4: The structrue of ILC-DPD.

where K and M are nonlinear order and memory depth, respectively. The vector form of (3.15) is expressed as $\mathbf{u} = \mathbf{Yw}$, which can be solved by least squares algorithm. Once the parameters $\mathbf{w}$ are determined, a repetition of that DPD model is used as the desired predistorter. Compared to ILA, ILC-DPD can achieve better linearization [12].

3.2 Crest factor reduction

One major drawback of OFDM is the high peak-to-average power ratio, which indicates the extent of peak power in a signal. An alternative of PAPR is CF, which is regularly used in this thesis. It is defined as the ratio of peak amplitude to root mean squares (RMS) amplitude of the signal $x(n)$.

$$\mathrm{CF} = \frac{\|x\|_\infty}{\|x\|_2} \tag{3.16}$$

where $\|\cdot\|_\infty$ is the l_∞-norm of signal $x(n)$, namely its peak value, and $\|\cdot\|_2$ is the l_2-norm of x_n, namely its RMS value. In the PA operation, high peaks of the signal first enter into the compression region and experience more severe distortion, shown in Figure 3.5. The gain loss and signal distortion are more severe for signals with higher CF. To address this issue, one solution is to back-off the input signal power, but this solution results in PA efficiency reduction. Another solution is to decrease the CF of input signals so that the signal peak x_{peak} is pulled back to the linear region of PA, which is called the CFR. Although there are various CFR algorithms, only iterative clipping and filtering is described in the following section due to its simple implementation.

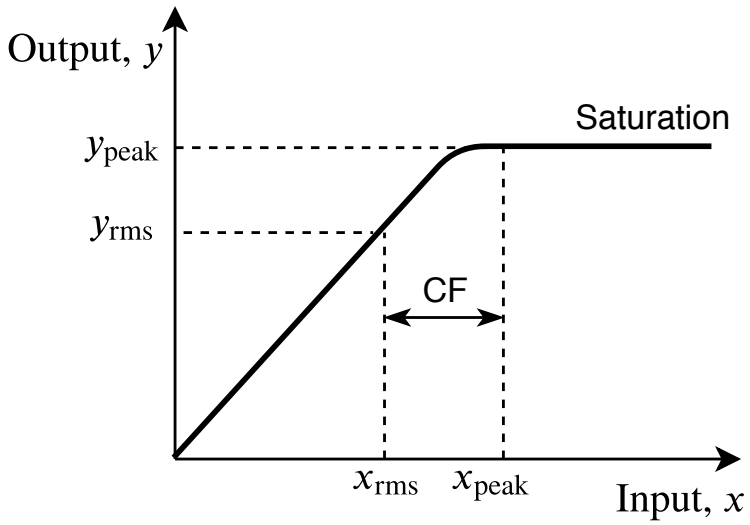


Figure 3.5: High crest-factor affects PA operation.

Iterative clipping and filtering

ICF is also known as amplitude clipping and filtering. The envelope signal y is first clipped by a limiter:

$$y_{\text{c}} = \begin{cases} y & \text{if } |y| \leq A \\ Ae^{j\angle y} & \text{if } |y| > A \end{cases} \tag{3.17}$$

where $\angle y$ denotes the phase of y. After clipping, not only the in-band distortion is produced, but the out-of-band distortion is also created. Compared with in-band distortion, the out-of-band distortion can be easily filtered out. The filtering process reproduces peaks, so the signal needs to be clipped again. The whole clipping and filtering procedure usually require several iterations until the desired CF is satisfied. Figure **??** Section 1.1 shows more details about the behavior of ICF.

4

CFR-free digital predistortion algorithm

4.1 Motivation of CFR-free

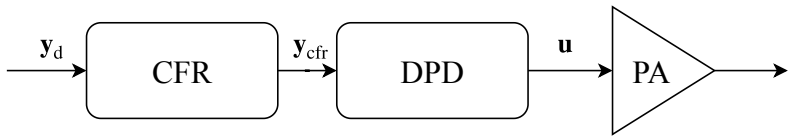
To reduce the crest factor of signal, CFR reduces the amplitude of the signal, while DPD increases the signal amplitude to restrain the compression effect of PA. Intuitively, CFR and DPD perform inverse operations on the signal, so putting CFR before DPD seems to be a redundant choice. To further observe these two inverse operations, we analyze the behaviour of CFR-DPD in Figure 4.1, where Figure 4.1a shows the block diagram of CFR-DPD and 4.1b compares different signals in CFR-DPD.

In Figure 4.1, the power gain of PA is 1 for the convenience of analysis, and the desired output signal $\mathbf{y}_\mathrm{d}$ is equal to the input signal. After CFR, the amplitudes of $\mathbf{y}_\mathrm{d}$ are reduced such that the output signal $\mathbf{y}_\mathrm{cfr}$ meet the desired crest factor requirement. This operation is also visible in Figure 4.1b (green curve becomes red curve). Then, DPD increases the amplitude of $\mathbf{y}_\mathrm{cfr}$ to $\mathbf{u}$ (blue curve) such that the PA output $\mathbf{y}$ approaches the DPD input $\mathbf{y}_\mathrm{cfr}$. In this situation, the DPD aims to minimize the error of the PA output towards $\mathbf{y}_\mathrm{cfr}$ instead of $\mathbf{y}_\mathrm{d}$. In Figure 4.1b, the PA output $\mathbf{y}$ matches $\mathbf{y}_\mathrm{cfr}$ properly. The distortion in $\mathbf{y}_\mathrm{cfr}$ restrains the performance of CFR-DPD heavily.

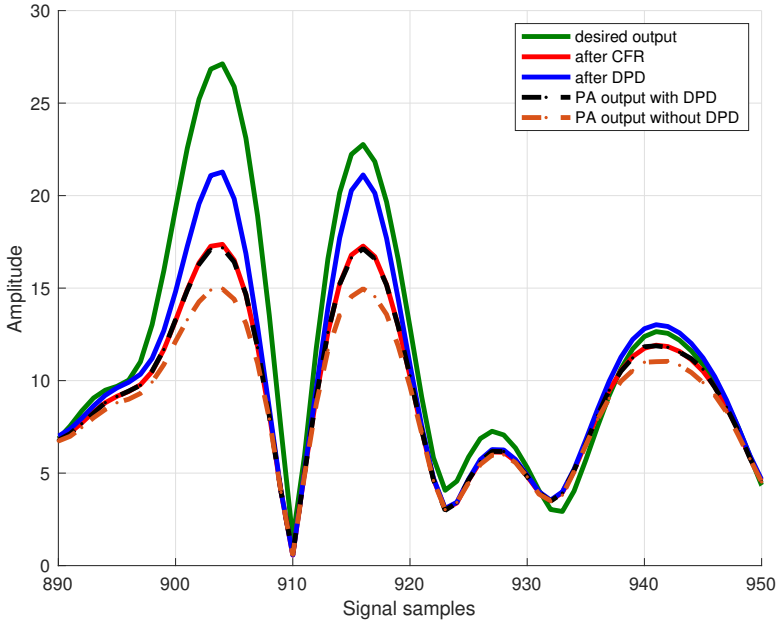
These observations indicate that CFR affects the optimal output of PA, which eventually limits the linearization performance. Thus we get a motivation to remove CFR. In the next Section, a CFR-free DPD structure is proposed, and similarly, analyze the signal behaviours of the proposed system are analyzed.

4.2 The proposed CFR-free DPD

The CFR-free DPD structure is proposed, and its block diagram is shown in Figure 4.2a. Based on this structure, the behaviours of different signals are illustrated in Figure 4.2b. Compared with the CFR-DPD structure, no CFR is implemented before DPD, so the system input $\mathbf{y}_\mathrm{d}$ directly goes into the DPD operation. After that, the amplitudes of DPD output $\mathbf{u}$ are expanded greatly which may be extremely high and be physically harmful to PA safety. To protect the PA, the limiter truncates those extreme peaks in $\mathbf{u}$, which makes $\mathbf{u}_\mathrm{lim}$. The selection of the limiter's threshold



(a) System model of CFR-DPD.



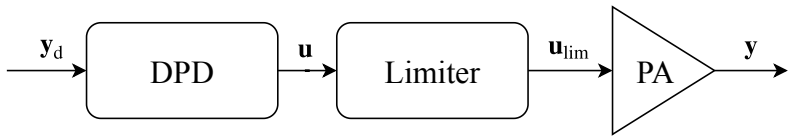
(b) Signals after each block of CFR-DPD.

Figure 4.1: Inverse behaviours of CFR and DPD in CFR-DPD structure.

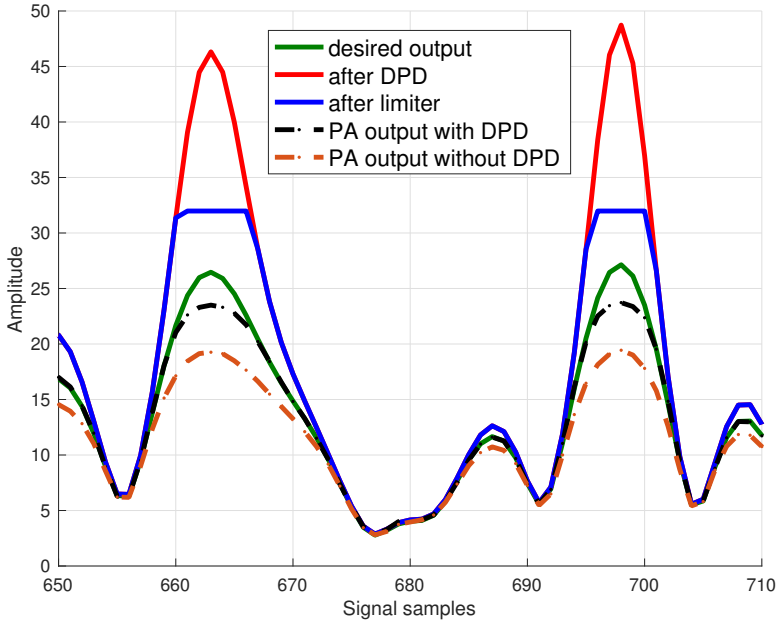
depends on different requirements. To yeild signal quality loss, the threshold should be as high as possible so long as the PA safety is guaranteed. The limiter's threshold is set to 40 in Figure 4.2b. Given that the limiter's threshold is high enough, the optimal output of PA is equal to the DPD's input, which is the desired signal $\mathbf{y}_d$ instead of $\mathbf{y}_{cfr}$ in CFR-DPD structure. Compared to Figure 4.1b, the PA's output $\mathbf{y}$ in Figure 4.2b is close to the desired output $\mathbf{y}_d$ instead of the biased signal $\mathbf{y}_{cfr}$. The difference between CFR-DPD and CFR-free DPD is further quantified and analyzed in the result section.

4.3 Evaluation strategies

To have a clear comparison between the proposed CFR-free DPD and CFR-DPD, we follow an evaluation plan that includes PA models, DPD learning structures, and performance metrics. Two types of PA models are considered, namely PAs with and without memory effects. Two different memoryless PAs can be built in MATLAB, namely the ideal PA and the tanh PA. The PA with memory effects is implemented in reality, namely a Class-AB PA. The introduction of those PAs is presented in Section 4.3.1. Then, we consider different learning architectures for DPD, which



(a) System model of CFR-free DPD.



(b) Signals after each block of CFR-free DPD.

Figure 4.2: Behaviours of DPD and limiter in CFR-free DPD structure.

include ILA,ILC, and ILC-DPD. Finally, we select several performance metrics in Section 4.3.2.

4.3.1 Selection of PA models and DPD learning architectures

An ideal PA model

The simplest memoryless PA model is the ideal PA, whose input-output relation is defined as

$$\frac{y_{\text{out}}}{G} = \begin{cases} y_{\text{in}} & \text{if } |y_{\text{in}}| \leq A \\ A\ \exp(j\angle y_{\text{in}}) & \text{if } |y_{\text{in}}| > A \end{cases} \tag{4.1}$$

where G denotes the gain in the linear region, $\frac{y_{\text{out}}}{G}$ denotes the normalized output, A is the saturation point of PA, $\angle y_{\text{in}}$ denotes the phase of input signal y_{in}. Based on (4.1), the response behaviour is shown in Figure 4.3. The amplitude of output is ideally amplified in the linear region while the gain is always zero in the compression region. Because this ideal model represents an extreme case of PA, it is the best situation that a DPD can reach. So no more DPD can perform better performance

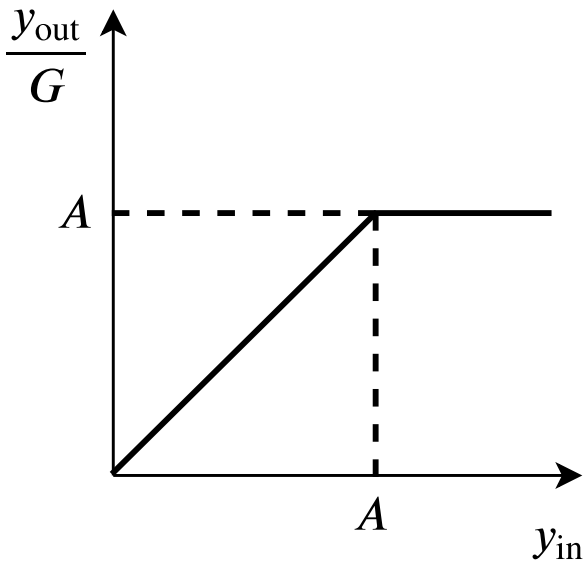


Figure 4.3: Behaviour of the ideal PA.

than this situation, it can be used to build a performance bound. A signal error lower bound has been studied in [34], and we introduce it in Section 4.3.2.

The tanh PA model

Another memoryless PA selected in this thesis is a nonlinear model whose nonlinear characteristic is represented by the positive part of the hyperbolic tangent function. As this model can be easily implemented in MATLAB by using the tanh function, we naturally call this model as a tanh PA model. This model is defined as

$$\frac{y_{\text{out}}}{G} = A \tanh(\frac{|y_{\text{in}}|}{A}) \exp(j\angle y_{\text{in}}) \tag{4.2}$$

where G denotes the PA gain, $\frac{y_{\text{out}}}{G}$ denotes the normalized output, A denotes the saturation point, and tanh is the hyperbolic tangent function. An example of the input-output response of a tanh model is shown in Figure 4.4, where the saturation point is 10.

A Class AB PA from the RF WebLab

Since the memory effect is another important factor of PA nonlinearity, it cannot be ignored in choosing PA models. Besides, the noise in the actual measuring environment comes from multiple sources such as digital-to-analog converter (DAC) and other thermal facilities, so it is complex to build a PA model considering all these factors. Hence, we consider a class AB PA, which is configured in the RF WebLab [13]. The RF WebLab provides a PA measurement setup which includes some instruments such as signal generator and signal analyzer. The block diagram of the RF WebLab setup is shown in Figure 4.5. In the MATLAB block, the signal is possibly processed by CFR, DPD, and limiter, and the processed signal is sent to

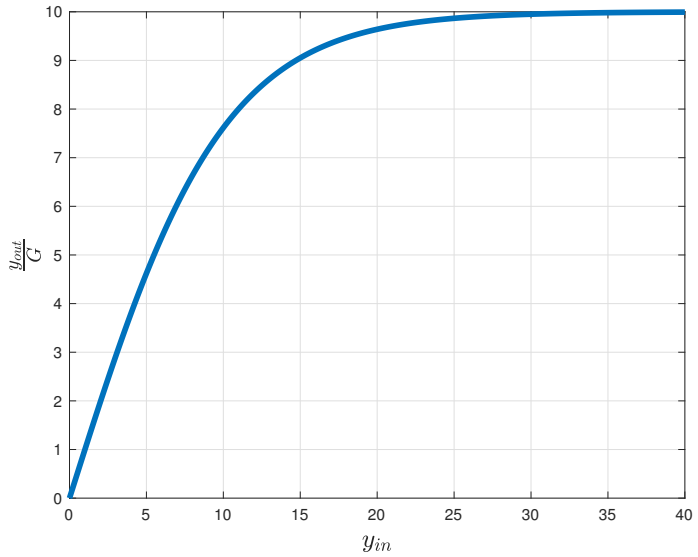


Figure 4.4: Behaviour of a tanh PA model with a saturation point 10.

the RF WebLab, which can be accessed remotely. In the WebLab block, the input signal from MATLAB is generated by the transmitter of vector signal transceiver (VST) (PXIe-5646R VST). Then, after a linear driver amplifier with a 40 dB gain, the signal is passed into the target PA (Cree CGH40006-TB), and through a 30 dB attenuator, the output signal is eventually measured by the VST receiver. The estimation of setup parameters (measurement noise and saturation point) are shown in Section 5.2.1.

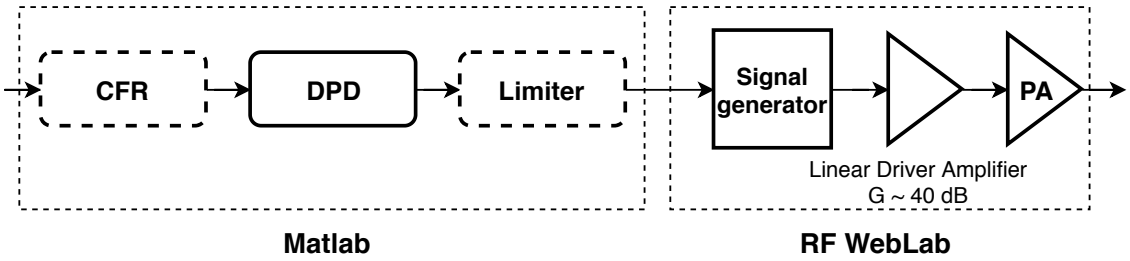


Figure 4.5: Block diagram of the experiment on RF WebLab. The blocks of CFR and limiter are in dashed lines since the structure of CFR-DPD and CFR-free DPD are selected in different cases.

4.3.2 Selection of metrics

The performance of different DPD methods can be evaluated by comparing the input/output signal. There are some evaluation metrics which focus on a different part of the signal, shown in Figure 4.6, where α denotes the evaluation focus. If $\alpha = 0$, the evaluation focus is the out-of-band signal quality, which can be evaluated by the metric ACPR. If $\alpha = 1$, only the in-band distortion is evaluated, where the metric error vector magnitude (EVM) is usually used. If we treat the in-band and out-of-band distortion equally, namely $\alpha = 0.5$, the overall distortion is the

evaluation goal, and the metric NMSE can be utilized. However, because the in-band distortion level usually dominate the overall distortion level, NMSE provides similar results as EVM. Also the implementation of NMSE is simple, so we choose NMSE as the metric.

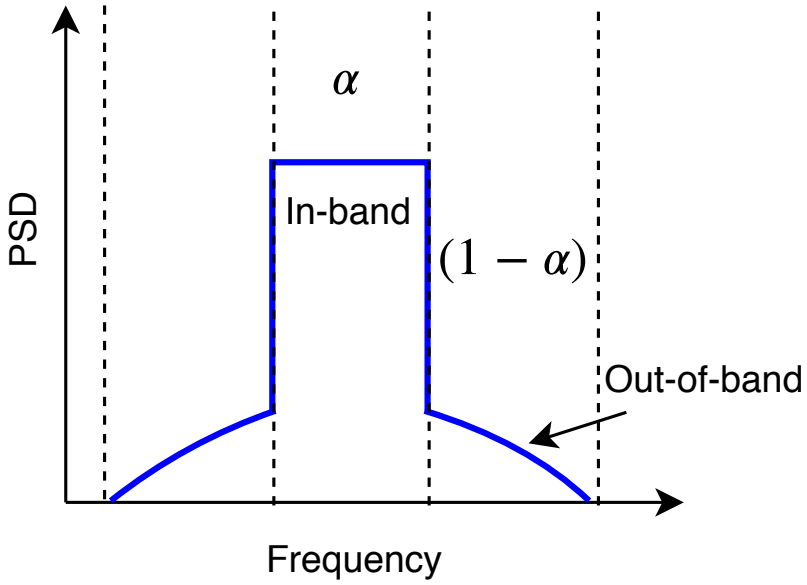


Figure 4.6: Different evaluations focus on different parts of the signal. NMSE evaluates the all-band distortion when $\alpha = 0.5$.

Evaluate all-band distortion by the NMSE

When the $\alpha = \frac{1}{2}$ in Figure 4.6, the overall signal quality is our object. NMSE is widely used to quantify the overall deviations between input and output signals. Figure 4.7 shows the block diagram of a DPD-PA system, where the measurement noise w is assumed to be a Gaussian white noise with variance $2\sigma^2$. NMSE is defined

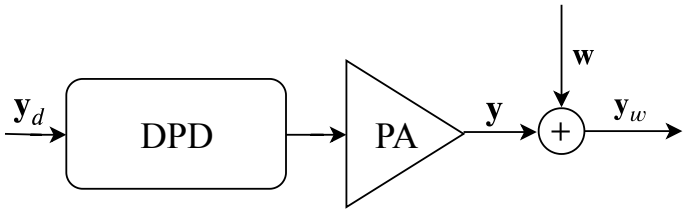


Figure 4.7: System model of DPD-PA.

as the square error between the measured signal $\mathbf{y}_w$ and the desired signal $\mathbf{y}_\mathrm{d}$. Then the summation of this error is then normalized by the power of $\mathbf{y}_\mathrm{d}$. The expression can be written as [12]

$$\mathrm{NMSE} = \frac{\mathrm{var}[\mathbf{y}_\mathrm{d} - \mathbf{y}_w]}{\mathrm{var}[\mathbf{y}_\mathrm{d}]} \tag{4.3}$$

We have mentioned in Section 4.3.1 that the input/output relation of ideal PA is the best situation that a DPD method can do. In this situation, DPD compensates perfectly so that the output is completely linear to the input before the saturation point. Also, the NMSE is the lowest. The author in [34] derived the NMSE lower bound, expressed as [34, Equation (10)]

$$\text{NMSE} = \frac{2\sigma^2 \exp(\frac{-A^2}{2\sigma^2}) - 2A\sigma\sqrt{2\pi}Q(\frac{A}{\sigma}) + \sigma_w^2}{2\sigma^2} \tag{4.4}$$

where A, σ_w^2, $2\sigma^2$ and $Q(\cdot)$ denote the saturation point, the measurement noise variance, the power of $\mathbf{y}_\text{d}$ and the Q-function. This NMSE lower bound is verified in Section 5.1, and it can be used as an excellent performance reference for comparing different DPD algorithms.

5

Results and discussion

5.1 Simulation results on PAs without memory

Results on an ideal PA model

First, the closed-form NMSE expression (4.4) is verified by simulations on the ideal PA. The setup parameters are shown in Table 5.1. Note that because the input signal is ideally clipped when its amplitude is larger than the saturation point, any limiter's threshold larger than PA saturation makes no difference. Here, the threshold is equal to the PA saturation point. The NMSE results are shown in Figure 5.1. The simulated NMSE results follow the theoretical NMSE lower bound properly. The lower bound can be viewed as a reference of the achievable performance.

Then, based on the setup parameters, the NMSE results are shown in 5.2, where we compare the CFR-free DPD with CFR-DPD. Parameters of this simulation are illustrated in Table 5.2. Notice that ILA is chosen to be the DPD learning architecture, and the DPD parameters are identified by a memoryless polynomial model of order 3. The NMSE gaps between the CFR-DPD and the CFR-free DPD are noticeable, especially in the high output power region. CFR-free cases have significant advantage to have lower NMSE in that power region. Here, ILA cannot offer any benefit because the compensation of DPD that exceeds the PA bound is ideally clipped. The limiter also has the same situation when its threshold goes beyond the PA saturation. Although we do not present an theoretical NMSE lower bound for case with CFR, the result of case with CFR only can be viewed as a performance reference.

Table 5.1: Setup parameters of the simulations on PAs without memory.

Parameters	Value	Unit
Saturation	26.5	V
Noise variance	1.2	mV
Impedance	50	Ω
Signal length	10^7	Samples
Sampling rate	200	MHz
Bandwidth	20	MHz

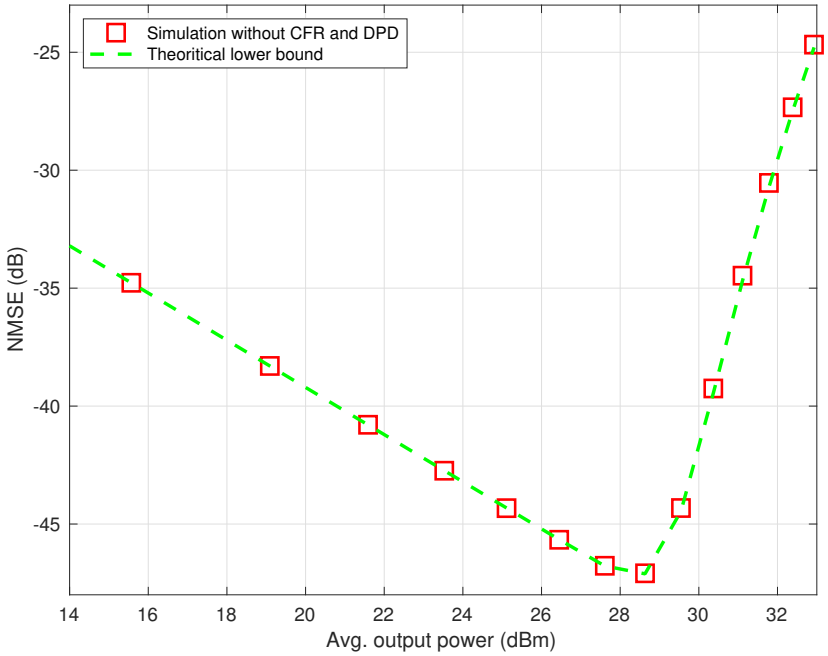


Figure 5.1: NMSE lower bound verification. Simulations implemented on an ideal PA.

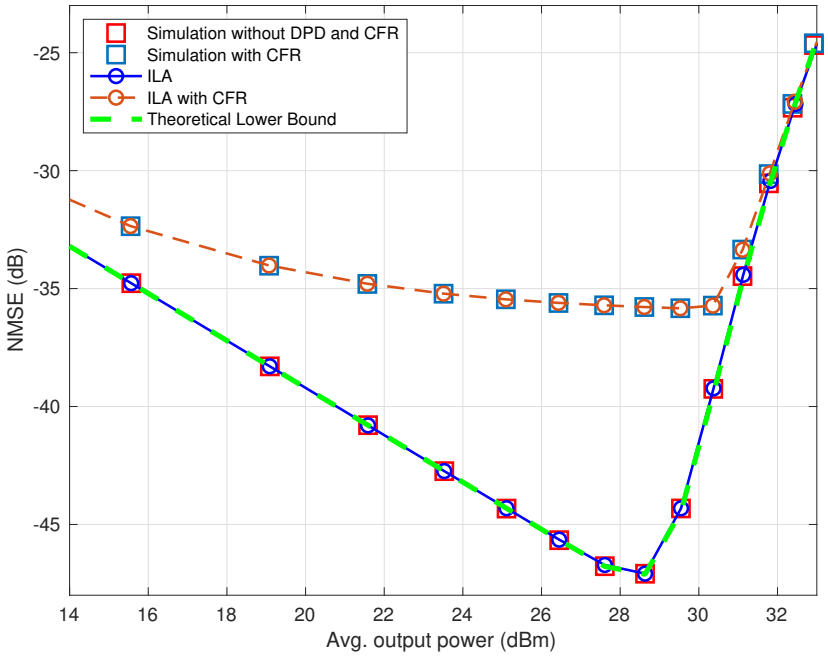


Figure 5.2: NMSE evaluations between cases with and without CFR. Simulations implemented on an ideal PA.

Table 5.2: The environmental parameters for Figure 5.2.

Parameters	Value	Unit
Limiter's threshold	$1 * 26.5$	V
PAPR before CFR	12.3	dB
PAPR after CFR	9.02	dB
Memory depth of MP model	0	-
Nonlinear order of MP model	3	-

Results on a tanh PA model

The tanh PA model is introduced in Section 4.3.1. Note that we set the gain of the PA to 1 for computation convenience. Based on this model the performance of CFR-DPD and CFR-free DPD are compared. Some scenario parameters are the same as in Table 5.1, and other parameters are shown in Table 5.3. In addition to ILA, ILC and ILC-DPD are added as other DPD learning architectures. In ILA and ILC-DPD, the DPD parameter identifications are handled by a memoryless polynomial model with nonlinear order 3. For PAPR reduction purpose, the limiter's threshold is selected to 1.6*26.5 V, which means the amplitude of the PA input is limited at that threshold. The NMSE results are shown in Figure 5.3, where dotted lines are CFR-DPD cases and solid lines are CFR-free DPD cases.

Table 5.3: The setup parameters for Figure 5.3.

Parameters	Value	Unit
Limiter's threshold	$1.6 * 26.5$	V
PAPR before CFR	12.3	dB
PAPR after CFR	8.89	dB
ILC learning algorithm	Linear	-
ILC learning rate	0.2	-
ILC learning iteration	100	times

Notice the poor performance of cases without DPD (red square solid line and blue square dotted line). Because of no DPD compensation, these two cases experience strong compression effects, which results in quick NMSE degradation when the average output power is larger than 19 dBm. In contrast, the NMSE degradation of cases with DPD appeared later due to the DPD's compensation. The results of CFR-free DPD cases (ILA, ILC, and ILC-DPD) reach the NMSE lower bound closely before the power level 26.5 dBm. After that, the performance of ILA is the first one to degrade with a large rate due to its model accuracy limitation. Although ILC has the best performance, the computation complexity is also the highest. ILC-DPD greatly decrease the complexity of ILC as well as better performance than ILA. The

results of CFR-free DPD do not reach the NMSE lower bound for high output power region (namely 28 dBm to 31 dBm) because of the signal distortion caused by the limiter, but the loss is acceptable compare to the large NMSE gaps causae by using CFR.

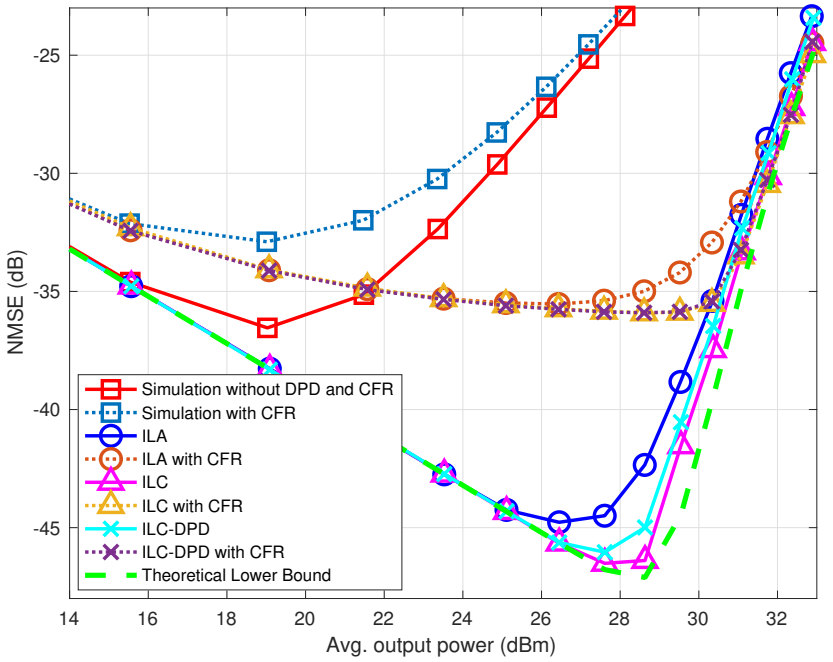


Figure 5.3: NMSE against the average output power. Simulations implemented on the tanh PA.

Comparing the NMSE results of CFR-DPD cases with CFR-free DPD cases, remarkable NMSE gaps can be viewed. This gap mainly comes from the signal distortion produced by the CFR operation. As we have mentioned in Section 4.1, CFR generates a distorted signal, which is the DPD's input. As DPD aims to make the PA's output equal to the DPD's input, namely the output of CFR instead of the desired output, the compensation is always biased. The performance gap is small at low output power level due to noise-limitation, but it became wider in higher power region, namely 26.5 dBm to 30 dBm. This region attracts our interest since our purpose is to have higher output power as well as low signal distortion. Hence, using CFR eventually increases signal distortion level compared to cases without CFR, which means CFR-free DPD has better linearization performance with respect to signal distortion.

Although the limiter has a negative effect on the linearization performance, it aims to prevent the DPD avalanche by truncating the extremely high peaks of the DPD output signal. An intuitive way to assess the limiter's performance is to evaluate the PAPR of the PA's signal. Based on the setup parameters in Table 5.1 and 5.3, PAPR results against the average output power are shown in Figure 5.4.
At the beginning, cases with CFR have lower PAPR (8.21 dB) than CFR-free cases

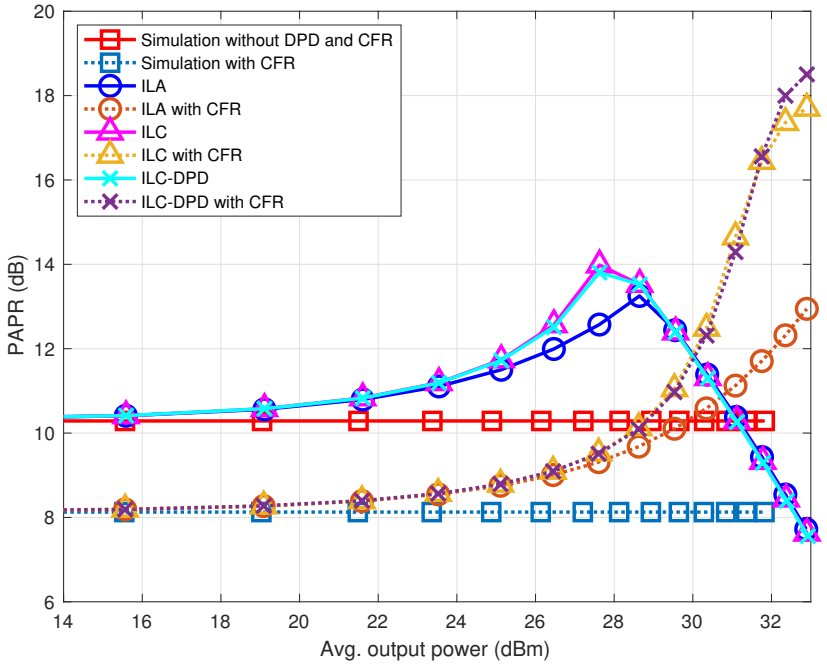


Figure 5.4: PAPR of the PA's input signal against the average output power. Simulations implemented on the tanh PA model.

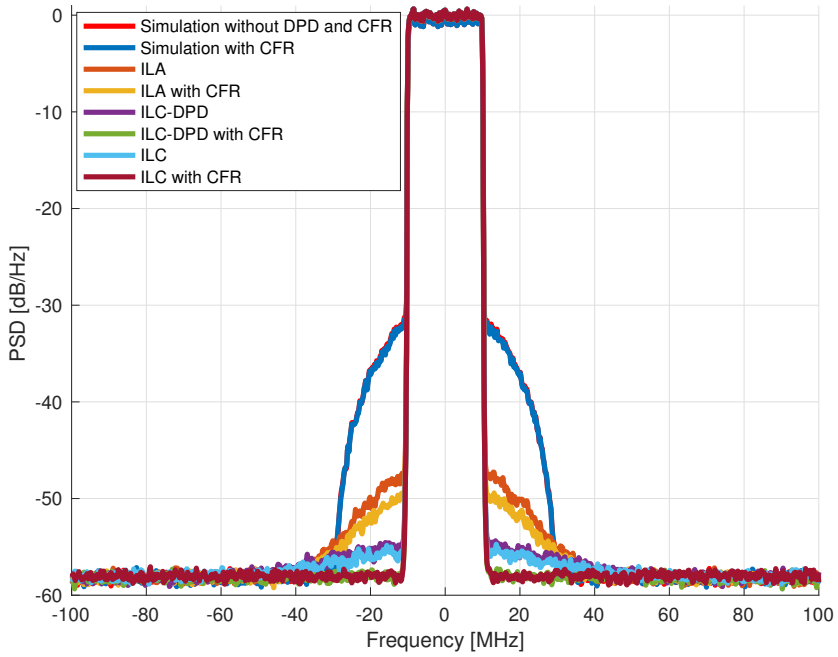


Figure 5.5: PSD of the PA's output signal at the average output power level 30.28 dBm. Simulations implemented on the tanh PA model.

(10.29 dB) due to the CFR operation. Notice that cases without DPD (lines with square markers) have stable PAPR values at all power level. As the output power increases, the PAPR of cases using DPD increase quickly because the signal's peaks start to obtain compensation from DPD. For CFR-free DPD cases, this increase ends at power level 27 dB, where the limiter starts to truncate signals with amplitude higher than the limiter threshold ($1.6 * 25.6$ V). And this threshold makes rapid reductions of the PAPR from around 14 dB to 10.5 dB as the output power level goes from 27.6 dBm to 31 dBm. However, because of no amplitude limitation, the results of CFR-based cases keep their growths to extremely high levels where DPD add more and more power to the signal peaks, and in turn, these extreme peaks experience more severe compression. Meanwhile, these extreme peaks are not hardware-friendly, which may harm the PA. These results indicate that the limiter makes great peak restriction to the DPD's output signal.

Another way to assess the signal quality is to evaluate the PSD of the output signal. Figure 5.5 shows the PSD of the PA outputs for different cases at an average output power level 30.28 dBm. The NMSE and PAPR of each case at this power level can be found in Figure 5.3 and 5.4, respectively. The PSD results make it easy to observe the out-of-band distortion level. Cases without DPD have severe out-of-band distortion, which is significantly reduced by the ILA. And the distortion level is even lower for ILC and ILC-DPD, which indicates that they have better PA linearization performance. While cases with CFR have out-of-band distortion advantages over cases without CFR, these advantages are neglectable compared to the large NMSE loss.

5.2 Experimental results on a PA with memory effects

In practice, memory effects become more prominent as the signal bandwidth goes up due to the PA's frequency selectivity. In this section we measure experimental results from a PA in reality, where the performance of CFR-DPD and CFR-free DPD is compared. First, the parameters of the experimental environment is calculated in Section 5.2.1. Then, the results are presented in Section 5.2.2.

5.2.1 Parameter initialization of the RF WebLab

The setup of RF WebLab is introduced in Section 4.3.1. This section describe the identification of the PA parameters which includes the saturation point and the measurement noise variance. Theses parameters are used in (4.4) to calculate the NMSE lower bound. The input OFDM signal with the same sampling rate (200 MHz) and bandwidth (20 MHz) but shorter length (10^6) is driven to the RF WebLab, and Figure 5.6 shows the input/output amplitude relation in Volts. The measurement noise variance was calculated at a low input power level where the PA works in its linear region. Based on the PA input and output, a linear model is computed, and the noise is estimated by the difference between the PA output and

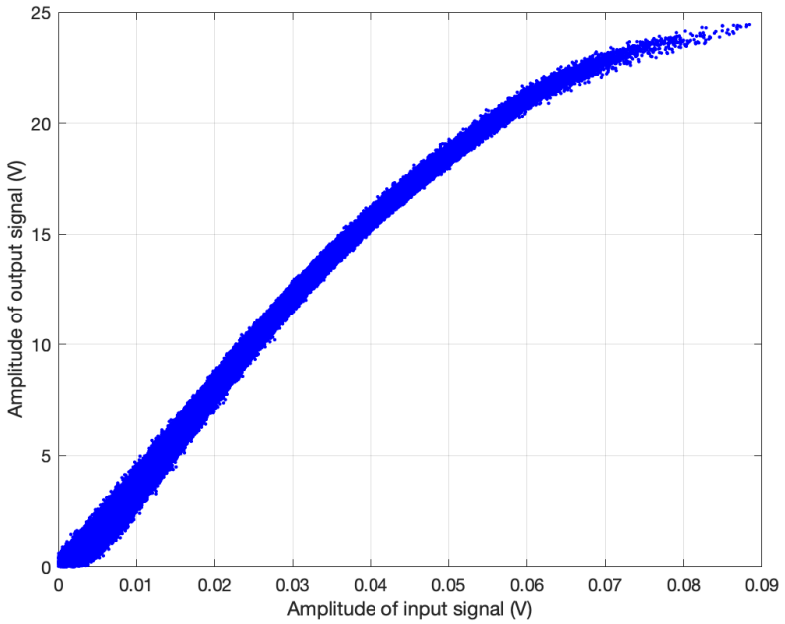


Figure 5.6: Input/Output amplitude relation of the RF WebLab.

the linear model output. Estimations of the noise variance and the saturation point are 0.0032 and 24.6 V, respectively.

5.2.2 Results on the RF WebLab

The setup parameters of the simulations on RF WebLab are shown in Table 5.4. Considering memory effects and severe nonlinearity of this class-AB PA, simulations utilize a MP model with memory depth 3 and nonlinear order 5. The learning algorithm of ILC is the gain-based type. The limiter's threshold is set to 0.089 V because the power of the input signal to RF WebLab is restrained to -11 dBm ($\approx$ 0.089 V).

The NMSE evaluation of different cases are shown in Figure 5.7. Notice the performance gaps between CFR-free DPD cases and CFR-DPD cases. As we have indicated in Section 4.1, the optimal output of PA in cases with CFR is the output of CFR (y_{cfr}) instead of the desired output (y_{d}), so eventually, the error between y_{cfr} and y_{d} causes that performance gaps. ILC performs the best compensation where the NMSE results are closest to the lower bound. Compared to the modeling process of tanh PA, the accuracy of MP model is limited in modeling the Class-AB PA due to the severe nonlinearity and memory effects. This limitation results in limited performances of ILA and ILC-DPD at high output power region, although the NMSE values are still greatly reduced compare to the cases without DPD. We also evaluate the PAPR of the PA's input signal and the PSD of the PA's output signal in Figure 5.8 and Figure 5.9, respectively. Compared to the results in the tanh PA, the PAPR results from RF WebLab followed similar trends, and the PSD results show that ILA and ILC-DPD have higher out-of-band distortion levels at a

high output power level.

Table 5.4: Setup parameters of simulations on the RF WebLab.

Parameters	Value	Unit
Saturation	24.6	V
Noise variance	0.0032	V
Impedance	50	Ω
Signal length	10^6	Samples
Sampling rate	200	MHz
Bandwidth	20	MHz
Limiter's threshold	0.089	V
PAPR before CFR	10.63	dB
PAPR after CFR	8.84	dB
Memory depth of MP model	3	-
Nonlinear order of MP model	5	-
ILC learning algorithm	Gain-based	-
ILC learning iteration	20	times

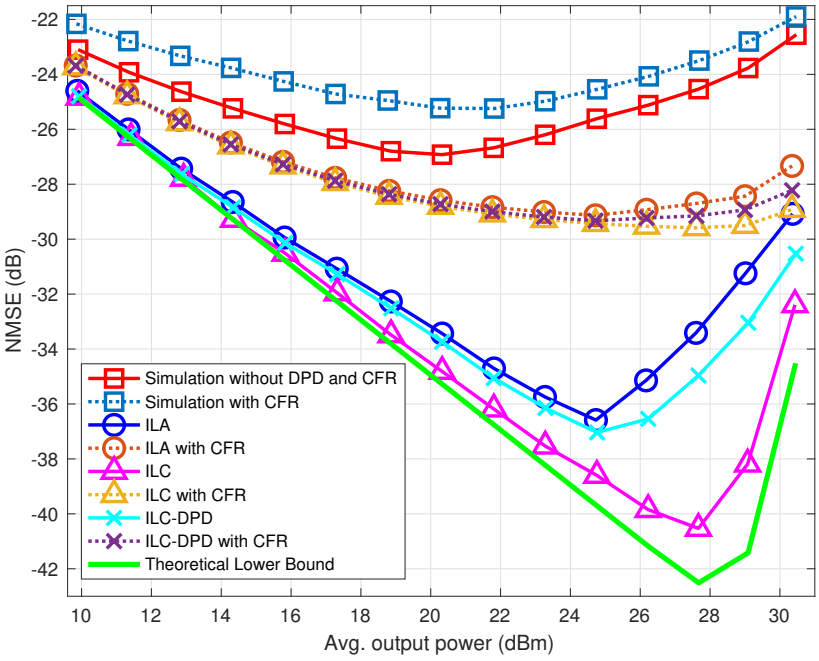


Figure 5.7: NMSE against the average output power. Simulations implemented on the RF WebLab.

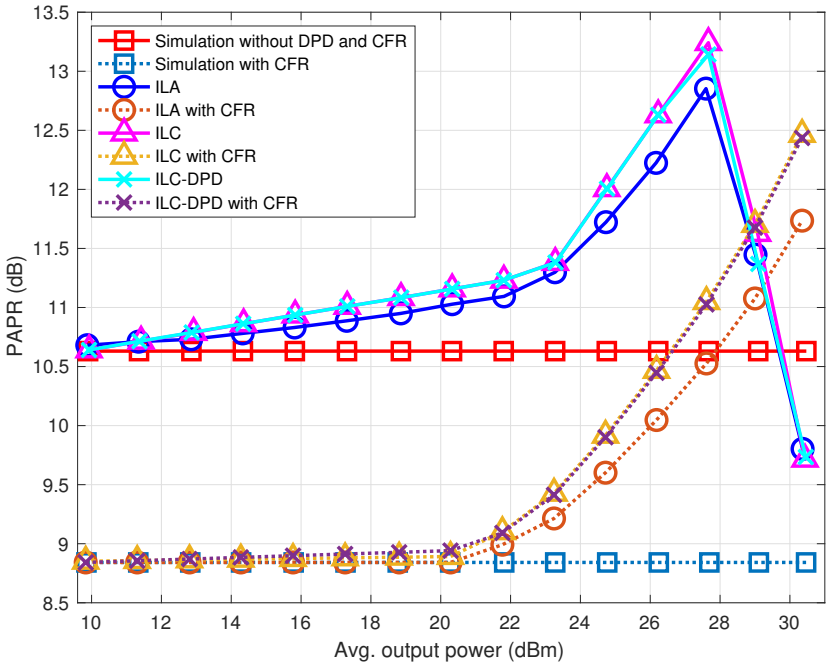


Figure 5.8: PAPR of the PA's input signal against the average output power. Simulations implemented on the RF WebLab.

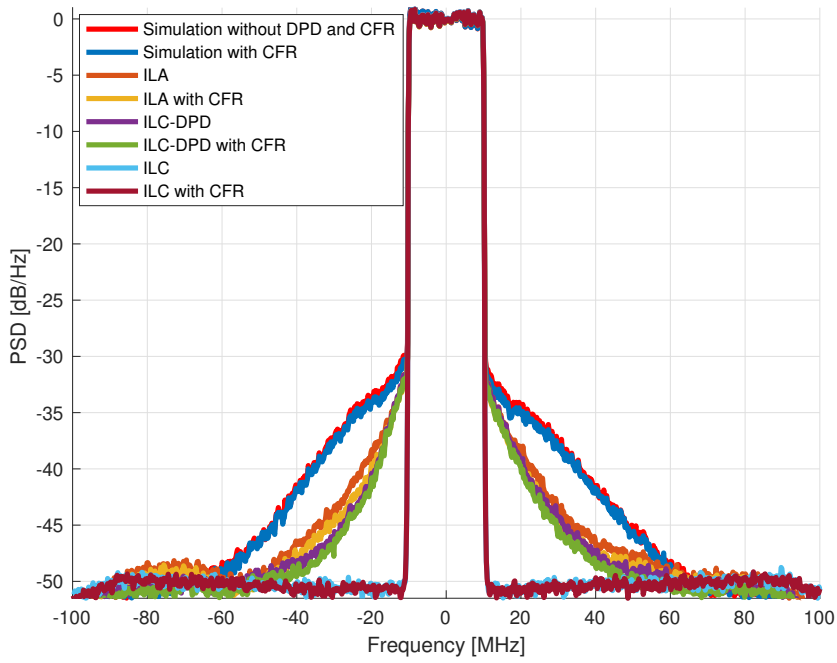


Figure 5.9: PSD of the PA's output signals at the average output power 27.59 dBm. Simulations implemented on the RF WebLab.

5.3 Conclusion

This thesis proposed a CFR-free DPD structure for PA linearization. Unlike the situation in CFR-DPD that the output of PA can only reach the output of CFR, the output of CFR-free DPD can reach the desired output, which is unbiased. We compared the performance of those two structures by evaluating the NMSE, PAPR and PSD on different PA models. Results showed that CFR-free DPD had better PA linearization performance with respect to the output distortion. The limiter in CFR-free DPD can protect PA safety as well as reduce the crest factor of the PA's input signal significantly. It was found that the amount of distortion caused by CFR has some relation to the NMSE results, which could be the future work for investigating the NMSE lower bound of CFR-DPD.

Bibliography

[1] S. H. Han and J. H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," *IEEE wireless communications*, vol. 12, no. 2, pp. 56–65, 2005.

[2] I. Aoki, S. D. Kee, D. B. Rutledge, and A. Hajimiri, "Fully integrated cmos power amplifier design using the distributed active-transformer architecture," *IEEE Journal of Solid-State Circuits*, vol. 37, no. 3, pp. 371–383, 2002.

[3] C. Nader, P. N. Landin, W. Van Moer, N. Bjorsell, P. Handel, and D. Ronnow, "Peak-power controlling technique for enhancing digital pre-distortion of rf power amplifiers," *IEEE transactions on microwave theory and techniques*, vol. 60, no. 11, pp. 3571–3581, 2012.

[4] X. Li and L. J. Cimini, "Effects of clipping and filtering on the performance of ofdm," in *1997 IEEE 47th Vehicular Technology Conference. Technology in Motion*, vol. 3. IEEE, 1997, pp. 1634–1638.

[5] J. Armstrong, "Peak-to-average power reduction for ofdm by repeated clipping and frequency domain filtering," *Electronics letters*, vol. 38, no. 5, pp. 246–247, 2002.

[6] A. E. Jones, T. A. Wilkinson, and S. Barton, "Block coding scheme for reduction of peak to mean envelope power ratio of multicarrier transmission schemes," *Electronics letters*, vol. 30, no. 25, pp. 2098–2099, 1994.

[7] A. Jayalath and C. Tellambura, "Reducing the peak-to-average power ratio of orthogonal frequency division multiplexing signal through bit or symbol interleaving," *Electronics Letters*, vol. 36, no. 13, pp. 1161–1163, 2000.

[8] R. W. Bauml, R. F. Fischer, and J. B. Huber, "Reducing the peak-to-average power ratio of multicarrier modulation by selected mapping," *Electronics letters*, vol. 32, no. 22, pp. 2056–2057, 1996.

[9] C. Nader, P. N. Landin, W. Van Moer, N. Bjorsell, and P. Handel, "Performance evaluation of peak-to-average power ratio reduction and digital pre-distortion for ofdm based systems," *IEEE transactions on microwave theory and techniques*, vol. 59, no. 12, pp. 3504–3511, 2011.

[10] C. Nader, P. N. Landin, W. Van Moer, N. Björsell, P. Händel, and M. Isaksson, "Peak-to-average power ratio reduction versus digital pre-distortion in ofdm based systems," in *2011 IEEE MTT-S International Microwave Symposium*. IEEE, 2011, pp. 1–4.

[11] C. Eun and E. J. Powers, "A new volterra predistorter based on the indirect learning architecture," *IEEE transactions on signal processing*, vol. 45, no. 1, pp. 223–227, 1997.

[12] J. Chani-Cahuana, P. N. Landin, C. Fager, and T. Eriksson, "Iterative learning control for rf power amplifier linearization," *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 9, pp. 2778–2789, 2016.

[13] P. N. Landin, S. Gustafsson, C. Fager, and T. Eriksson, "Weblab: A web-based setup for pa digital predistortion and characterization [application notes]," *IEEE Microwave Magazine*, vol. 16, no. 1, pp. 138–140, 2015.

[14] H. Ku and J. S. Kenney, "Behavioral modeling of nonlinear rf power amplifiers considering memory effects," *IEEE transactions on microwave theory and techniques*, vol. 51, no. 12, pp. 2495–2504, 2003.

[15] D. Bondar, "Advanced digital predistortion of power amplifiers for mobile and wireless communications," Ph.D. dissertation, University of Westminster, 2009.

[16] J. C. Pedro and S. A. Maas, "A comparative overview of microwave and wireless power-amplifier behavioral modeling approaches," *IEEE transactions on microwave theory and techniques*, vol. 53, no. 4, pp. 1150–1163, 2005.

[17] L. C. Nunes, P. M. Cabral, and J. C. Pedro, "Am/pm distortion in gan doherty power amplifiers," in *2014 IEEE MTT-S International Microwave Symposium (IMS2014)*. IEEE, 2014, pp. 1–4.

[18] M. Schetzen, "The volterra and wiener theories of nonlinear systems," 1980.

[19] G. M. Raz and B. D. Van Veen, "Baseband volterra filters for implementing carrier based nonlinearities," *IEEE Transactions on Signal Processing*, vol. 46, no. 1, pp. 103–114, 1998.

[20] J. Kim and K. Konstantinou, "Digital predistortion of wideband signals based on power amplifier model with memory," *Electronics Letters*, vol. 37, no. 23, pp. 1417–1418, 2001.

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

[22] 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 Transactions on signal processing*, vol. 54, no. 10, pp. 3852–3860, 2006.

[23] S. Afsardoost, T. Eriksson, and C. Fager, "Digital predistortion using a vector-switched model," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 4, pp. 1166–1174, 2012.

[24] A. Bjorck, *Numerical methods for least squares problems*. Siam, 1996, vol. 51.

[25] G. Golub and W. Kahan, "Calculating the singular values and pseudo-inverse of a matrix," *Journal of the Society for Industrial and Applied Mathematics, Series B: Numerical Analysis*, vol. 2, no. 2, pp. 205–224, 1965.

[26] J. Chani-Cahuana, *Digital compensation techniques for power amplifiers in radio transmitters*. Chalmers University of Technology, 2017.

[27] L. Ding, G. T. Zhou, D. R. Morgan, Z. Ma, J. S. Kenney, J. Kim, and C. R. Giardina, "Memory polynomial predistorter based on the indirect learning architecture," in *Global Telecommunications Conference, 2002. GLOBECOM'02. IEEE*, vol. 1. IEEE, 2002, pp. 967–971.

[28] J. Chani-Cahuana, C. Fager, and T. Eriksson, "A new variant of the indirect learning architecture for the linearization of power amplifiers," in *2015 10th*

European Microwave Integrated Circuits Conference (EuMIC). IEEE, 2015, pp. 444–447.

[29] M. Schetzen, "Theory of pth-order inverses of nonlinear systems," *IEEE Transactions on Circuits and Systems*, vol. 23, no. 5, pp. 285–291, 1976.

[30] H. Paaso and A. Mammela, "Comparison of direct learning and indirect learning predistortion architectures," in *2008 IEEE International Symposium on Wireless Communication Systems*. IEEE, 2008, pp. 309–313.

[31] D. Zhou and V. E. DeBrunner, "Novel adaptive nonlinear predistorters based on the direct learning algorithm," *IEEE transactions on signal processing*, vol. 55, no. 1, pp. 120–133, 2006.

[32] Y. Chen and C. Wen, *Iterative learning control: convergence, robustness and applications*. Springer, 1999.

[33] K. L. Moore, *Iterative learning control for deterministic systems*. Springer Science & Business Media, 2012.

[34] J. Chani-Cahuana, C. Fager, and T. Eriksson, "Lower bound for the normalized mean square error in power amplifier linearization," *IEEE Microwave and Wireless Components Letters*, vol. 28, no. 5, pp. 425–427, 2018.