Modeling and Mitigation of Nonlinear Distortions by using Neural Networks for LTE-A Wireless Transmitters

Wang, Dongming

Retrieved from https://prism.ucalgary.ca. doi:10.11575/PRISM/26939
http://hdl.handle.net/11023/2775
Downloaded from PRISM Repository, University of Calgary
Modeling and Mitigation of Nonlinear Distortions by using Neural Networks for LTE-A Wireless Transmitters

by

Dongming Wang

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN ELECTRICAL AND COMPUTER ENGINEERING

CALGARY, ALBERTA

JANUARY, 2016

© Dongming Wang 2016
Abstract

A two-box DPD system based on the cascade of a memory predistortion model and a memoryless predistortion model is proposed. The memory predistortion model is designed by using an ARVTDNN, while the memoryless predistortion model is designed by using a memoryless ARVTDNN.

Considering the signal at the output of the PA linearized by the proposed two-box DPD system, one will notice that its ACPR is demonstrated by measurement results to have a better performance by 3-5 dB than that of an existing two-box polynomial-based DPD system. Most importantly, the two-box polynomial-based DPD system does not meet the spectrum emission mask of -45 dBc. In addition, the proposed two-box ARVTDNN-based DPD system meets the ACPR requirements for the observation bandwidth of as low as 155 MHz.
Acknowledgements

I would like to express my profound gratitude to my supervisor, Professor Dr. Fadhel M. Ghannouchi for giving me a chance to study here and his support, encouragement and suggestions throughout my research work. This thesis would not have been possibly completed without his guidance and help.

I also wish to thank all my friends and my colleagues in the iRadio lab group, especially Mohsin Aziz, Andrew Kwan, Xiang Li and Dr. Ramzi Darraji, who helped me a lot in various aspects of my research. To the entire Faculty and Staff of the Electrical Engineering Department at the University of Calgary, thank you for all your help behind the scenes.

Finally, I want to express my appreciation to my family who are always on my side and give me spiritual support while I am working on this program.
# Table of Contents

Abstract ........................................................................................................................................... ii
Acknowledgements ......................................................................................................................... iii
Table of Contents ............................................................................................................................ iv
List of Tables .................................................................................................................................... vi
List of Figures and Illustrations .......................................................................................................... vii
List of Symbols, Abbreviations and Nomenclature .......................................................................... ix

CHAPTER ONE: INTRODUCTION .............................................................................................. 1
  1.1 Introduction ................................................................................................................................ 1
  1.2 Goal and Scope of Work .............................................................................................................. 4

CHAPTER TWO: RF POWER AMPLIFIER CHARACTERISTICS ............................................. 6
  2.1 Introduction ................................................................................................................................ 6
  2.2 RF Power Amplifier Nonlinear Characteristics ......................................................................... 6
    2.2.1 Amplitude and Phase Characteristics .................................................................................... 6
    2.2.2 $P_{1}$ dB Compression Point .................................................................................................. 8
    2.2.3 Third-order Intercept Point .................................................................................................. 8
    2.2.4 Power Back-off .................................................................................................................... 11
    2.2.5 Power Efficiency .................................................................................................................. 12
    2.2.6 Memory Effects ................................................................................................................... 12
  2.3 Performance Evaluation ............................................................................................................. 13
    2.3.1 Peak-to-Average Power Ratio .............................................................................................. 13
    2.3.2 Normalized Mean Square Error ........................................................................................... 13
    2.3.3 Error Vector Magnitude ...................................................................................................... 14
    2.3.4 Adjacent Channel Power Ratio ............................................................................................ 14
  2.4 Conclusion .................................................................................................................................. 15

CHAPTER THREE: LINEARIZATION AND BEHAVIOURAL MODELING OF POWER AMPLIFIERS ........................................................................................................... 16
  3.1 Introduction ................................................................................................................................ 16
  3.2 Linearization Techniques for PA ............................................................................................... 16
    3.2.1 Feedback Linearization Technique ......................................................................................... 16
    3.2.2 Feedforward Linearization Technique ................................................................................... 17
    3.2.3 Linear Amplification with Nonlinear Components ................................................................. 18
    3.2.4 Predistortion Linearization Technique .................................................................................. 19
  3.3 Behavioral Modeling of Power Amplifiers ................................................................................. 21
    3.3.1 Saleh Model .......................................................................................................................... 21
    3.3.2 Look-up Table Model ............................................................................................................ 22
    3.3.3 Volterra Series ..................................................................................................................... 22
    3.3.4 Memoryless and Memory Polynomial Models ....................................................................... 23
    3.3.5 Hammerstein and Wiener Models ....................................................................................... 24
    3.3.6 Neural Network Model ......................................................................................................... 25
  3.4 Conclusion .................................................................................................................................. 31
CHAPTER FOUR: TWO-BOX POLYNOMIAL-BASED DPD SYSTEM AND MEASUREMENT SETUP ......................................................... 32
 4.1 Introduction .................................................................. 32
 4.2 Two-box DPD System .................................................. 32
 4.3 Measurement Setup ...................................................... 36
    4.3.1 Experimental Setup and Data Acquisition .................... 36
    4.3.2 PA/signals used in the experiment .................................. 39
 4.4 Conclusions .................................................................. 39

CHAPTER FIVE: BEHAVIORAL MODELING OF POWER AMPLIFIERS BY USING AUGMENTED REAL-VALUED TIME-DELAY NEURAL NETWORK ........ 40
 5.1 Introduction .................................................................. 40
 5.2 Augmented Real-valued Time-Delay Neural Network ......... 40
 5.3 Data processing procedure in NNs .................................. 44
    5.3.1 NN training ............................................................ 45
    5.3.2 NN Validation and Testing ......................................... 46
 5.4 Comparative Study between ARVTDNN and RVFTDNN .... 47
    5.4.1 Number of Modeling Data and Evaluation Data ............... 47
    5.4.2 Number of the Hidden Layer in NNs ......................... 48
    5.4.3 Memory Depth and Neuron’s Number ......................... 49
    5.4.4 Complexity Analysis ............................................. 50
    5.4.5 Measurement Result .............................................. 50
 5.5 Conclusion .................................................................... 54

CHAPTER SIX: EXTENDED CORRECTION BANDWIDTH DPD SYSTEM MODELED BY AUGMENTED REAL-VALUED TIME DELAY NEURAL NETWORK ...... 55
 6.1 Introduction .................................................................. 55
 6.2 Proposed ARVTDNN-based DPD System .......................... 55
 6.3 Conclusion .................................................................... 62

CHAPTER SEVEN: CONCLUSIONS AND FUTURE WORKS ................. 63
 7.1 Research summary ...................................................... 63
 7.2 Future Works ............................................................. 64

REFERENCES ....................................................................... 65
List of Tables

Table 1 Comparative study of the different number of modeling data samples .................. 48
Table 2 Comparative study in terms of NMSE for different layers .................................. 49
Table 3 Optimal parameters of proposed NN model and its comparison with RVFTDNN .... 49
Table 4 Optimal parameters of two-box ARVTDNN-based system .................................. 56
List of Figures and Illustrations

Figure 2.1 Amplitude and phase characteristics of the PA: (a) AM-AM characteristic (b) AM-PM characteristic [8] .............................................................................................................. 7

Figure 2.2 1 dB compression point ............................................................................................................. 8

Figure 2.3 Third-order intermodulation ...................................................................................................... 10

Figure 2.4 Power back-off of the PA .......................................................................................................... 11

Figure 3.1 Block diagram of feedback technique .......................................................................................... 16

Figure 3.2 Block diagram of feedforward technique ...................................................................................... 18

Figure 3.3 Block diagram of the LINC technique .......................................................................................... 19

Figure 3.4 Block diagram of DPD technique ................................................................................................ 19

Figure 3.5 PSD: (a) PSD of input and output signals (2) PSD of output signal with and without predistorter .............................................................................................................................................. 21

Figure 3.6 Block diagram of look-up table model ...................................................................................... 22

Figure 3.7 Block diagram of Hammerstein model ....................................................................................... 24

Figure 3.8 Block diagram of Wiener model .................................................................................................. 25

Figure 3.9 Structure of single neuron ......................................................................................................... 26

Figure 3.10 Single layer neural network ..................................................................................................... 27

Figure 3.11 A Multilayered NN .................................................................................................................. 28

Figure 3.12 Block diagram of NNs: (a) feedforward NN (b) recurrent NN .................................................. 29

Figure 4.1 Block diagram of conventional DPD system ............................................................................... 33

Figure 4.2 Block diagram of extended correction bandwidth DPD system ................................................. 33

Figure 4.3 Block diagrams of effects of different DPD systems ................................................................. 34

Figure 4.4 Block diagram of measurement setup ........................................................................................ 37

Figure 4.5 Real diagram of measurement setup .......................................................................................... 39

Figure 5.1 Block diagram of CVSISOFTNN ............................................................................................... 41

Figure 5.2 Block diagram of RVDIDOFNN ............................................................................................... 41
Figure 5.3 Block diagram of NNs: (a) ARVTDNN (b) RVFTDNN ................................................................. 44

Figure 5.4 Gain, Phase and input-output characteristics ................................................................. 50

Figure 5.5 Comparison of the modeled amplitude with actual amplitude obtained from the measurement setup .................................................................................................................. 51

Figure 5.6 Comparison of the modeled phase with actual phase obtained from the measurement setup .................................................................................................................................................. 51

Figure 5.7 Comparison of the modeled PSD with actual PSD obtained from the measurement setup .................................................................................................................................................. 52

Figure 5.8 Comparative epochs: (a) training performance of RVFTDNN (b) training performance of the ARVTDNN .................................................................................................................................................. 53

Figure 6.1 Block diagram of memoryless ARVTDNN ................................................................. 56

Figure 6.2 PSD of output signal of the PA at the observation bandwidth of 300 MHz ............... 58

Figure 6.3 PSD of output signal of the PA with the observation bandwidth: (a) 276 MHz (b) 241 MHz (c) 206 MHz (d) 184 MHz (e) 171 MHz (f) 135 MHz (g) 100 MHz .................. 61

Figure 6.4 Performance of ACPR at the different observation bandwidth ................................. 62
### List of Symbols, Abbreviations and Nomenclature

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACPR</td>
<td>Adjacent Channel Power Ratio</td>
</tr>
<tr>
<td>ADC</td>
<td>Analog to Digital Converter</td>
</tr>
<tr>
<td>AM-AM</td>
<td>Amplitude Modulation to Amplitude Modulation</td>
</tr>
<tr>
<td>AM-PM</td>
<td>Amplitude Modulation to Phase Modulation</td>
</tr>
<tr>
<td>AMPS</td>
<td>Advanced Mobile Phone Service</td>
</tr>
<tr>
<td>ARVTDNN</td>
<td>Augmented Real-Valued Time-Delay Neural Network</td>
</tr>
<tr>
<td>ARVTDNN-DPD</td>
<td>ARVTDNN DPD</td>
</tr>
<tr>
<td>BER</td>
<td>Bit Error Rate</td>
</tr>
<tr>
<td>CDMA</td>
<td>Code-Division Multiple-Access System</td>
</tr>
<tr>
<td>C-DPD</td>
<td>Conventional DPD</td>
</tr>
<tr>
<td>DPD</td>
<td>Digital Predistortion</td>
</tr>
<tr>
<td>EPA</td>
<td>Error Power Amplifier</td>
</tr>
<tr>
<td>EVM</td>
<td>Error Vector Magnitude</td>
</tr>
<tr>
<td>GHz</td>
<td>Gigahertz</td>
</tr>
<tr>
<td>GSM</td>
<td>Global system for mobile communications</td>
</tr>
<tr>
<td>I</td>
<td>In-Phase</td>
</tr>
<tr>
<td>IIP&lt;sub&gt;3&lt;/sub&gt;</td>
<td>Input IP&lt;sub&gt;3&lt;/sub&gt;</td>
</tr>
<tr>
<td>IMD&lt;sub&gt;3&lt;/sub&gt;</td>
<td>Third-Order IMD</td>
</tr>
<tr>
<td>IP&lt;sub&gt;2&lt;/sub&gt;</td>
<td>Second-Order Intercept Point</td>
</tr>
<tr>
<td>IP&lt;sub&gt;3&lt;/sub&gt;</td>
<td>Third-Order Intercept Point</td>
</tr>
<tr>
<td>IPBO</td>
<td>Input Power Back-Off</td>
</tr>
<tr>
<td>LS</td>
<td>Least Square</td>
</tr>
<tr>
<td>LTE</td>
<td>Long-Term Evolution</td>
</tr>
<tr>
<td>LTE-A</td>
<td>Long-Term Evolution-Advanced</td>
</tr>
<tr>
<td>LUT</td>
<td>Look-Up Table</td>
</tr>
<tr>
<td>MHz</td>
<td>Megahertz</td>
</tr>
<tr>
<td>MP</td>
<td>Memory Polynomial</td>
</tr>
<tr>
<td>NMSE</td>
<td>Normalized Mean Square Error</td>
</tr>
<tr>
<td>NMT</td>
<td>Nordic Mobile Telephones</td>
</tr>
<tr>
<td>NN</td>
<td>Neural Network</td>
</tr>
<tr>
<td>NTT</td>
<td>Nippon Telephone and Telegraph</td>
</tr>
<tr>
<td>OFDM</td>
<td>Orthogonal Frequency Division Multiplexing</td>
</tr>
<tr>
<td>OIP&lt;sub&gt;3&lt;/sub&gt;</td>
<td>Output IP&lt;sub&gt;3&lt;/sub&gt;</td>
</tr>
<tr>
<td>OPBO</td>
<td>Output Power Back-Off</td>
</tr>
<tr>
<td>P&lt;sub&gt;1&lt;/sub&gt; dB</td>
<td>1 dB Compression Point</td>
</tr>
<tr>
<td>PA</td>
<td>Power Amplifier</td>
</tr>
<tr>
<td>PAE</td>
<td>Power Added Efficiency</td>
</tr>
<tr>
<td>PAPR</td>
<td>Peak Average Power Ratio</td>
</tr>
<tr>
<td>PDC</td>
<td>Personal Digital Cellular</td>
</tr>
<tr>
<td>PPBO</td>
<td>Peak Power Back-Off</td>
</tr>
<tr>
<td>PSD</td>
<td>Power Spectrum Density</td>
</tr>
<tr>
<td></td>
<td>ix</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>Q</td>
<td>Quadrature-Phase</td>
</tr>
<tr>
<td>QAM</td>
<td>Quadrature Amplitude Modulation</td>
</tr>
<tr>
<td>QOS</td>
<td>Quality of Services</td>
</tr>
<tr>
<td>QPSK</td>
<td>Quadrature Phase Shift Key</td>
</tr>
<tr>
<td>R-DPD</td>
<td>RVFTDNN DPD</td>
</tr>
<tr>
<td>RF</td>
<td>Radio Frequency</td>
</tr>
<tr>
<td>RVFTDNN</td>
<td>Real-Valued Focused Time-Delay Neural Network</td>
</tr>
<tr>
<td>S-ARVTDNN-DPD</td>
<td>Memoryless ARVTDNN-DPD</td>
</tr>
<tr>
<td>S-BE-DPD</td>
<td>Memoryless BE-DPD</td>
</tr>
<tr>
<td>S-RVFTDNN-DPD</td>
<td>Memoryless RVFTDNN-DPD</td>
</tr>
<tr>
<td>TACC</td>
<td>Total Access Communications System</td>
</tr>
<tr>
<td>TD-CDMA</td>
<td>Time-Division Synchronous CDMA</td>
</tr>
<tr>
<td>TDMA</td>
<td>Time-Division Multiple-Access System</td>
</tr>
<tr>
<td>VSA</td>
<td>Vector Signal Analyzer</td>
</tr>
<tr>
<td>VSG</td>
<td>Vector Signal Generator</td>
</tr>
<tr>
<td>WCDMA</td>
<td>Wideband Coded Multiple Access</td>
</tr>
<tr>
<td>Wi-MAX</td>
<td>Worldwide Interoperability for Microwave Access</td>
</tr>
<tr>
<td>1G</td>
<td>First Generation</td>
</tr>
<tr>
<td>2G</td>
<td>Second Generation</td>
</tr>
<tr>
<td>3G</td>
<td>Third Generation</td>
</tr>
<tr>
<td>4G</td>
<td>Fourth Generation</td>
</tr>
</tbody>
</table>
Chapter One: Introduction

1.1 Introduction

Because of the rapid development of information techniques in the past 30 years, people have entered into the information era. Wireless mobile communication techniques as a symbol of the information era have been widely utilized in communication systems due to their flexibility. Accordingly, communication systems have evolved from first generation (1G) to fourth generation (4G) in the past 3 decades [1]. The 1G communication systems began in the early 1980s. The cellular communication technique of advanced mobile phone service (AMPS) first began for commercial telecommunication service in the United States in 1983. At the same time, other communication systems, like total access communication system (TACC), Nippon telephone and telegraph (NTT) and Nordic mobile telephones (NMT), were applied in other regions. The second generation (2G) digital communication systems were released to the market in the late 1980s. There were also several communication standards, such as global system for mobile communication (GSM), code-division multiple-access system (CDMA), time-division multiple-access system (TDMA), personal digital cellular (PDC), and so on.

The third generation (3G) communication systems were proposed due to some limitations of 1G and 2G systems. The 1G and 2G systems could offer voice and low data rate services and could not offer multi-media, high data rate and high quality video services. The 3G communication systems included wideband CDMA (WCDMA), Time-division synchronous CDMA (TD-CDMA) and CDMA 2000. The maximum data rate for the 3G systems was 144 kbps, 384 kbps and 2 Mbps for high/low mobility traffic and stationary condition, respectively. The 4G
communication systems provide high data rate service of 100 Msp for high mobility traffic and 1 Gsp for stationary condition.

The development of communication systems leads to the reduction of the amount of spectral resources. Hence, various modulation techniques, such as quadrature amplitude modulation (QAM), quadrature phase shift key (QPSK) and orthogonal frequency division multiplexing (OFDM), are used in communication systems to improve spectral efficiency. Although, these techniques can relieve the scarcity of spectral resources, the testing signals used by these techniques are more sensitive to the communication systems’ nonlinearity, and then engender more distortions. In fact, the most of the nonlinear distortions of communication systems is engendered by power amplifier (PAs). Hence, the behavioral characteristics of the PAs are mainly investigated in communication systems.

PA nonlinearities can be classified into four main categories: memory or memoryless and linear or nonlinear. If PA is a narrowband system that is depended on the channel’s coherence bandwidth, the PA is usually considered to be a memoryless system. If not, the PA is a system with memory, and the signal at the output of the PA is not only related to the current samples of the signal at the input of the PA, but is also associated with the past samples of the signal at the input of the PA. Memory effects are generally attributed to energy storage elements such as capacitors, and transistor. If PA works at the low power region, the PA is a linear system. If not, the PA is a nonlinear system. PAs in modern communication systems, e.g. LTE-A system, usually operate at the high power region to obtain higher power efficiency. Hence, the PA will engender strong nonlinear distortions, which result in spectral regrowth and could deteriorate bit error rate (BER). Thus, significant attention should be paid to correct the nonlinearity of a PA.
By overcoming these nonlinear distortions of the PA, communication systems can acquire good performance [2].

Initially, the technique of power back-off in early communication systems is applied to mitigate the nonlinear distortions of a PA. After using this technique, the PA can work at linear region. However, this technique decreases the power efficiency of the PA. Also, the memory effect of the PA need to be considered due to the increase of communication systems’ bandwidth. Hence, some other linearization techniques are proposed [2], [3].

Among all of the linearization techniques, the digital predistortion (DPD) technique [2] is one of the most effective techniques due to its simplicity and good linearization capability. The basic idea of the DPD technique is to create the inverse model of a PA. Thus, in order to obtain an accurate DPD model, the PA model should be characterized and modeled properly. The PA model can be built by using physical modeling and behavioral modeling [2], [4]. For the physical modeling of the PA, we must know the characteristic of various components of the PA and the structure of electric circuit of the PA. However, although physical modeling is very accurate, it is complicated [4]. Behavioral modeling is commonly known as black-box modeling [2]. The behavioral model of the PA can be built by using the input and output signals of the PA. It is much easier to extract PA parameters with behavioral modeling rather than physical modeling. Therefore, behavioral modeling is commonly applied to construct the PA model.

The common models used for the behavioral modeling of a PA include Volterra series model [3], [5], memory polynomial models [6] and neural network models (NN) [7], [8]. The Volterra series is one of the most popular models. In fact, the Volterra series is an extension of Taylor series [7], [8]. If the Volterra series is applied for the strong nonlinear distortions of the PA, the number of model coefficients will be significantly increased [8]. Hence, memory polynomial is
proposed to build the model for the strong nonlinear distortions of the PA. In fact, memory polynomial is a variation of the Volterra series [7]. The aim is at being able to reduce the number of model coefficients and simultaneously keep good modeling performance.

NNs also have a good modeling capability due to adaptive nature and approximation capability [9, 10, 11] for PAs. NNs can approximate any model when having enough neurons [11]. In the past 20 years, many NN models have been proposed in order to build the model of PAs. For example, a single-input single-output feedforward NN [9] was proposed to build the model of a PA. It introduces the complex-valued weights and transfer function, which result in cumbersome calculation. Thus, another method [10], [11], e.g. polar method, was proposed. The method separately constructs the amplitude and phase characteristics of the PA, and then combines the simulated amplitude and phase signals. However, this method cannot guarantee that the constructed amplitude and phase models simultaneously convergence [14]. Recently, radial basis NN [12], fuzzy logic NN [13] and real-valued focused time-delay neural network (RVFTDNN) [14] were presented in literature for the behavioral modeling of the PA with memory. They all have good modeling capability for the PA with memory. In this thesis, an augmented real-valued time-delay neural network (ARVTDNN) is proposed in order to have a better modeling and mitigation for the nonlinear distortions with memory in the PA.

1.2 Goal and Scope of Work
The objective of this thesis is to mitigate the nonlinear distortions of PAs by using the proposed two-box ARVTDNN-based DPD system, and simultaneously to reduce the sampling rate requirement of the feedback path’s analog to digital converter (ADC) in transmitters.

This thesis focuses on four areas:
1) Analyzing the nonlinear characteristics of PAs and investigating the previous linearization techniques and behavioral models of PAs;
2) Reviewing the existing two-box polynomial-based DPD system;
3) Implementing nonlinear distortions modeling and mitigation based on ARVTDNN;
4) Proposing a two-box ARVTDNN-based DPD system to reduce the required sampling speed of the feedback path’s ADC.

This thesis is organized as follows:

- Chapter one introduces the background and conceptual understanding of this thesis.
- Chapter two describes in detail the nonlinear characteristics of PAs.
- Chapter three provides a literature review of the various linearization techniques and behavioral modeling techniques of PAs.
- Chapter four presents the existing two-box polynomial-based DPD system, and measurement setup in order to validate the performance of model.
- Chapter five describes the modeling process of the proposed ARVTDNN, and compares the ARVTDNN with the RVFTDNN for the behavioral modeling of PAs.
- Chapter six compares the two-box ARVTDNN-based DPD system with two-box polynomial-based DPD system.
- Chapter seven concludes the thesis and describes the future work that can be carried out in the area of digital predistortion.
Chapter Two: RF Power Amplifier Characteristics

2.1 Introduction
PAs are a key building block of any communication system, because it not only dominates the power consumption of communication system, but also determinates the degree of nonlinearity in communication system [8]. Thus, it is very important for overcoming the limitations and imperfections of PA in communication system. This chapter mainly investigates the characteristics of PA and defines some metrics for evaluating the performance of a testing signal and a communication system.

2.2 RF Power Amplifier Nonlinear Characteristics
The following subsections introduce the characteristics of radio frequency (RF) PA.

2.2.1 Amplitude and Phase Characteristics
The main behavioral characteristics of a PA are amplitude and phase characteristics [7]. The amplitude characteristic, i.e. amplitude modulation to amplitude modulation (AM-AM), is expressed as a change of the output amplitude power level as a function of the input amplitude power level [7]. Similarly, the phase characteristic, namely amplitude modulation to phase modulation (AM-PM), is represented as a non-constant phase shift of the PA’s output signal according to the envelope the of the PA’s input signal [7]. The AM-AM and AM-PM characteristics of a PA are shown in Fig. 2.1 that is cited from [8]. From Fig. 2.1, we can clearly see that the AM-AM and AM-PM characteristics of the PA are not linear at the saturation region.

In addition, the input signal of the PA usually can be expressed as [8], [15], [16]

\[ x(t) = r(t) \cdot \cos(w_c t + \theta(t)). \]  (2.1)

The output signal of the PA can be expressed as [8], [15], [16]
\[ y(t) = g(r(t)) \cdot r(t) \cdot \cos \left( w_c t + \theta(t) + \varphi(r(t)) \right), \] (2.2)

where \( g(r(t)) \) represents the amplitude distortion of the PA and \( \varphi(r(t)) \) represents the phase distortion of the PA.

**Figure 2.1** Amplitude and phase characteristics of the PA: (a) AM-AM characteristic (b) AM-PM characteristic [8]
2.2.2 $P_{1\text{dB}}$ Compression Point

A typical 1 dB compression point ($P_{1\text{dB}}$) for which the actual output power is 1 dB less than the predicted linear output power in PAs [17] is shown in Fig. 2.2. From Fig. 2.2, we can see that the PA is working at linear region when input power works at low power region. However, when the input power is gradually increased, the PA enters into nonlinear region before reaching the saturation region. Herein, saturation point corresponds to the point where the PA reaches its maximum output power. Maximum output power is known as saturation power ($P_{\text{sat}}$).

![Figure 2.2 1 dB compression point](image)

2.2.3 Third-order Intercept Point

Third-order intercept point ($IP_3$) of a PA is an effective measure to evaluate the degree of linearity of the PA [7]. The higher the $IP_3$, the better the linearity of the PA [7]. The $IP_3$ of the PA is shown in Fig. 2.3. From Fig. 2.3, one can see that the $IP_3$ corresponds to the point where the output power of the fundamental frequency and third-order intermodulation products are equal at the same input power.
The IP$_3$ of the PA can be measured by using a two-tone signal. Consider a two-tone signal given by [7]

$$x(t) = A_1 \cdot \cos(w_1 t) + A_2 \cdot \cos(w_2 t), \quad (2.3)$$

where $A_1$ and $A_2$ are different constants, respectively, $w_1$ and $w_2$ are separately different angular frequencies, and $x(t)$ is input signal of the PA.

Then, the output signal of the PA is expressed as [7]

$$y(t) = \sum_{n=1}^{N} a_n \cdot x^n(t)$$

$$= a_1 \cdot (A_1 \cdot \cos(w_1 t) + A_2 \cdot \cos(w_2 t)) + a_2 \cdot (A_1 \cdot \cos(w_1 t) + A_2 \cdot \cos(w_2 t))^2$$

$$+ a_3 \cdot (A_1 \cdot \cos(w_1 t) + A_2 \cdot \cos(w_2 t))^3 + \cdots$$

$$= \frac{a_2(A_1^2 + A_2^2)}{2} + \left( a_1 A_1 + \frac{3a_3 A_1^3}{4} + \frac{3a_3 A_1 A_2^2}{2} \right) \cdot \cos(w_1 t)$$

$$+ \left( a_1 A_2 + \frac{3a_3 A_2^3}{4} + \frac{3a_3 A_1^2 A_2}{2} \right) \cdot \cos(w_2 t) + \frac{a_2 A_1^2}{2} \cdot \cos(2w_1 t) + \frac{a_2 A_2^2}{2} \cdot \cos(2w_2 t) + \frac{a_3 A_1^3}{4} \cdot \cos((w_1 + w_2)t)$$

$$+ \frac{a_3 A_2^3}{4} \cdot \cos((w_1 - w_2)t) + \frac{a_2 A_1 A_2}{4} \cdot \cos((w_1 + 2w_2)t) + \frac{3a_3 A_1 A_2^2}{4} \cdot \cos((w_1 - 2w_2)t)$$

$$+ \cdots, \quad (2.4)$$

where $a_1$, $a_2$, and $a_3$ are different constants, respectively, and $y(t)$ is the output signal of the PA.

From eq. (2.4), one can see that the output signal of the PA not only contains some integer harmonic products, but also includes some intermodulation products. These integer harmonic products can be easily filtered out by a passband filter around the fundamental frequency product.
[17] because there is a great distance between these integer harmonic and the fundamental frequency products. Hence, these integer harmonic products do not have a noticeable effect on the fundamental frequency products.

However, the third-order intermodulation products that is also referred to as the third-order intermodulation distortion (IMD₃) products cannot be neglected [17], because they are close to the fundamental frequency products and, hence, will affect the desired signal. In order to measure the degree of the IMD₃ products, the third-order intermodulation ratio (IMR₃) and the IP₃ are important metrics. The IMR₃ is defined as the ratio of the IMD₃ products’ amplitude to the fundamental frequency products’ amplitude [17]. As we know, the rate of increase in the IMD₃ products is three times that of the fundamental frequency products [17]. Hence, the fundamental frequency and IMD₃ products become equal in amplitude at a point that is known as IP₃ due to the increase of the input signal power. Of course, this is a hypothetical theoretical value. This will never happen in practice. From the above, one can see that the IMR₃ and the IP₃ are important for the evaluation of the PA’s characteristics.

![Figure 2.3 Third-order intermodulation](image-url)
2.2.4 Power Back-off

In general, power back-off is defined as the ratio of PA’s saturation power to the PA’s mean power [17]. There are three types of power back-off: input power back-off (IPBO), output power back-off (OPBO) and peak power back-off (PPBO) [17].

IPBO is expressed as

\[
IPBO(dB) = P_{in,sat} - P_{in,ave},
\]

(2.5)

where \(P_{in,sat}\) is the PA’s saturation input power and \(P_{in,ave}\) is the PA’s average input power.

Similarly, OPBO is given as

\[
OPBO(dB) = P_{out,sat} - P_{out,ave},
\]

(2.6)

where \(P_{out,sat}\) is the PA’s saturation output power and \(P_{out,ave}\) is the PA’s average output power.

And, PPBO is represented as

\[
PPBO(dB) = P_{out,peak} - P_{out,ave},
\]

(2.7)

where \(P_{out,peak}\) is the PA’s peak output power.

The power back-offs are illustrated in Fig. 2.4.

![Figure 2.4 Power back-off of the PA](image-url)
2.2.5 Power Efficiency

Power efficiency is an important metric to quantify the capability of a system for transforming a given input power to a useful output power [7], [8]. Power efficiency can be defined in two ways: drain efficiency and power added efficiency [7], [8].

The drain efficiency of PA, \( \eta_d \), is defined as the ratio of the useful output power to the given input direct current power. It is expressed as [8]

\[
\eta_d = \frac{P_{out}}{P_{DC}} \times 100\%,
\]

where \( P_{out} \) is the useful output power of the PA and \( P_{DC} \) is the supplied input direct current power.

The input signal power is applied to calculate the power added efficiency of PA. The power added efficiency of the PA, on the other hand, can be expressed as [8]

\[
\eta_{PAE} = \frac{P_{out} - P_{in}}{P_{DC}} \times 100\%,
\]

where \( P_{in} \) is the input signal power of the PA.

2.2.6 Memory Effects

A PA driven by a narrowband signal (e.g. few MHz) can be approximately regarded as a memoryless PA [17]. But, memory effects need to be considered when the driving signal is increased. The wider the driving signal, the higher the memory effects. Hence, we cannot neglect the memory effects of a PA driven by a wideband signal [17].

In substance, the memory effects of a PA [7, 8, 17] are caused by biasing networks, matching networks, and the capacitors and inductors of internal equivalent circuit. The memory effects of the PA can be classified as electrical memory effects and thermal memory effects.
The electrical memory effects are caused by the variation of impedance. Because the frequency-dependent envelope impedance is a non-constant impedance over a large frequency range [7].

The thermal memory effects are caused by the variation of transistor parameters resulting from the variation of the transistor’s temperature.

In general, the electrical memory effects can be relieved by optimizing the matching and biasing networks [17]. Also, the thermal memory effects are very difficult to be optimized. We only can compensate it by using some temperature compensation circuits [17].

2.3 Performance Evaluation

This following subsections introduce some metrics, which are applied to evaluate the performance of testing signals and communication systems.

2.3.1 Peak-to-Average Power Ratio

Peak-to-average-power-ratio (PAPR) is an important metric to quantify the rate of signal’s peak power to signal’s average power. A signal with high PAPR is more sensitive to PA’s nonlinearity [8], which results in high nonlinear distortions. The mathematical expression of PAPR can be written as [8]

\[
PAPR(dB) = 10 \times \log_{10}\left(\frac{P_{\text{peak}}}{P_{\text{ave}}}\right),
\]

where \(P_{\text{peak}}\) is the peak power of a signal and \(P_{\text{ave}}\) is the average power of a signal.

2.3.2 Normalized Mean Square Error

Normalized mean square error (NMSE) is used to evaluate the modeling capability of a model. NMSE is defined as [8]

\[
NMSE(dB) = 10 \times \log_{10}\left(\frac{\sum_{j=1}^{N}|y_{\text{sim}} - y_{\text{meas}}|^2}{\sum_{j=1}^{N}|y_{\text{meas}}|^2}\right),
\]
where $N$ is the number of the signal samples at the input of a DUT, and $y_{sim}$ and $y_{mea}$ are the corresponding simulated and measured signals at the output of a DUT, respectively. The smaller the NMSE, the better the modeling accuracy.

2.3.3 Error Vector Magnitude

Error Vector Magnitude (EVM) [1], [7] is one of the key parameters for communication systems. The EVM is the difference between the measured symbol location and ideal symbol location on modulation constellation diagram. The EVM is defined as [7]

$$EVM = \sqrt{\frac{\sum_{j=1}^{N}((I_{mea}(j)-I_{ref}(j))^2+(Q_{mea}(j)-Q_{ref}(j))^2)}{\sum_{j=1}^{N}(I^2_{ref}(j)+Q^2_{ref}(j))}} \times 100\%,$$  \hspace{1cm} (2.12)

where $N$ is the number of the measured signal samples, $I_{mea}$ and $Q_{mea}$ are the measured in-phase and quadrature signals, and $I_{ref}$ and $Q_{ref}$ are the ideal transmission in-phase and quadrature signals.

For a given standard and data rate, a maximum level of the EVM is specified by the standard in order to guarantee accurate demodulation. The EVM is usually caused by inter-symbol interference [7]. In addition, the close-in phase noise of local oscillators, carrier leakage and I/Q imbalance are the other sources for the degradation of modulation accuracy [8], [17].

2.3.4 Adjacent Channel Power Ratio

The EVM is a metric to quantify the quality of the in-band signal. Adjacent channel power ratio (ACPR) [1], [8] is defined in order to determine the distortions introduced to the signal in the adjacent channels. The definition of the ACPR is the ratio of the adjacent channel’s total power to main channel’s total power. It is expressed as
\begin{align*}
ACPR(dB) &= 10 \times log_{10} \left( \frac{\int_{f_c+\Delta f + \frac{B}{2}}^{f_c+\Delta f + \frac{B}{2}} PSD(f) \, df}{\int_{f_c-\frac{B}{2}}^{f_c+\frac{B}{2}} PSD(f) \, df} \right), \tag{2.14} \\
ACPR(dB) &= 10 \times log_{10} \left( \frac{\int_{f_c-\Delta f - \frac{B}{2}}^{f_c-\Delta f - \frac{B}{2}} PSD(f) \, df}{\int_{f_c-\frac{B}{2}}^{f_c+\frac{B}{2}} PSD(f) \, df} \right), \tag{2.15}
\end{align*}

where $PSD(f)$ is the power spectrum density of the measured signal, and $f_c$, $\Delta f$ and $B$ are the carrier frequency, the offset from the carrier frequency and the bandwidth of the measured signal, respectively.

**2.4 Conclusion**

This chapter had introduced the various characteristics of the PAs, such as AM-AM, AM-PM, $P_1$ dB compression point, and so on. In addition, some metrics, i.e. PAPR, EVM, ACPR, etc, were defined in order to evaluate the performance of the communication systems and the characteristics of the testing signals.
Chapter Three: Linearization and Behavioural Modeling of Power Amplifiers

3.1 Introduction

As we know, the linearity and power efficiency of a PA are two main criteria for communication systems [8]. However, good linearity and power efficiency cannot be simultaneously obtained, and there is a trade-off between linearity and power efficiency when designing the PA. With the aim of power efficiency in transmission chain, the PA must be operated at high nonlinear region. Even though, the PA sometimes works at saturation region. However, this results in strong nonlinear distortions for the PA. In order to cancel out the strong nonlinear distortions of the PA, some linearization techniques are proposed in literature, such as the feedback [18-20], feedforward [20], linear amplification with nonlinear components (LINC) [21], DPD [2], [6] and NN techniques [14].

3.2 Linearization Techniques for PA

3.2.1 Feedback Linearization Technique

Feedback linearization [18-20] is a common linearization method for a narrowband PA. The basic idea of the method is to transform a nonlinear system into a linear system. The block diagram of the method is shown in Fig. 3.1.
According to Fig. 3.1, a mathematical expression can be written as

\[ y(t) = \frac{G}{1+AG} \cdot x(t), \]  

(3.1)

where \( G \) is the transfer function of the PA, \( A \) is the transfer function of the feedback system, and \( x(t) \) and \( y(t) \) are the input and output signals of the PA, respectively.

The nonlinear distortions of the closed loop system are less than that of the open loop system at a given input power. Because the rate of increase of the output power in the closed loop system, namely \( \frac{G}{1+AG} \), is less than that of the open loop system, namely \( G \). Actually, the feedback path spreads the linearization range of transfer characteristic curve. However, the gain of the closed loop system is \( \frac{1}{1+AG} \) times the gain of the open loop system. Hence, one can say that the improvement of linearity in the PA is at the expense of the gain. Also, the feedback technique has a limited application, because it can only be applied in a narrowband PA. The feedback technique is not stable for a wideband PA, because it is very difficult to keep a constant gain over a very large frequency range.

3.2.2 Feedforward Linearization Technique

Feedforward linearization technique [20] has a good linearization performance for a PA. It is also absolutely stable [8]. In addition, it does not need to decrease the gain of the PA in order to linearize the nonlinear distortions of the PA. However, the feedforward linearization technique also has many drawbacks. The block diagram of the feedforward linearization technique is in Fig. 3.2.

From this figure, one can see that its circuit is more complex than that of the feedback linearization technique. In addition, the two delay blocks must be very tightly tuned for good performance. Also, the gain of the error power amplifier (EPA) must be very well matched to the
small signal gain (SSG) of the PA to amplify the error signal. Finally, the PA temperature also affects the linearization performance.

![Diagram](image)

**Figure 3.2 Block diagram of feedforward technique**

The basic idea of the feedforward linearization technique is that the error signal, \( e(t) \), is first generated. Then, the error signal is magnified by the EPA. Finally, the recovered error signal, \( e_d(t) \), will be subtracted by the delayed signal at the output of the EPA Equal Delay block.

### 3.2.3 Linear Amplification with Nonlinear Components

Linear amplification with nonlinear components (LINC) [21] is a very efficient linearization technique for a PA. The block diagram of the LINC technique is shown in Fig. 3.3. The basic idea of the LINC technique is to first decompose a non-constant envelope complex signal into two phase-modulated signals. And then, the two amplified phase-modulated signals at the output of the PA are combined to get back a complex signal. The key point of the LINC technique is the separation and combination of the signals, and the signal matching technique between branches. First, the phase discontinuities in the two phase-modulated signals result in a wideband spectrum. In addition, the loss that is introduced by the combining network influences
dramatically the efficiency of the LINC transmitter. Finally, gain and phase imbalances exist between both branches. The imbalances cause distortions in the LINC transmitter.

![Diagram of LINC Technique]

**Figure 3.3 Block diagram of the LINC technique**

### 3.2.4 Predistortion Linearization Technique

In modern communication systems, the DPD technique [2, 5-7] is one of the most popular techniques for the mitigation of PA’s nonlinear distortions. The DPD technique is unconditionally stable. The DPD principle is to generate additional distortions in order to compensate for the nonlinear distortions of PA. The block diagram of the DPD technique is shown in Fig. 3.4.

![Diagram of DPD Technique]

**Figure 3.4 Block diagram of DPD technique**

From Fig. 3.4, one can see that the mitigation of the nonlinear distortions of a PA uses a digital predistorter placed before the PA. The predistorter is the inverse model of the PA. Hence, in order to achieve a good linearization performance, a proper predistorter is the key.
The mathematical expression of the signal at various points in the Fig. 3.4 is written as

\[ r(t) = f_{DPD}(x(t)), \]  
\[ y(t) = f_{PA}(r(t)), \]  
\[ y(t) = f_{PA}(f_{DPD}(x(t))), \]

where \( x(t) \) is the actual input signal, \( f_{DPD}(\cdot) \) is the transfer function of the predistorter, \( r(t) \) is the predistorted signal at the output of the predistorter, \( f_{PA}(\cdot) \) is the transfer function of the PA, and \( y(t) \) is the output signal.

In order to better understand the influence of the predistorter, Fig. 3.5 presents the power spectral density (PSD) of signals. Fig. 3.5a shows not only the PSD of the input signal, \( x(t) \), but also the PSD of the output signal, \( y(t) \). From Fig. 3.5 (a), one can see that the PA exhibits a strong nonlinear distortions. If the PA is linear, the PSD of the input and output signals should be the same after normalizing the output signal by the small signal gain. Fig. 3.5b shows the PSD of the output signal with and without predistorter. From Fig. 3.5 (b), one can see that the ACPR performance of the output signal has an obvious improvement when the predistorter is applied. It means that the DPD system can suppress the nonlinear distortions of the PA.
In here, the input signal is a LTE signal with a bandwidth of 20 MHz, and the amplifier is a high power Doherty PA with a small signal gain of 13 dB with a saturation power of 44 dBm.

3.3 Behavioral Modeling of Power Amplifiers

3.3.1 Saleh Model

Saleh Model [22], [23] is used to fit the AM-AM and AM-PM characteristics of travelling wave tube amplifier (TWTA) [24] and solid state power amplifier (SSPA) [23], [25]. These characteristics are defined by Saleh’s model as

\[ A(r) = \frac{\alpha_a r}{1 + \beta_a r^2}, \]

\[ \varphi(r) = \frac{\alpha_\theta r}{1 + \beta_\theta r^2}, \]

where \( A(\cdot) \) and \( \varphi(\cdot) \) represent the AM-AM and AM-PM characteristics of the PA, respectively, and \( r \) represents the input signal amplitude of the PA. The four model parameters, \( \alpha_a, \beta_a, \alpha_\theta, \beta_\theta \), can be determined by using least square (LS) fit.
3.3.2 Look-up Table Model

Look-up Table (LUT) model [2], [26] can be used to describe the characteristics of a memoryless PA. The block diagram of the LUT model is shown in Fig. 3.6.

According to Fig. 3.6, its mathematical expression can be easily written as

\[ y(n) = G(|x(n)|) \cdot x(n), \quad (3.7) \]

where \( x(n) \) and \( y(n) \) represent the input and output signals of the model, respectively, and, \( G(|x(n)|) \) is the instantaneous complex gain of the model.

A modified version of the LUT is proposed in order to build the behavioral model of a memory PA in modern communication systems. The modified LUT is called Nested Look-up Table (NLUT) model [2], which can be expressed as

\[ y(n) = G(|X(n)|) \cdot x(n), \quad (3.8) \]

where \( G(|X(n)|) \) is also the instantaneous complex gain of the model, and \( X(n) \) is defined as

\[ X(n) = [x(n), x(n-1), \ldots, x(n-M)]. \quad (3.9) \]

3.3.3 Volterra Series

Volterra series [5] is the one of the most comprehensive models for the behavioral modeling of a PA. The Volterra series is expressed as

\[ y(n) = \sum_{k=1}^{K} \sum_{i_1=0}^{M} \ldots \sum_{i_k=0}^{M} h(i_1, \ldots, i_k) \prod_{j=1}^{k} x(n - i_j), \quad (3.10) \]
where \( h(i_1, \ldots, i_p) \) is the model coefficients, and \( K \) and \( M \) are the nonlinearity order and memory depth of the model, respectively.

Although, the Volterra series has a good modeling capability for a PA, however, the Volterra series is computationally very complex for a PA that exhibits strong nonlinearity. This limits the practical application of the Volterra series. Volterra series is usually applied for mild nonlinear distortions in a PA. In order to overcome the high complexity issue, many modified versions of the Volterra series, such as memory polynomial and memory two-box polynomial model [2] have been proposed in literature.

### 3.3.4 Memoryless and Memory Polynomial Models

Memoryless polynomial model [2], [8] is commonly used to construct the behavioral model of a memoryless PA. The memoryless polynomial model can be built by using the input signal and the measured output signal of the PA. The memoryless polynomial model is defined as

\[
y(n) = \sum_{i=0}^{N} a_i |x(n)|^i, \tag{3.11}
\]

where \( x(n) \) and \( y(n) \) represent the input and output signals of the model, respectively, \( N \) is the nonlinearity order of the model, and \( a_i \) is the complex-valued coefficients of the model.

In modern communication systems, the memory effect of a PA cannot be neglected. Several models have been investigated in literature, such as memory polynomial model [2], [6], orthogonal memory polynomial model [3] and envelope memory polynomial model [2] in order to characterize the memory effect of the PA. The memory polynomial model is defined as

\[
y(n) = \sum_{k=0}^{K} \sum_{m=0}^{M} a_{k,m} |x(n - m)|^k |x(n - m)|^m, \tag{3.12}
\]

where \( x(n) \) and \( y(n) \) are the input and output signals of the model, respectively, \( k \) and \( m \) are the nonlinearity order and memory depth of the model, respectively, and \( a_{k,m} \) is the coefficients of the model.
In fact, the mathematical expression of the memory polynomial model can also be written in matrix form. And then, the coefficients of the model can be calculated by using linear estimation algorithm. The model in matrix form is represented as

\[ Y = XA, \]  

(3.13)

where

\[ x_{k,m}(n) = x(n - m)|x(n - m)|^k, \]

\[ x_{k,m} = [x_{k,m}(n), x_{k,m}(n - 1), ..., x_{k,m}(n - P + 1)]^T, \]

\[ X = [x_{0,0}, ..., x_{0,M-1}, ..., x_{K,0}, ..., x_{K,M-1}], \]

and

\[ A = [a_{0,0}, ..., a_{0,M-1}, ..., a_{K,0}, ..., a_{K,M-1}]^T. \]

\( P \) is the number of testing data samples that is used for the PA model identification.

Finally, the model coefficients is given as

\[ A = [X^H X]^{-1}X^H Y. \]  

(3.14)

### 3.3.5 Hammerstein and Wiener Models

Hammerstein [2, 27, 28] and Wiener models [2], [27] are examples of two-box model used for the behavioral modeling of a PA. The Hammerstein model shown in Fig. 3.7 is based on the cascade of a LUT model followed by a finite impulse response (FIR) filter.

![Figure 3.7 Block diagram of Hammerstein model](image)

The mathematical expression of the Hammerstein model is represented as

\[ r(n) = G(|x(n)|) \cdot x(n), \]  

(3.15)

and

\[ y(n) = \sum_{j=0}^{M} h(j) \cdot r(n - j), \]  

(3.16)
where $h(j)$ is the coefficients of the FIR filter, $M$ is the memory depth of the model, and $G(|x(n)|)$ is the memoryless instantaneous complex gain of the model.

Similarly, the Wiener model consists of a FIR filter followed by a LUT model. It is shown in Fig. 3.8. The mathematical expression of the Wiener model is given as

$$r(n) = \sum_{j=0}^{M} h(j) \cdot x(n - j),$$  \hspace{1cm} (3.17) 

and

$$y(n) = G(|r(n)|) \cdot r(n).$$  \hspace{1cm} (3.18)

![Block diagram of Wiener model](image)

**Figure 3.8 Block diagram of Wiener model**

### 3.3.6 Neural Network Model

NNs are well known for their modeling capability focusing on a PA [13]. Initially, the NNs [7, 8, 13] were inspired by the structure and functionality of human brain. At that time, NNs mimicked the principle and characteristics of human brain to solve practical problems that can be solved by human brain. Hence, we can say that NNs are an abstraction and simplification of human brain. In recent years, many NN-based models have been proposed to construct the behavioral model of a PA due to their high approximation capability [9-11].

The basic component of the NNs is a neuron, which is composed of four elements: inputs, weights, biases, and outputs. Fig. 3.9 illustrates the structure of the neuron [7], [8].

From a mathematical point of view, the relationship between the input and output signals of a neuron is expressed as

$$y = f(\sum_{i=1}^{N} w_i \cdot x_i + b),$$  \hspace{1cm} (3.19)
where $N$ represents the number of the input signal samples, $f$ is the transfer function or the activation function of the neuron, and $w_i$ and $b$ are the adjustable synaptic weights and bias of the neuron.

![Figure 3.9 Structure of single neuron](image)

In practice, there are many different types of transfer functions in order to construct the behavioral model of different systems. Usually, hyperbolic tangent sigmoid transfer function is used to build the behavioral model of a PA. In the following, some transfer functions will be introduced.

Purelin- Linear transfer function is expressed as

$$Purelin(x) = x.$$  \hspace{1cm} (3.20)

Satlin and Satlins- Saturating and Symmetric saturating linear transfer functions are given as

$$Satlin(x) = \begin{cases} 0 & x \leq 0 \\ x & 0 \leq x \leq 1, \\ 1 & x \geq 1 \end{cases}$$ \hspace{1cm} (3.21)

$$Satlins(x) = \begin{cases} -1 & x \leq 0 \\ x & 0 \leq x \leq 1. \\ 1 & x \geq 1 \end{cases}$$ \hspace{1cm} (3.22)

Poslin- Positive linear transfer function is represented as
\[ Poslin(x) = \begin{cases} 0 & x \leq 0 \\ x & 0 \leq x \leq 1. \end{cases} \quad (3.23) \]

Tansig- Hyperbolic tangent sigmoid transfer function is given as

\[ tansig(x) = \frac{\exp^{2x} - 1}{\exp^{2x} + 1}. \quad (3.24) \]

Logsig-Log-sigmoid transfer function is expressed as

\[ logsig(x) = \frac{1}{\exp^{-x} + 1}. \quad (3.25) \]

In general, one neuron cannot describe the characteristics of a PA. Thus, a parallel combination of many neurons is applied to build the behavioral model of the PA. Such a combination forms a layer. A block diagram of a single layered NN is shown in Fig. 3.10. From Fig. 3.10, one can see that the input signal is connected to each neuron of the one-layered network. And the output signal at the output of the transfer functions is expressed as

\[ y_i = f(\sum_{j=1}^{M} w_{ij} \cdot x_j + b_i), \quad i = 1, \ldots, N \quad (3.26) \]

In order to better construct the behavioral model of the PA, a multilayered network is proposed. The multilayered network is shown in Fig. 3.11. In fact, multilayered networks are more popular.
than one-layered networks. In the multilayered network, there exist two types of NN structure:
feedforward and recurrent NNs. They are shown in Fig. 3.12.

Figure 3.11 A Multilayered NN
Figure 3.12 Block diagram of NNs: (a) feedforward NN (b) recurrent NN

From Fig. 3.12, one can see that the input signal of the feedforward and recurrent NNs is respectively given as

Feedforward NN:

\[ X(t) = [X_1(t), X_2(t), \ldots, X_N(t)]. \]  

(3.27)

Recurrent NN:

\[ X(t) = [X_1(t), X_2(t), \ldots, X_N(t), Y_1(t), Y_1(t-1), \ldots, Y_1(t-t_d), Y_2(t), \ldots, Y_2(t-t_d), \ldots, Y_q(t), \ldots, Y_q(t-t_d)]. \]  

(3.28)

where \( t_d \) is time delay line of the recurrent NN’s feedback path.

And, the feedforward NN allows the input signal to travel only in one direction from the input layer to the output layer. First, the input signal is transmitted into each neuron of the hidden layer’s first layer. Then, the output signal of the hidden layer’s first layer is transmitted into each neuron of the hidden layer’s second layer, and so on. Finally, the output signal of the hidden
layer’s last layer is transmitted into each neuron of the output layer. The only difference between the feedforward NN and the recurrent NN is whether the feedback path exists or not. The recurrent NN has a better modeling capability for a PA with memory. The drawback of the recurrent NN is that it is complicated. Also, it presents a serious problem for network stability [7]. Thus, some systems are modeled by using the recurrent NN, and some other systems are built by using the modified feedforward NN. And, time-delay feedforward NN [14] is a modified version of the feedforward NN. It not only has good memory modeling capability, but also is stable.

In addition, one can see that the NNs consists of input layer, hidden layer and output layer form Fig 3.12. The hidden layer could be either one-layered or multilayered. Usually, one-layered and two-layered networks are applied. The number of hidden layer is determined by training the network in this thesis. The input and output layers are both one-layered. The number of neurons in the hidden layers is also determined during the training of the network. Further details about the conceptual understanding and in-depth analysis of NNs will be discussed in chapter five.

After the model of the NNs is built, the NNs need to be trained. The training process includes two type of learning algorithms: supervised algorithm and unsupervised algorithm. The supervised learning algorithm relies on the input and output signals of the PA. One of the most popular algorithms in the supervised learning algorithm is back propagation learning algorithm (BPLA) [7], [29]. The BPLA is implemented by using Levenberg-Marquardt method [20], [30]. The unsupervised learning algorithm doesn’t depend on the external agent. It only depends on itself, called a self-organizing algorithm. It is very useful for some applications, and it doesn’t need any prior knowledge.
3.4 Conclusion

This chapter described several linearization techniques for PAs, such as feedback, feedforward and DPD techniques. The DPD in these techniques is the popular method for the mitigation of the nonlinear distortions in the PAs. In addition, the chapter introduced several behavioral models for the behavioral modeling of PAs. Finally, the basic structures of the NNs and the existing state-of-the-art NN techniques were presented.
Chapter Four: Two-box Polynomial-based DPD System and Measurement Setup

4.1 Introduction
This chapter discusses an existing two-box polynomial-based DPD system [31], which can reduce the sampling rate requirement of the feedback path’s ADC. Furthermore, this chapter describes the experimental setup, data pre-process, and post-process for the behavioral modeling of the PA.

4.2 Two-box DPD System
Broadband communication systems are becoming increasingly attractive due to the demands of high data rate. In 4G communication systems, LTE-Advanced signal is a wide band signal with bandwidth of up to 100 MHz, and results in high memory effects, and high sampling speed requirement of ADC and digital analog converter (DAC).

The sampling rate requirement of the ADC and DAC is usually based on signal bandwidth. The bandwidth of the signal observed at the output of the PA usually spans over five times the bandwidth of the signal at the input of the PA [32], [33]. This is because the third-order and fifth-order intermodulation products of the input signal generated by the nonlinear characteristics of the PA are usually considered. Seventh-order or higher order intermodulation products of the input signal do not have a noticeable effect [31]. Hence, the sampling rate requirement of the ADC and DAC is five times the original input signal bandwidth. The reason can be observed from Fig. 4.1 or Fig. 4.2, because the signal bandwidth at the input of the ADC and DAC are both five times the original input signal bandwidth.

The sampling rate of the ADC and DAC is usually not enough for LTE-A signal with a bandwidth of 100 MHz [31]. Some methods [31, 34, 35] have been proposed in order to satisfy
the sampling rate requirement of broadband signal. The method in [31] reduces the sampling rate of the feedback path’s ADC by using a two-box polynomial-based DPD system. The sampling rate is decreased by 30%. At the same time, it exhibited the same linearization performance as conventional DPD system. [34] proposes a band-limited digital predistortion technique. The principle is that the band-limiting function is inserted into the general Volterra operators of the DPD model to control the signal bandwidth in modeling. The technique in [35] introduces spectral extrapolation to the band-limited feedback signal for implementing DPD modeling in wideband signal.

Figure 4.1 Block diagram of conventional DPD system

Figure 4.2 Block diagram of extended correction bandwidth DPD system
In Fig. 4.1, the conventional DPD system is based on Wiener model, which is a two-box structure based on the cascade of a memory polynomial model followed by a memoryless polynomial model. A digital input signal with a bandwidth of 60 MHz is first converted to analog signal by the DAC after it goes through the DPD system. Then, the analog signal is up-converted by frequency up-conversion stage (UCS) before it goes through the PA. In the feedback path, the signal at the output of the PA is down-converted by frequency down-conversion stage (DCS) before being digitized by the ADC. Finally, the DPD system can be built by using the signal at the output of the DPD system and the signal at the output of the ADC. In Fig. 4.1, one can see that the sampling rate of the DAC and ADC are both 384 (300×1.28) MHz in order to cancel out the fifth-order intermodulation products of the PA. The 1.28 is the roll-off factor of the filter in the oscilloscope applied in my measurement setup. Here, some parameters need to be defined [31]:

- **Signal Bandwidth**: This refers to the signal bandwidth at the input of the DPD model.
- **DPD correction bandwidth**: This refers to the signal bandwidth at the output of the PA.
- **Signal generation bandwidth**: This refers to the signal bandwidth at the input of the DAC.
• Signal observation bandwidth: This refers to the signal bandwidth at the input of the ADC.

In this thesis, the input signal bandwidth of 20 MHz is used to construct the memoryless DPD model. Also, the input signal bandwidth of 60 MHz is applied to transmit the desired message that is because this thesis focuses on single band power amplifier, and the typical frequency band is only 60 MHz wide. Usually, bandwidths beyond 60 MHz can be obtained through carrier aggregation between several frequency bands.

In Fig. 4.2, the existing DPD system is performed in two steps. First, the memoryless narrowband DPD system is produced offline by narrowband input signal with a bandwidth of 20 MHz in order to cancel out the major memoryless nonlinear distortions of the PA. The modeling process of the memoryless model is exactly as the same as that of the above mentioned Wiener model. The memoryless model results in the reduced signal bandwidth at the output of the PA, which is shown in Fig. 4.3. The PSD of the signal at the output of the PA, accompanying with no model, memoryless model, and two-box DPD system, are respectively shown. As shown in Fig. 4.3, the signal bandwidth at the output of the PA linearized by using memoryless predistorter is narrower than that obtained at the output of the PA without any predistorter. This means the DPD correction bandwidth is less than five times the signal bandwidth. Then, the wideband input and output signals of the device under test (DUT) (memoryless model + PA) are used to synthesize the dynamic model in the reduced DUT bandwidth scenario. The dynamic model in Fig. 4.2 can be built by using the same modeling process. To sum up, the DPD correction bandwidth in Fig. 4.2 is less than that in Fig. 4.1.

An important question here is the reason for the memoryless polynomial model can cancel out the most of the memoryless (static) nonlinear distortions of the PA driven by a wideband signal.
In fact, the memoryless nonlinear distortions of the PA can be better modeled under narrowband condition [31, 36]. This is because the nonlinear behavior of a PA is a function of the operating carrier frequency and the input signal’s characteristics (mainly its average power, statistics, and bandwidth) [36]. For a given carrier frequency, signal bandwidth only affects the memory effects of a PA and does not affect its static nonlinear characteristics [36].

Furthermore, we need to consider the unequal sampling frequency between the memory model’s output signal and ADC’s output signal in the identification algorithm box in Fig. 4.2. The ADC’s output signal is first up-sampled to the same sampling frequency as the output signal of the memory model. The time and power alignments are required as it is common for all DPD systems. The time alignment can be implemented by using cross-correlation technique. After the time alignment between the input and output signals, we should extract the normalized output signal of the PA. Because the same power level for the input and output signals can build a better model [29]. The power alignment can be implemented by adjusting the average power of signals.

4.3 Measurement Setup

In this section, the measurement setup is described in order to implement the DPD system mentioned above.

4.3.1 Experimental Setup and Data Acquisition

Fig. 4.4 and 4.5 present the schematic and real diagram of a measurement setup, which is used for the acquisition of the PA’s output signal. For the measurement setup, a testing signal is first synthesized by using Agilent’s Advanced Design System (ADS) software. Then, the testing signal is downloaded into Arbitrary Waveform Generator (AWG) model 81180A through internet. The baseband signal generated by AWG is transmitted into Performance Signal Generator (PSG) E8267D via wideband I/Q inputs cable. The PSG from Agilent Technologies
performs digital modulation, digital to analog conversion, and frequency up-conversion, and finally generate a RF signal. In addition, an isolator is applied in order to prevent the PA’s reflection power into a driver. The signal power at the output of the PA is decreased by an attenuator before being transmitted into a high rate oscilloscope model MSO 9404A.

The signal in MSO 9404A is analyzed by using Agilent 89600 Vector Signal Analyzer (VSA) software that can control the MSO 9404A. The VSA performs frequency down-conversion, analog to digital conversion and demodulation. Digital baseband in-phase and quadrature signals in the oscilloscope will be obtained by using the VSA. The obtained digital baseband signal and the testing signal are processed in the MathWork’s MATLAB® software in order to obtain the coefficients of the DPD system.

![Figure 4.4 Block diagram of measurement setup](image-url)
4.3.2 PA/signals used in the experiment

The PA in the measurement setup is a Doherty PA with a small signal gain of 13 dB and a saturation power of 44 dBm. The driver stage provides a gain of 37 dB. The PA is driven by two LTE-A signals. The bandwidth of the two signals is 60 MHz and 20 MHz, respectively. Accordingly, their sampling rates are 384 MHz and 128 MHz, respectively. The duration time of the two signals are both 1 ms, with 384000 symbols and 128000 symbols, respectively. The two signals are both modulated at a center frequency of 2.14 GHz.

4.4 Conclusions

In this chapter, the existing two-box polynomial-based DPD system was presented to illustrate how to mitigate the nonlinear distortions of the PA, simultaneously keep the same linearization performance for the PA. Furthermore, the measurement setup was elaborated. Finally, the details of the PA and testing signal were described.
Chapter Five: Behavioral Modeling of Power Amplifiers by using Augmented Real-Valued Time-Delay Neural Network

5.1 Introduction

This chapter proposes an augmented real-valued time-delay neural network (ARVTDNN) for the behavioral modeling of the PA. The concepts of both the existing neural networks (NNs) and proposed ARVTDNN are described. In addition, the modeling process of the PA by using the proposed ARVTDNN is also elaborated. Finally, a comparison between ARVTDNN and real-valued time-delay neural network (RVFTDNN) is illustrated.

5.2 Augmented Real-valued Time-Delay Neural Network

NN theory has gone through a rapid development during the past 20 years. At their most basic level, NNs usually contain an input layer, hidden layers and an output layer [29]. Earlier, the most basic structure was a complex-valued single-input and single-output feedforward NN (CVSISOFNN) [9]. The complex-valued input and output signals engender the complex-valued weights, biases and transfer functions, which result in the cumbersome calculation. The block diagram, as shown in Fig. 5.1, has illustrated the complex-valued weights, $w_{i,j}^l$, the complex-valued biases, $b_i^l$, and the complex-valued transfer function, i.e. $(a) = \frac{\exp^{2a} - 1}{\exp^{2a} + 1}$, where, $i$ denotes the $i^{th}$ neuron of the current layer, $j$ denotes the $j^{th}$ neuron of the previous layer, $w_{i,j}^l$ is the synaptic weight connecting the $j^{th}$ neuron of the $(l-1)^{th}$ layer to the $i^{th}$ neuron of the $l^{th}$ layer and $a$ is the signal at the input of the transfer function.

In order to address the issue mentioned above, a real-valued double-input double-output feedforward NN (RVIDOFNN) [29] in Fig. 5.2 was proposed. The only difference between CVSISOFNN and RVIDOFNN models is that CVSISOFNN and RVIDOFNN are complex-
valued and real-valued input signals, respectively. In addition, the RVDIDOFNN decreases the pre-processing and post-processing expenditure due to the capability of the real-valued input signal [14]. The block diagram of RVFTDNN as shown in figure 5.2 denotes the in-phase signal by $I_{in}(t)$, the quadrature signal by $Q_{in}(t)$.

**Figure 5.1 Block diagram of CVSISOFNN**

**Figure 5.2 Block diagram of RVDIDOFNN**
In modern communication systems, we have to consider the modeling capability of NNs for the memory effects of the PA. The RVFTDNN [14] and the augmented radial basis function neural network (ARBFNN) [37] had been proposed in literatures to mitigate the memory effects of the PA. The input and output signals of the ARBFNN are complex signals like that of the CVSISOIFNN, which results in huge complexity. In this thesis, the ARVTDDN is proposed for the behavioral modeling of the PA. The difference between ARVTDDN and RVFTDNN is that the input signal of the ARVTDDN, unlike RVFTDNN, contains the signal amplitude, the square of the signal amplitude and the cube of the signal amplitude. These additional amplitude terms can generate additional distortions that are added to the in-band product and the third-order intermodulation product, when remixed with the in-band signal again, resulting in richer generation of distortions [38-40]. In addition, the corresponding memory effects generated by these additional amplitude terms are also considered due to the non-constant envelope impedance over a large frequency range in the matching circuits. Hence, when these terms are added to the proposed DPD, it results in a better cancellation of these distortions. Consequently, the ARVTDDN obtain a better modeling performance than the RVFTDNN.

The block diagrams of ARVTDDN and RVFTDNN, as shown in Fig. 5.3, both have illustrated the in-phase signal, $I_{in}(t)$, the quadrature signal, $Q_{in}(t)$, the in-phase and quadrature signals’ past samples, $[I_{in}(t-1), \cdots, I_{in}(t-M), Q_{in}(t-1), \cdots, Q_{in}(t-1)]$. In here, $M$ denotes the memory depth. The block diagram of ARVTDDN also introduced the additional terms: signal amplitude, i.e. $|x_{in}(n)|$, the square of the signal amplitude, i.e. $|x_{in}(n)|^2$, the cube of the signal amplitude, i.e. $|x_{in}(n)|^3$, and their past samples, i.e. $[|x_{in}(n-1)|, \cdots, |x_{in}(n-M)|, |x_{in}(n-1)|^2, \cdots, |x_{in}(n-M)|^2, |x_{in}(n-1)|^3, \cdots, |x_{in}(n-M)|^3]$. 
5.3 Data processing procedure in NNs

The modeling data processing procedure in NNs is very important for the modeling of the PA. The modeling data can be divided into three categories and can be used respectively for training, validation and testing. In this thesis, the data that we used for training process takes up 60% of all the available data. What we used for validation process takes up 20% of the whole data and what we used for testing processor takes up the remaining 20% of the whole data.
5.3.1 NN training

From the NN model as shown in Fig. 5.3, one can see that both the input signal and the output signal are represented as Cartesian coordinate components. The input signal of the RVFTDNN contains the current and past samples and the corresponding vector is represented as [14]

\[ X_{in} = [I_{in}(n), I_{in}(n-1), \ldots, I_{in}(n-M); Q_{in}(n), Q_{in}(n-1), \ldots, Q_{in}(n-M)], \]  

(5.1)

where \( I_{in}(n) \) denotes the in-phase component of the current samples and \( I_{in}(n-k), k = 1, 2, \ldots, M \), is the in-phase component of the past sample; \( Q_{in}(n) \) is the quadrature component of the current samples and \( Q_{in}(n-k), k = 1, 2, \ldots, M \), is the quadrature component of the past sample; \( M \) indicates the memory depth.

According to Fig. 5.3, the vector of the input signal in ARVTDNN, \( X_{in} \), consists of both \( I \) and \( Q \) components, the amplitude of the baseband input signal, the square and the cube of the amplitude of the baseband input signal, which is shown as follows

\[ X_{in} = [I_{in}(n), I_{in}(n-1), \ldots, I_{in}(n-M); Q_{in}(n), Q_{in}(n-1), \ldots, Q_{in}(n-M), |x_{in}(n)|, |x_{in}(n-1)|, \ldots, |x_{in}(n-M)|, |x_{in}(n)|^2, |x_{in}(n-1)|^2, \ldots, |x_{in}(n-M)|^2, |x_{in}(n)|^3, |x_{in}(n-1)|^3, \ldots, |x_{in}(n-M)|^3]. \]  

(5.2)

In terms of the training process, the ARVTDNN and RVFTDNN are similar. Regarding the training of the NNs, the initial value of the synaptic weights of the NNs should be given first. According to the literature [14], the initial value of the hidden layer’s synaptic weights is set in the range of [-0.8, 0.8]. This is because a high initial value of the synaptic weights will drive the network into the saturation region of the hyperbolic tangent sigmoid transfer function so that the learning rate is very slow and the training is stopped at very low initial values. The lowest and highest synaptic weights are -1 and 1, respectively. For the output layer, the synaptic weights are also set in the interval of [-0.8, 0.8] for simplification [14], [29].
The output of each neuron is equal to the bias plus the sum of the products of the input signal and the corresponding weight, which can be expressed as [14]

\[ a_i^l(n) = \sum_{j=1}^{p} w_{ij}^l \cdot x_j^{l-1}(n) + b_i^l. \]  

(5.3)

where all the mathematic symbols have the same meaning as before. The output of each neuron is then transformed by the transfer function, as follows, [14]

\[ c = f(a_i^l(n)). \]  

(5.4)

Finally, the cost function of the training data is calculated in batch mode during the forward pass [29]. The mean square error (MSE) is calculated by [14]

\[ E = \frac{1}{2N} \sum_{n=1}^{N} [(I_{sim} - I_{mea})^2 + (Q_{sim} - Q_{mea})^2], \]  

(5.5)

where \( I_{mea} \) and \( Q_{mea} \) are the in-phase and quadrature components of the measured output signal, respectively, and \( I_{sim} \) and \( Q_{sim} \) are the in-phase and quadrature components of the simulated output signal, respectively.

After we obtain the cost function of the training network during one epoch, the value of the synaptic weights and biases will be adjusted by using Levenberg-Marquardt algorithm [20, 30]. Then the cost function will be calculated again by using the updated synaptic weights and biases during the next epoch.

5.3.2 NN Validation and Testing

The training process will not terminate until NN convergence or reaching the preset maximum epoch or obtaining the desired performance or overtraining [29]. The desired performance and the maximum epoch are set respectively to be \( 10^{-6} \) and 100 [29]. Overtraining can be validated by using validation data. The occurrence of the overtraining will deteriorate the modeling performance. Thus, validation process is also very important for system identification.
Simultaneously, the testing data can evaluate the NN performance in terms of MSE during each epoch. In order to find out the last epoch’s performance of the NN, another independent 4000 input and output sample pairs are calculated by using NMSE algorithm.

5.4 Comparative Study between ARVTDNN and RVFTDNN

5.4.1 Number of Modeling Data and Evaluation Data

The number of modeling data in NNs is very important for the modeling and mitigation of the nonlinear distortions in PAs [7-8, 14]. It usually can be determined by systematically changing training parameters or using empirical method [7], [14]. Here, the training method is applied to obtain the optimal number of the modeling data. In addition, the number of the modeling data and hidden layer’s neurons are considered during the training network, which are respectively in the range of [4000, 11000] and [5, 107]. Also, the number of the hidden layer is assumed to be 1, and the number of the epoch is assumed to be 10.

In order to simplify it that obtains the optimal number of modeling data, the ARVTDNN model is only applied to evaluate it. The modeling data consists of the training, validation and testing data. Furthermore, an independent data set of 4000 sample pairs are applied to evaluate the NN’s performance at the last epoch by using NMSE algorithm.

The simulation results shown in Table 1 exhibit the performance of the different modeling data, and the corresponding optimal neuron’s number and NMSE. From table, one can observe that 11000 modeling sample pairs obtain the best NMSE. However, the number of its neurons is the biggest as given in the second column. In addition, 5000 modeling sample pairs have only a small reduced NMSE performance, and the number of the corresponding neurons is dramatically decreased. Hence, 5000 modeling samples pairs are applied, and another 4000 sample pairs are used to evaluation the NN’s performance at the last epoch.
Table 1 Comparative study of the different number of modeling data samples

<table>
<thead>
<tr>
<th>Num. of Modeling Data</th>
<th>Optimal NO. of Neurons</th>
<th>NMSE (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4000</td>
<td>11</td>
<td>-37.43</td>
</tr>
<tr>
<td>5000</td>
<td>19</td>
<td>-38.25</td>
</tr>
<tr>
<td>6000</td>
<td>28</td>
<td>-37.78</td>
</tr>
<tr>
<td>7000</td>
<td>22</td>
<td>-38.33</td>
</tr>
<tr>
<td>8000</td>
<td>25</td>
<td>-37.86</td>
</tr>
<tr>
<td>9000</td>
<td>24</td>
<td>-38.00</td>
</tr>
<tr>
<td>10000</td>
<td>31</td>
<td>-37.76</td>
</tr>
<tr>
<td>11000</td>
<td>41</td>
<td>-38.36</td>
</tr>
</tbody>
</table>

5.4.2 Number of the Hidden Layer in NNs

At present, no approach can clearly determine the number of the hidden layer in a particular problem [7]. The ways to deal with this question is by using the empirical method or systematically changing training parameters [7]. In order to achieve the optimal number of the hidden layer in ARVTDNN and RVFTDNN, the different number of the neurons and hidden layers is tried out. In practice, the one-layered and two-layered NNs are commonly applied, and also assuming that hardware for the one-layered and two-layered NNs is available.

The performance of the ARVTDNN and RVFTDNN for the different number of the hidden layers is shown in Table 2. From Table 2, one can observe that the modeling capability of the ARVTDNN is better than that of the RVFTDNN for any layer. Also, the NMSE performance of
the two-layered structure is better than that of one-layered structure. However, the two-layered structure does not have a tremendously improvement on NMSE, and improving dramatically the complexity as compared to the one-layered structure. Hence, the one-layered structure is adopted in this thesis for the behavioral modeling of the PA.

Table 2 Comparative study in terms of NMSE for different layers

<table>
<thead>
<tr>
<th>Model</th>
<th>RVFTDNN</th>
<th>ARVTDNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>one-Layer</td>
<td>-35.72 dB</td>
<td>-38.33 dB</td>
</tr>
<tr>
<td>two-Layer</td>
<td>-36.35 dB</td>
<td>-38.89 dB</td>
</tr>
</tbody>
</table>

5.4.3 Memory Depth and Neuron’s Number

Memory depth and the neuron’s number of hidden layer in ARVTDNN and RVFTDNN models are also determined by systematically changing the memory depth and neuron’s number. Here, the memory effect is in the range of [0, 5]. The simulation results shown in Table 3 present the optimal memory depth and neuron’s number along with the corresponding NMSE performance. The optimal parameters of the ARVTDNN and RVFTDNN models are respectively (3, 19) and (3, 52), and their corresponding NMSE are -38.33 dB and -35.72 dB.

Table 3 Optimal parameters of proposed NN model and its comparison with RVFTDNN

<table>
<thead>
<tr>
<th>Model</th>
<th>Optimal NO. of Neurons</th>
<th>NO. of Delay Taps</th>
<th>NMSE (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RVFTDNN</td>
<td>52</td>
<td>3</td>
<td>-35.72</td>
</tr>
<tr>
<td>ARVTDNN</td>
<td>19</td>
<td>3</td>
<td>-38.33</td>
</tr>
</tbody>
</table>
5.4.4 Complexity Analysis

In order to clearly compare the complexity of the ARVTDNN and RVFTDNN models, the number of the synaptic weights in NNs, $N_w$, is calculated by the following equation

$$N_w = (n_{in} + m \times n_{in}) \times k + k \times n_{out} + k + n_{out},$$

(5.6)

where $n_{in}$ and $n_{out}$ are the number of the current input and output signal vectors, $m$ is the memory depth of the input signal, and $k$ is the number of the hidden layer’s neurons. Through calculation, the number of the synaptic weights of RVFTDNN and ARVTDNN are 574 and 439, respectively. As shown, the ARVTDNN models have less complexity by 23.5% for the behavioral modeling of the PA, when compared to the RVFTDNN.

5.4.5 Measurement Result

The gain, phase and input-output characteristics of the realistic Doherty PA are all shown in Fig. 5.4. From the figure, one can observe that the applied PA is not only a highly nonlinear PA, but also exhibits strong memory effects, which are caused by the energy storage elements and frequency-dependent impedances.

![Figure 5.4 Gain, Phase and input-output characteristics](image_url)
Figure 5.5 Comparison of the modeled amplitude with actual amplitude obtained from the measurement setup

Figure 5.6 Comparison of the modeled phase with actual phase obtained from the measurement setup
Figure 5.7 Comparison of the modeled PSD with actual PSD obtained from the measurement setup
Figure 5.8 Comparative epochs: (a) training performance of RVFTDNN (b) training performance of the ARVTDDN
From Fig 5.5, one can see that the RVFTDNN does not have a good modeling during the fast transition states. The maximum deviation between the measured and simulated values is 0.01V. However, the measured and simulated signals for the ARVTdNN model are similar. Also, a similar situation shown in Fig. 5.6 for the phase characteristics of the PA is observed.

Fig. 5.7 (a) show the PSD of the actual PA’s output signal and the two mimicked PA’s output signals. The two mimicked PA is respectively modeled by the ARVTdNN and RVFTDNN models. From Fig. 5.7 (b), one can clearly observe that the ARVTdNN has a better modeling capability for as compared to the RVFTDNN.

Fig. 5.8 shows the performance of the two simulated output signals generated by the ARVTdNN and RVFTDNN models at each epoch. One can observe that the performances between epoch of 10 and 22 are the similar for the ARVTdNN, while performance of epoch 20 is as the similar as epoch 52’s performance. This means that the ARVTdNN would save more than half the training time, when compared to the RVFTDNN. This is because each epoch of the ARVTdNN is spending less training time than that of RVFTDNN, which is based on the mentioned complexity analysis above. To sum up, simulation results show that the ARVTdNN has a better modeling performance, and simultaneously saving more than half the training time, when compared to the RVFTDNN.

**5.5 Conclusion**

This chapter first described the basic knowledge of the existing NNs, such RVFTDNN and CVSISOFFNN. Then, the training process, validation process, and testing process for the ARVTdNN and RVFTDNN were elaborated. Finally, the modeling capability of the ARVTdNN and the RVFTDNN were compared for a realistic Doherty PA.
Chapter Six: Extended Correction Bandwidth DPD System Modeled by Augmented Real-Valued Time Delay Neural Network

6.1 Introduction

As we know, the behavioral model of a DPD system is ideally the inverse model of a PA [7]. Hence, the DPD system can be built by exchanging the input and output signals of the PA. A two-box DPD system modeled by augmented real-valued time delay neural network (ARVTDNN) is proposed in this thesis. Furthermore, the modeling and complexity performances of the proposed two-box DPD system are investigated.

6.2 Proposed ARVTDNN-based DPD System

The proposed two-box DPD system is a two-box structure based on the cascade of a memory model followed by a memoryless (static) model. The memoryless model is designed by using the memoryless ARVTDNN. The only difference between the memoryless ARVTDNN and the ARVTDNN is that the memoryless ARVTDNN doesn’t have memory taps. The implementation process of the two-box ARVTDNN-based DPD system is exactly the same as that of the two-box polynomial-based DPD system [31]. Also, the LTE-A input signal with the bandwidth of 20 MHz is used to build the memoryless model based on the memoryless ARVTDNN. In addition, the LTE-A input signal with the bandwidth of 60 MHz is used to build the memory (dynamic) model modeled by the memory ARVTDNN. The PAPR of the two signals is both 10.4 dB. The block diagram of the memoryless ARVTDNN is shown in Fig. 6.1.

In order to better linearize the nonlinear distortions of the PA, the optimal two-box ARVTDNN-based DPD system should be obtained by systematically changing the training parameters as mentioned in chapter five. The optimal parameters of the two-box based-ARVTDNN DPD system is presented in Table 4. From Table 4, one can see that the optimal number of the hidden
layer’s neurons and the optimal memory depth for the dynamic model are \([11, 3]\). The optimal memoryless model is built by using 24 neurons. The corresponding optimal NMSE value of the memoryless and memory ARVTDNNs are \([-34.25, -40.49]\).

![Block diagram of memoryless ARVTDNN](image)

**Figure 6.1** Block diagram of memoryless ARVTDNN

**Table 4** Optimal parameters of two-box ARVTDNN-based system

<table>
<thead>
<tr>
<th>Model</th>
<th>Optimal NO. of Neurons</th>
<th>NO. of Delay Taps</th>
<th>NMSE (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARVTDNN (Static)</td>
<td>24</td>
<td>0</td>
<td>-34.25</td>
</tr>
<tr>
<td>ARVTDNN (Dyn.)</td>
<td>11</td>
<td>3</td>
<td>-40.49</td>
</tr>
</tbody>
</table>
The linearization performances of the two-box DPD system based on ARVTDNN and polynomial in the context of the normal observation bandwidth are shown in Fig. 6.2. The normal observation bandwidth is 300 MHz, and the corresponding sampling rate is 384 MHz (300 × 1.28). From Fig. 6.2, one can see that the ACPR performance of the PA linearized by DPD system modeled by ARVTDNN is dramatically improved, when compared to the nonlinear PA without any DPD system. Furthermore, the compensation capability of the ARVTDNN is better by 3-5 dB than that of the MPM.

The linearization performance of the ARVTDNN and MPM in the context of the different reduced observation bandwidth is shown in Fig. 6.3. Simultaneously, the PSD of the two models in the normal observation bandwidth condition is applied as the reference signals. In addition, the whole ACPR performances of the ARVTDNN and MPM at the different observation bandwidth are shown in Fig. 6.4. The ACPR is measured by the ratio of the power contained in bandwidth with 20 MHz, at offset with 70 MHz, from the center frequency with 2.14 GHz, to the main channel’s power in bandwidth with 20 MHz, around 2.14 GHz. From Fig. 6.4, one can clearly observe that the two-box polynomial-based DPD system does not meet the spectrum emission mask, because the minimum ACPR requirements for LTE-A is -45dBc. In addition, the two-box ARVTDNN-based DPD system still meets the ACPR requirements until the observation bandwidth is 155 MHz.
Figure 6.2 PSD of output signal of the PA at the observation bandwidth of 300 MHz

...
(b)

Observation Bandwidth = 241 MHz

Observation Bandwidth = 206 MHz

(c)
Observation Bandwidth = 184 MHz

Observation Bandwidth = 171 MHz
Figure 6.3 PSD of output signal of the PA with the observation bandwidth: (a) 276 MHz (b) 241 MHz (c) 206 MHz (d) 184 MHz (e) 171 MHz (f) 135 MHz (g) 100 MHz
Figure 6.4 Performance of ACPR at the different observation bandwidth

6.3 Conclusion

In this chapter, the linearization performance of the ARVTDNN and MPM were shown. The measurement results have demonstrated that the ARVTDNN has better compensation capability and requires less observation bandwidth than MPM-DPD.
Chapter Seven: Conclusions and Future Works

7.1 Research summary

In this thesis, we reviewed several common linearization techniques, such as feedback technique, feedforward technique, LINC technique and the most promising predistortion technique. The basic idea of the predistortion is to generate a nonlinear device and put this nonlinear device before the PA in order to overcome the nonlinear distortions of the PA. The ideal predistorters should be the inverse of the PA model, in theory, to better compensate the nonlinear distortions of the PA. Therefore, the behavioral modeling of the PA is significant for designing accurate predistorters. Chapter two and three have reviewed the characteristics of the PA and the existing modeling of the PA.

As is well known, it is important to acquire a high linearity and a high efficiency of the PA. In this thesis, a general method is developed to cancel out the nonlinear distortions of the PA and to increase the bandwidth capability of the PA. The two-box ARVTDNN-based DPD system not only reduces the sampling rate of the ADC of the feedback path, but also allows the PA to reside in nonlinear region, significantly increasing power efficiency.

Chapter four presented an existing two-box DPD system, which was built by polynomial model. This thesis proposed a new two-box ARVTDNN-based DPD system. The two-box DPD system consists of a memory predistorter and a memoryless predistorter. The memoryless predistorter is obtained by using the memoryless ARVTDNN. The memory predistorter is obtained by using the ARVTDNN. The ARVTDNN has a better linearization for the memory effect of the PA. We also compared the ARVTDNN with the polynomial in terms of the capability of linearization. The ARVTFNN performs better than RVFTDNN by 3-5 dB.
7.2 Future Works

The multi-band and wideband PA are prerequisite for the current and future wideband communication systems. LTE-A has a maximum bandwidth of 100MHz. The wide bandwidth beyond 60 MHz is commonly obtained through carrier aggregation of several different bands. The multi-band PA has a very high nonlinear memory distortion due to the extra inter-band intermodulation distortion. In order to linearize the multi-band PA, the multi-band predistorters need to be proposed. Dual-Band predistorters have been proposed in literature [8], [41] and would be valuable for the behavioral modeling of the multi-band predistorters. However, the multi-band predistorters have a more obvious disadvantage, i.e. the large number of model coefficients. Therefore, we will focus on how to decrease the number of coefficients for the multi-band PA.


References


