2014-01-30

Signal Processing Techniques for Power Efficiency and Signal Quality Enhancement of SISO and MIMO Radio Systems

Vejdaniamiri, Mehdi

UNIVERSITY OF CALGARY

Signal Processing Techniques for Power Efficiency and Signal Quality Enhancement of SISO and MIMO Radio Systems

by

Mehdi Vejdani Amiri

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

CALGARY, ALBERTA

January, 2014

Abstract

Communications has tremendously evolved during last 30 years period with the goal to realize real-time communications wirelessly. As the communication is targeted to be real-time, there is always a need to achieve the maximum possible data rate. Moreover, link reliability should always be guaranteed to receive a reliable version of the transmitted signal.

RF front-ends as the physical layer of the communication systems, suffer from non-ideal behaviour of most of the actual electronic components which causes a nonlinear dynamics. Such systems introduce some amount of distortions to the signal.

The other issue which has attracted much attention is the power efficiency which mainly deals with the cost and reliability as well as recently environmental impacts of the communication systems. This dissertation proposes a couple of novel signal processing techniques to overcome the problems associated with the single input single output (SISO) and multiple input multiple output (MIMO) radio systems.

The first topic of this research is devoted to efficiently partition the linearization scheme between the base station and mobile terminal. Phase distortions are compensated at the base station transmitter and the compensation of amplitude distortions is devoted to the mobile terminal receiver. This technique improves the power efficiency of communication link and in particular the transmitter.

Then, the above technique is extended to base stations to improve the efficiency while meeting the standard spectral requirements. This work employs a soft clipping technique coupled with digital predistortion such that the overall transmitter output spectrum passes the mask. The distortion in the signal amplitude is then compensated at the receiver side.

The other major topic of the present research thesis is the dimensionality problem in digital-predistorter design. The information criteria have been employed to consider error along with model complexity to estimate the optimum order.

The last research topic carried on in this thesis is related to mitigating numerical instability during system identification of the nonlinear MIMO radio systems suffering from cross-talk. The numerical problem in fixed point processors is resolved using orthogonal memory polynomials. Moreover, a new identification procedure is proposed to reduce computational cost during MIMO digital predistortion identification.

Preface

This thesis is submitted to the Faculty of Graduate Studies in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Electrical Engineering. It reports the work that I have directed in University of Calgary's Intelligent RF Radio Technology Laboratory (iRadio Lab) from January 2010 to December 2013. This thesis has been solely written by me and revised by my supervisor, Dr. Fadhel M. Ghannouchi. Most of the text in Chapters 2-5 is based on publications that have been resulted from my research work.

A version of Chapter 2 has been submitted: M.V. Amiri, S.A. Bassam, M. Helaoui, F.M. Ghannouchi, "Efficient Partition of Distortion Mitigation in Radio Link at the Transmitter and the Receiver," Submitted to the IEEE Transactions on Microwave Theory, August 2013. I conducted the tests and prepared the manuscript. Dr. Ghannouchi and Dr. Helaoui have helped in revising the paper.

A version of Chapter 3 has been submitted: M.V. Amiri, M. Helaoui, F.M. Ghannouchi, "Optimized Soft Clipping Crest Factor Reduction Technique Using Polynomial," Submitted to the IEEE Transactions on Microwave Theory and Techniques, January 2013. I performed the measurements and wrote the manuscript. Dr. Ghannouchi and Dr. Helaoui have provided technical suggestions and contributed in revising the paper.

A version of Chapter 4 has been published: M.V. Amiri, S.A. Bassam, M. Helaoui and F.M. Ghannouchi, "New order selection technique using information criteria applied to SISO and MIMO systems predistortion", International Journal of Microwave and Wireless

Technologies, 5, pp 123-131, March 2013. I accomplished the practical tests and drafted the manuscript. In addition to their technical advice, Dr. Ghannouchi and Dr. Helaoui have also helped in revising the paper.

Versions of different parts of Chapter 5 have been published in one journal and two conferences: M.V. Amiri, S.A. Bassam, M. Helaoui, F.M. Ghannouchi, "Matrix-based orthogonal polynomials for MIMO transmitter linearization," Computer Aided Modeling, Analysis and Design of Communication Links and Networks (CAMAD), 2010 15th IEEE International Workshop on , vol., no., pp.57-60, 3-4 Dec. 2010.

M.V. Amiri, S.A. Bassam, M. Helaoui, Fadhel M. Ghannouchi, "Estimation of crossover DPD using orthogonal polynomials in fixed point arithmetic," AEU - International Journal of Electronics and Communications, Volume 67, Issue 11, November 2013, Pages 905-909, ISSN 1434-8411.

M.V. Amiri, M. Helaoui, Fadhel M. Ghannouchi, "Streamlined MIMO Cross-Over Digital Predistortion," IEEE Radio and Wireless Symposium (RWS), 2013, accepted.

I conducted the measurements and practical tests and drafted the manuscript. Dr. Ghannouchi and Dr. Helaoui have contributed with technical inputs and also helped in revising the papers.

Acknowledgements

Moments are passing fast and will take everything with itself but my good memories in iRadio Lab. remain unchanged and help me find my way to the future.

The Intelligent RF Radio technology Laboratory (iRadio Lab.) was founded by Dr. Fadhel Ghannouchi, CRC Chair (Tier 1) and iCORE Professor in May 2005 in Schulich School of Engineering, University of Calgary.

It was a great pleasure and opportunity for me to work in iRadio Lab. First and foremost, I would like to express my deepest and sincere gratitude to my supervisor Prof. Fadhel Ghannouchi who provided me the chance to work in his group. There is no proper word to express my thankfulness to him. It is my honor to be one of his students and work under his supervision.

In addition to his great and distinguished research abilities, knowledge and deep thoughts, I always found him a powerful group leader and patient manager. In the shadow of his efforts, supports, his great instructive critiques and advice, I learned how to advance my PhD research. His role in my life is greater than a supervisor and I never forget his favor in my life.

All my colleagues in the iRadio Lab. supported me in my research and I appreciate all of them but Dr. Helaoui has a notable role in this dissertation. His great advice helped me all along my research and PhD program. By working with him, I learned how to evaluate the ideas and criticize research works. I express my special thanks to him.

My wife Samira, to whom this thesis is dedicated, has made my memorable moments in the past, motivation for the present and hope towards the future. Hereby, by these short

sentences, I would like to express my deepest appreciation, longest lifetime love and widest respect to her.

I wish to highly appreciate my family in Iran, my mother, brother, sister and her family who always encouraged me during my life to learn more and stay hopeful. I express my highest gratitude to my father, mother and sister in-law who always supported me even in the hardest moments.

Mehdi.

To my love Samira

Table of Contents

List of Tables

List of Figures and Illustrations

List of Symbols, Abbreviations and Nomenclature

SDR	Software Defined Radio
1G	First Generation
AMPS	Advanced Mobile Phone Services
NMT	Nordic Mobile Telephone
TACS	Total Access Communication System
GSM	Global System for Mobile Communications
D-AMPS	Digital-AMPS
CDMA	Code-Division Multiple Access
PDC	Personal Digital Cellular
WCDMA	Wideband CDMA
TD-SCDMA	Time-Division Synchronous CDMA
QoS	Quality of Service
3GPP	Third Generation Partnership Project
LTE	Long Term Evolution
LTE-A	LTE-Advanced
MIMO	Multiple Input Multiple Output
OFDM	Orthogonal Frequency-Division Multiplexing
QAM	Quadrature Amplitude Modulation
PSK	Phase Shift Keying
ISI	Inter-Symbol Interference
RF	Radio Frequency
PA	Power Amplifier
DPD	Digital Predistortion
NMSE	Normalized Mean Square Error
EVM	Error Vector Magnitude

ACPR	Adjacent Channel Power Ratio
ACLR	Adjacent Channel Leakage Ratio
PAE	Power Added Efficiency
BER	Bit Error Rate
SNR	Signal to Noise Ratio
MP	Memory Polynomial
GMP	Generalized Memory Polynomial
RBFNN	Radial Basis Function Neural Networks
PAPR	Peak to Average Power Ratio
Co-DPD	Cross-over DPD
FPGA	Field Programmable Gate Array
OP	Orthogonal Polynomials
LU decomposition	Lower and Upper triangular decomposition
CDF	Cumulative Density Function
AWGN	Additive White Gaussian Noise
CFR	Crest Factor Reduction
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
ML	Maximum Likelihood
PBO	Power Back Off
VSA	Vector Signal Analyzer
GPIB	General-Purpose Interface Bus
DUT	Device-Under-Test
LUT	Look Up Table
ADS	Advanced Design System
DAC	Digital to Analog Converter
CLF	Clipping and Filtering
PDF	Probability Density Function
SCCFR	Soft Clipping Crest Factor Reduction
AR	Autoregressive

MA	Moving Average
ARMA	Autoregressive Moving Average
MAP	Maximum A Posteriori
KL	Kullback–Leibler
MMP	Matrix Memory Polynomial
MIMOAIC	MIMO AIC
ACEPR	Adjacent Channel Error Power Ratio
LS	Least Squares
DSP	Digital Signal Processor
RLS	Recursive Least Squares
SVD	Singular Value Decomposition
MIMOBIC	MIMO BIC
ADC	Analog to Digital Converter
DPoD	Digital Post-Distortion
MGD	Modified Gradient Descent
SA	Simulated Annealing

We should stay,

Learn how to live,

Teach how to live,

This is the way to find the glory of love.

Chapter One: Introduction

Nowadays, wireless communications have become an unavoidable part of our life. Billions of mobile handsets and base stations are being shipped yearly all around the world [1]. The industry surveys report a high volume of shipments of the latest technology wireless devices. Apart from the statistics, the importance and the growth of the applications of wireless technology can be intuitively felt in everyone's daily life.

To support the increasing demand for wireless communications and to provide better services to the customers, the designers have to consider some trade-offs. The trade-offs of digital communication can be summarized as follows: to maximize the transmission rate, to minimize probability of error (maintaining the signal quality), to minimize the required power (increasing the range), to minimize the required bandwidth, to maximize link reliability (providing reliable service to maximum number of users with minimum delay and with maximum robustness to interferences) and to minimize system complexity, cost and computations [2].

Along with the above mentioned goals, with the advent of new applications for wireless communications and the emergence of environmental considerations, more constraints need to be considered in the design of digital communication systems.

Among these constraints, power consumption and efficiency play integral roles. To elaborate more, excluding the power consumed by users, a typical mobile network in the United Kingdom consumes around 40-50 MW. To power the mobile network, the operator has to burn more than one million gallons of diesel per day which in turn will generate large amounts of heat and CO_2. The green radio program has been developed to control the power consumption in

wireless networks. To meet the considerations of green radio communications, power efficiency of the wireless transceivers has to be maximized to avoid CO_2 emissions and heat dissipation [3]. With the evolution of wireless communications, a variety of radio access networks have been optimized to fit targeted applications. Subsequently, different standards with their own carrier frequency, bandwidth and modulation have emerged. This phenomenon has motivated the industry to look for a solution for access in different situations. Multi-standard, multi-band, multi-user systems and software defined radios (SDR) have been found as reasonable solutions to meet these requirements [4].

To achieve the above mentioned goals, digital communication design has evolved over the years. The development of first generation (1G) mobile systems dates back to the1970s and 1980s and includes Advanced Mobile Phone Services (AMPS), the Nordic Mobile Telephone (NMT) system and the Total Access Communication System (TACS) employing analog communications to transfer voice. During the1980s and 1990s, the 2G systems were released to the market which introduced digital technology such as Global System for Mobile Communications (GSM), Digital-AMPS (D-AMPS), Code-Division Multiple Access (CDMA) and Personal Digital Cellular (PDC). The 3G mobile system featuring high data rate was developed in the 1990s and is still under modification. 3G mobiles include three major standards wideband CDMA (WCDMA), time-division synchronous CDMA (TD-SCDMA) and cdma2000. The maximum data rate of this generation was 144 kbps for high mobility traffic, 384 kbps for low mobility traffic and 2 Mbps in proper conditions. The main constraints with 3G mobile are extending the data rate to 100 Mbps using CDMA and the difficulty in multi-rate services with different quality of services (QoSs). These limitations in 3G led to a new network called 4G which provides high data rates coupled with open network architecture. The main feature of 4G

is to provide reliable communication with high data rates ranging from 100Mbps for highly mobile applications and 1 Gbps for stationary applications. This architecture could improve spectral efficiency up to 10 bps/Hz [5]. Based on the Third Generation Partnership Project (3GPP) Long Term Evolution (LTE) technology, the 4G broadband mobile system has found a total of 32 million subscribers and it is expected to grow to two billion users in 2018. Its first evolved version LTE-Advanced (LTE-A), could transmit wideband signals with bandwidth higher than 20 MHz, moreover the carrier aggregation feature could improve the spectrum flexibility. Having targeted the above mentioned specifications, 4G is going to employ bandwidth efficient modulations, Multiple Input Multiple Output (MIMO) and orthogonal frequency-division multiplexing (OFDM) techniques to realize systems meeting these specifications [6].

The 5G terminals employ software defined radios and modulation schemes along with error-control schemes that are available in the internet. In 5G, handling user-mobility is the responsibility of each network, while the final choice will be made by the terminal among different wireless/mobile access network providers for a given service [7].

There is always a trade-off between bandwidth and the required signal-to-noise ratio. In the shadow of quadrature amplitude modulation (QAM), the required bandwidth can be reduced. Compared to M-ary phase shift keying (PSK), QAM provides a notably efficient exchange of the bandwidth in favor of signal-to-noise ratio [2].

Due to the limitations in bandwidth, as a valuable and critical resource, and also the destructive role of propagation, there is always a demand for high spectral efficiency and link reliability. MIMO wireless technology could successfully satisfy these requirements through spatial multiplexing and diversity gain. In spite of the fact that there are some open research

issues with MIMO systems, the technology seem to have enough theoretical and implementation background to enter the practical world [8].

MIMO technology can be employed in wideband systems that suffer from frequency selective multipath fading which in turn translates to inter-symbol interference (ISI). OFDM modulation is a smart way to deal with the ISI problem triggered by the fading problem. In fact OFDM cracks the frequency selective fading channel into a set of parallel flat fading channels and accordingly the equalization process is simplified [8].

As the state of the art techniques appear to resolve the wireless challenges, a number of practical issues arise. A major portion of these practical and implementation challenges deals with the radio frequency (RF) front-end performance. In fact, RF front-ends play an integral role in the overall performance of communication radio systems. A necessary condition for the above cited techniques to be effective is proper design and implementation of RF front-ends [9].

Due to the limitations in practical circuit design, some devices along RF front-ends behave nonlinearly, such that the overall transceiver can be considered as a nonlinear dynamic system. Hence, in addition to linear distortions, nonlinear non-idealities have to be compensated to maintain the quality of the received signal through the communication link.

Among nonlinear devices in RF front-end, power amplifiers (PAs) effectively determine the power efficiency and the linearity of transmitters. The need for PA linearization stimulated a comprehensive research and, accordingly, diverse techniques in the analog and digital domain have been proposed in the literature.

Digital predistortion (DPD) and post-distortion (DPoD) compensation techniques stand for promising ways of PA linearization. DPD and DPoD are actually models that behave as the

inverse of the PA nonlinearity. As a result, the combination of these models with PA should behave as a linear system [9].

This chapter introduces the background research covered in this thesis. First, PA nonlinear behaviour and distortion mechanisms are explained. Later the efficiency-linearity trade-off in designing PA is discussed. Subsequently, the concept and the design challenges of DPD and DPoD are addressed.

The next section of this chapter deals with the problems associated with DPD implementation in mobile terminals. Then new transceiver architectures that resolve these issues are briefly explained. Finally, the research motifs and the organization of this thesis conclude this chapter.

1.1 PA Nonlinear Behaviour and Modeling

For performance evaluation of communication systems, some metrics have been evolved over the years in the literature. These measures quantify the common performance factors such as efficiency, linearity and quality of signal.

1.1.1 Metrics and Some Definitions

For efficiency, there are two major definitions: Drain Efficiency (DE) and Power Added Efficiency (PAE). Drain efficiency is defined as the ratio of output signal power to DC power consumption as follows:

$$\eta = \frac{P_{out}}{P_{DC}}.100 \tag{1-1}$$

PAE on the other hand excludes the input power as below:

$$\eta_a = \frac{P_{out} - P_{in}}{P_{DC}} . 100 \tag{1-2}$$

The linearity measures include Normalized Mean Square Error (NMSE), Error Vector Magnitude (EVM) and Adjacent Channel Power Ratio (ACPR) which can be used interchangeably with Adjacent Channel Leakage Ratio (ACLR) or Adjacent Channel Error Power Ratio (ACEPR).

The definitions of linearity measures are listed below. If the transmitted signal and the received signal after a nonlinear channel, such as PA, are denoted by x and y respectively and $E(.)$ shows the expectation operator, then the NMSE can be defined as follows [10]:

$$NMSE(dB) = 10log_{10}\left(\frac{E(|x-y|^2)}{E(|x|^2)}\right) \tag{1-3}$$

in which it is assumed that the output signal is normalized to the small signal gain. If $x_{s,r}$ represents the transmitted baseband symbols and $y_{s,r}$ represents the received baseband symbols then error power and signal power ($P_{err}$and P_x) can be defined as below:

$$P_{err} = E\left(\left|x_{s,r} - y_{s,r}\right|^2\right) \tag{1-4}$$

$$P_x = E\left(\left|x_{s,r}\right|^2\right) \tag{1-5}$$

The EVM can be obtained in dB or in percentile as follows:

$$EVM(dB) = 10\ log_{10}\left(\frac{P_{err}}{P_x}\right) \tag{1-6}$$

$$EVM(\%) = \sqrt{\frac{P_{err}}{P_x}} . 100 \tag{1-7}$$

If $S_{yy}(f)$ represents the Power Spectral Density (PSD) of y, f_c, $f_o$and BW denote the carrier frequency, offset frequency and signal bandwidth, then the ACPR can be formulated as follows [10]:

$$ACPR\ (dBc) = ACLR(dBc)$$
$$= 10\ log_{10}\left(\frac{\int_{f_c-f_o-BW/2}^{f_c-f_o+BW/2} S_{yy}(f)df + \int_{f_c+f_o-BW/2}^{f_c+f_o+BW/2} S_{yy}(f)df}{\int_{f_c-BW/2}^{f_c+BW/2} S_{yy}(f)df}\right) \quad (1\text{-}8)$$

For evaluating the quality of signal or link throughput, the number of erroneous bits versus signal-to-noise is the most commonly used performance measure. Bit error rate (BER) is the ratio of corrupted bits to the total number of bits and signal-to-noise ratio (SNR) is defined as the ratio of signal power to noise power:

$$SNR\ (dB) = signal\ power\ (dB) - noise\ power\ (dB) \quad (1\text{-}9)$$

Among the components in the communication link, this thesis is focused on the most significant source of distortions which is the PA.

1.1.2 PA classes of Operation and Doherty PA

As shown in Figure 1-1, a linear PA includes input and output matching networks, bias networks and a transistor [10]. The first step in PA design is to stabilize the transistor in the desired frequency band by attaching a stabilizing network to the transistor. The matching networks are usually designed to provide maximum power for the load at the output of the PA. The bias networks provide DC currents and voltages for the transistor to generate amplification gain [12]. PAs are classified based on their bias conditions. The seminal parameter that determines PA class of operation is conduction angle which is mainly a function of gate bias voltage. For a transistor with pinch-off voltage, V_P, bias voltage, V_B, and maximum forward voltage, V_φ, the conduction angle, θ, for maximum swing can be calculated as follows [12]:

$$\theta = acos\left(\frac{V_P - V_B}{V_\varphi - V_B}\right) \tag{1-10}$$

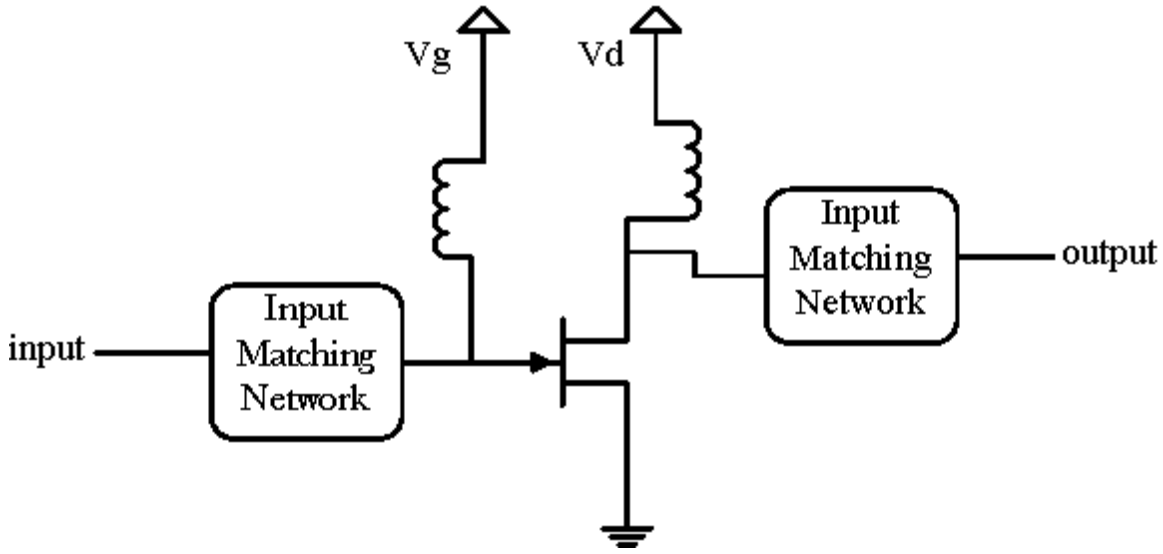

Figure 1-1 General simplified block diagram of a single-ended amplifier [11]

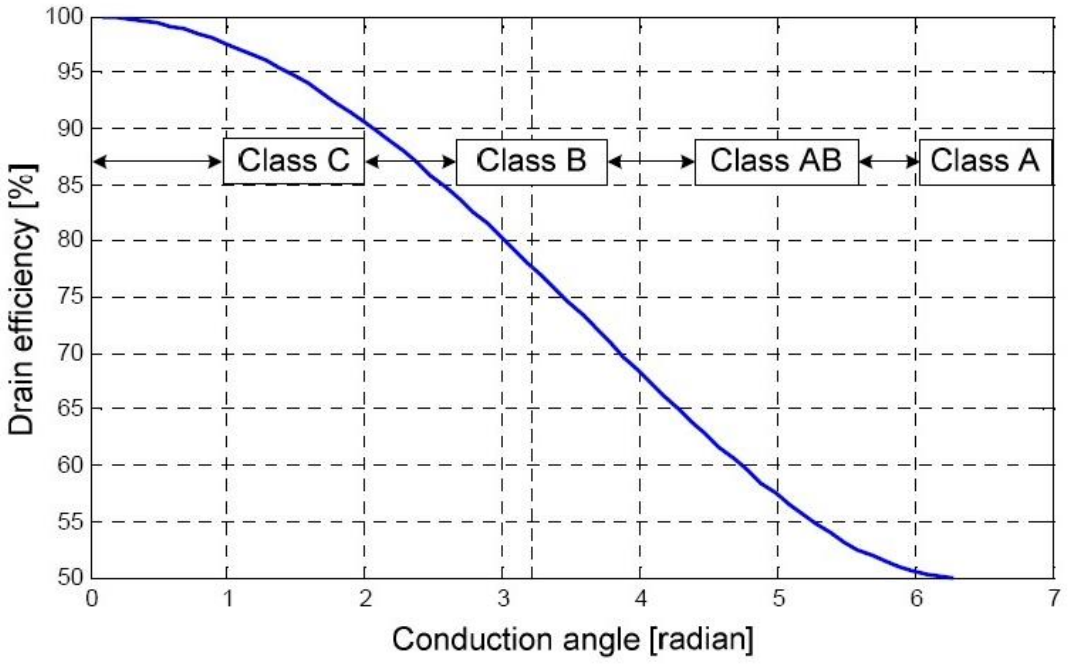

Figure 1-2- Maximum drain efficiency versus conduction angle [11]

Table 1-1- Power amplifier classes of operation

Class Operation	C	B	AB	A
Conduction Angle	$\theta < \frac{\pi}{2}$	$\theta = \frac{\pi}{2}$	$\frac{\pi}{2} < \theta < \pi$	$\theta = \pi$

Table 1-1 lists classes of operation versus conduction angle. Actually, classes other than A are considered as efficiency enhancement classes. Figure 1-2 shows maximum efficiency of the above mentioned classes as a function of the conduction angle [12].

Doherty PA on the other hand combines powers of one class AB (Carrier) and one class C PA (Peaking). The class AB PA works continuously, but the class C PA works only during the 6dB high power part of the signal. Over this last 6 dB range, the theoretical efficiency of Doherty PA is close to the maximum of the carrier and the peaking amplifiers efficiencies.

Hence, the efficiency of Doherty PAs shows an improvement compared to their conventional class AB counterparts [12].

There are two main categories of circuits: linear or nonlinear and memory or memory-less [14]. Power amplifiers are categorized as nonlinear memory circuits. Nonlinear memory systems include nonlinear and memory introducing components. Figure 1-3 depicts the input-output of linear and nonlinear systems.

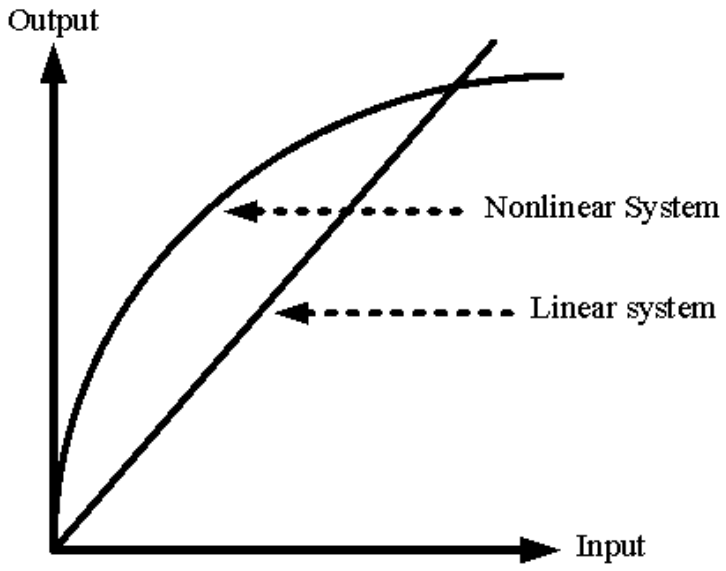

Figure 1-3- Linear and nonlinear systems [14]

On the other hand, memory effect can be assigned to energy storage elements. Memory effect is considered as the dependence of the output signal on the past values in a linear or nonlinear way [14].

Power amplifiers generally include memory nonlinear components. Accordingly, the I/V curves of PA become nonlinear due to the nonlinear components and dynamic load lines of PA show hysteresis due to memory effect [14].

The nonlinear behaviour of PA is classified into weak and strong cases. The small signal PA exhibit nonlinear behaviour around the bias point and is called weak nonlinearity. This nonlinear behaviour can cause different effects on signals, such as AM-AM compression, AM-PM conversion, harmonic generation and intermodulation distortions [15].

There exist generally three main categories of PA characterization models: physical models, equivalent circuit models and behavioural modeling. Most of the CAD tools use second and third trends for modeling purposes [12]. This dissertation mainly deals with the behavioural modeling of PA, which has been used as a basic and powerful tool for designing linear efficient transmitters. Weak nonlinearities have been reported to be modeled effectively by Volterra-based models [12]. For modeling the strong nonlinear behaviour of PAs, which is caused by their limiting behaviour (saturation)[12], several models have been reported in the literature. Behavioural models of PA include static models, such as look up tables, static polynomials and dynamic models like Wiener, Hammerstein, Memory Polynomial (MP), Generalized Memory Polynomial (GMP) and Radial Basis Function Neural Networks (RBFNN). Inherently, PA models can be categorized into three major families: Volterra-based, box-based and neural-based models. Recent comparative studies [16] indicate that box-based Wiener and Hammerstein models perform close to static models and hardly can model the dynamic nonlinear memory effects of PAs. On the other hand, MP, GMP and RBFNN show superior performance in terms of NMSE and ACPR. MP also favors a higher robustness in validation tests [16]. GMP results in a better performance at the cost of more complexity. Consequently MP exhibits the best

performance (NMSE and ACPR) with the lowest complexity [17], and therefore, is the most commonly used model in practice.

1.2 Efficiency and Linearity Trade-off

The efficiency of PAs plays a key role in transmitter design. The amount of output signal power to the DC consumption has been defined as the Drain Efficiency (DE) of the PA. This is the efficiency that determines the battery life of a cell phone or any portable device.

One of the draw-backs of modern communication signals is their high peak to average power ratio (PAPR). The PAPR of a signal is defined as the ratio of peak power to mean power of the signal measured in dB. This high PAPR increases the amount of power back-off needed for driving PAs. The higher back-off means lower mean power which in turn degrades the efficiency [18]. As can be seen from Figure 1-4, the measurement results confirm that the average efficiency is a decreasing function of the PAPR. There are some models describing the average efficiency behaviour as a function of PAPR. In [19] an approximate model has been proposed to formulate this behaviour:

$$\eta = \eta_{peak} exp(-g.PAPR) \tag{1-11}$$

where η_{peak} stands for the peak efficiency and g stands for a constant that depends on PA class. Therefore, it can be concluded that PA efficiency is highest at peak power when PA works in its nonlinear region. On the other hand, gain compression of PA due to saturation phenomenon introduces distortion to signal. Accordingly, PA linearization is an effective and promising way for improving efficiency while maintaining the quality of signal [9].

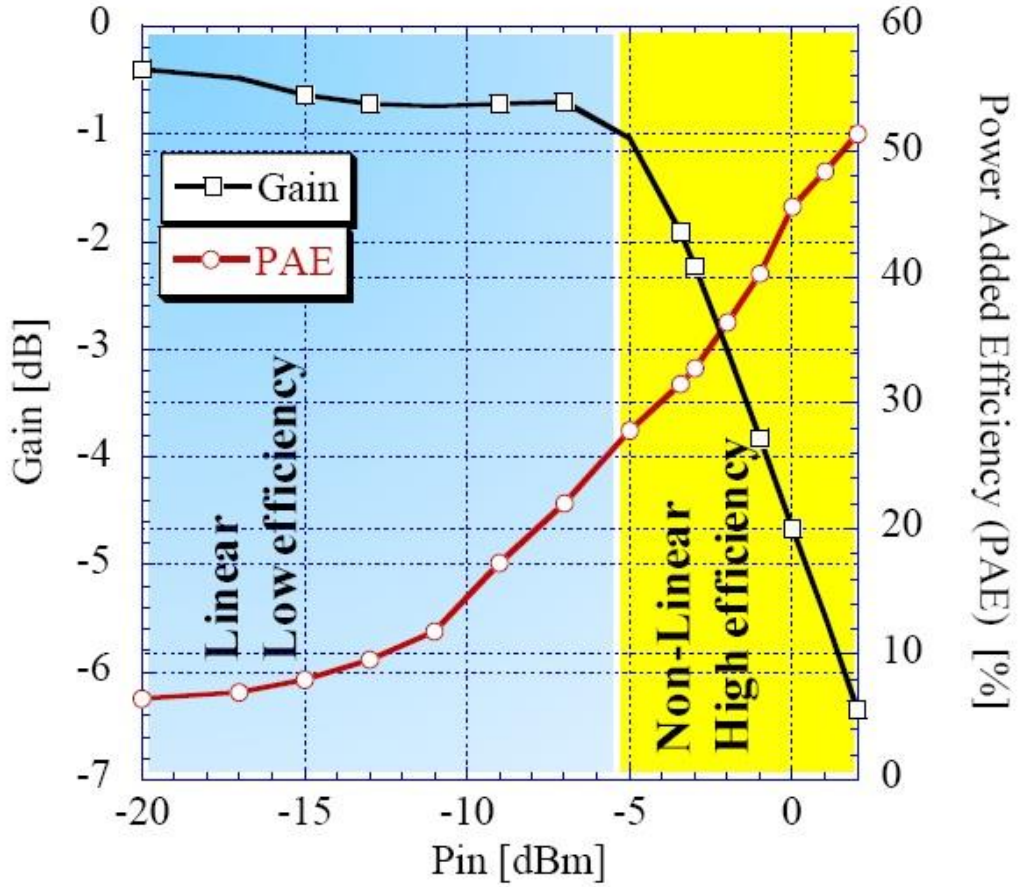

Figure 1-4- Measured gain and efficiency variation versus input power of a transmitter with a typical class AB PA [11]

1.3 Digital Predistortion (DPD)

Digital predistortion is one of the favorable techniques for linearizing PAs. Through linearization, gain and phase compressions of a PA are compensated and the linear region of the PA is extended. Subsequently, efficiency improves without distorting the signal [9].

The first step in designing a DPD is to select a model. As described above, DPD can be realized in a variety of models. Among them, MP has been found to be more reliable. It has the ability to capture the nonlinear dynamics of a PA [9].

The MP model of a PA, with x and y denoting the discrete time input and output signals respectively, can be described as below [9]:

$$y(n) = \sum_{i=1}^{K} \sum_{m=0}^{Q} (a_{im} + jb_{im}) x(n-m) |x(n-m)|^{i-1} \qquad (1\text{-}12)$$

where K, Q and $(a_{im} + jb_{im})$ are the nonlinearity order, the memory depth and the complex parameters of the MP model, and $|x(n-m)|$ is the envelope of the input signal to the PA at $n-m$ time step.

The main task of a DPD is to compensate for the AM-AM and AM-PM distortions of a PA. To accomplish this task the DPD should behave as the inverse of the PA gain compression such that the overall function of DPD+PA becomes a linear system [9].

1.4 DPD Challenges and the Organization of the Dissertation

As discussed earlier, the PA efficiency is one of the important concerns in radio communications. The first issue in the application of complex DPD in mobile cell phones, for the up-link communications, is its computational cost which usually requires more power consumption by digital signal processors and additional power back due to higher PAPR of the driving signal resulting from the complex DPD. These effects increase the power consumption of the mobile terminal and make complex DPD implementation less attractive for the up-uplink application.

In chapter two, a novel transceiver architecture for linearization and efficiency enhancement purposes is proposed. This architecture calls for it joint-deployment deployed and partition between the base station and the mobile terminal. In the proposed new architecture, the phase nonlinearity is compensated first at the transmitter of the mobile terminal using phase-only DPD which does not change the PAPR of the signal and, the amplitude nonlinearity will be compensated at the base station using CDF estimation. Consequently, the overall efficiency

improves. The distortion-free amplitude of the signal is then recovered at the receiver using CDF estimation, which does not rely on a training sequence. As a result, the efficiency of the transmitter shows a notable improvement compared to DPD and conventional back-off techniques.

From the handset vendor's point of view, this technique improves the efficiency of the mobile terminal transceiver and increases the battery life of mobile handsets. The other important issue is that this technique does not require additional hardware since it is fully software based it can be implemented in the existing processors of the handset.

From the network provider's point of view, this technique will require the additional implementation of a relatively simple and low complexity algorithm to estimate the CDF of the received signal and compared with distortion free signal CDF. This technique does not rely on a training sequence and feedback loop for compensating the PA distortion of the handset terminal.

As this technique does not need any training sequence, accordingly, the latency of the network does not degrade. The CDF is estimated using a given frame and distortion compensation is applied for the subsequent frame. Herein we assume that the characteristic of the PA of the mobile terminal does not change in significant manner within two successive frames. . In terms of hardware, the proposed partitioned compensation method does not need any feedback loop. This aspect of the technique is attractive, since, in the full adaptive complex digital pre-distortion technique a feedback loop for PA characterization is needed this will add a complexity burden to the system. This work was submitted to the IEEE Transactions on Microwave Theory and Techniques [P3].

As a matter of fact, PA gain starts compressing before saturation. Consequently, DPD expands peaks such that the compression transforms them to their initial linear desirable

condition. This expansion in amplitude increases the PAPR of the signal applied to the PA. The higher PAPR of the pre-distorted signal increases the required power back-off and limits the amount of improvement in efficiency using DPD. This problem is the second issue which is considered in this thesis.

In chapter three, a new crest factor reduction (CFR) technique is presented which is then combined with full DPD to optimize the PAPR and hence the efficiency while meeting the spectral constraints. The algorithm uses a post-distortion compensation approach to recover the amplitude nonlinearity of the CFR at the receiver. This technique can be employed by base stations to improve the efficiency while passing ACPR and EVM requirements. The soft clipping crest factor reduction technique which is recommended for base stations reduces the PAPR of the signal in addition to the improvement in efficiency and signal quality. As a result of the PAPR reduction due to soft clipping, the peak power requirement of the base station PA decreases. Hence less powerful PAs can be used for the same base station while keeping the same coverage. This can be very beneficial in reducing the cost of deployment of the network (capital expenses) as well as the operating expenses for the network operator. This work was also submitted to the IEEE Transactions on Microwave Theory and Techniques [P4].

The third issue regarding DPD is the dimensionality of the model used in DPD design. Conventional techniques are established based on sweeping the orders and memory taps to find saturation in the NMSE trend. This technique lacks an optimum point and is computationally inefficient.

In chapter four, new model order selection approaches are proposed which are based on information criteria. Using these criteria, the complexity of the model is also considered in model order selection. This means that there is a minimum in the dimensions space that can be

estimated using mixed integer nonlinear programming techniques. This approach employs an optimization which can be performed much faster than the brute force dimension sweeping technique along with NMSE monitoring. This work was published in the International Journal of Microwave and Wireless Technologies [P1].

The rest of this thesis deals with nonlinearity in MIMO systems and specifically nonlinear cross-talk. Due to practical limited isolations in integrated MIMO PAs, it is probable that the inputs cross-talk with each other by a coupling factor. If this phenomenon happens at the inputs of the chip, then the cross-talk will be considered nonlinear. This type of cross-talk can be compensated using Cross-over DPD (Co-DPD) [20]. The first problem arising in the Co-DPD implementation occurs when the precision is limited, i.e., in fixed point arithmetic in processors such as FPGAs. The numerical problems associated with the conventional MP matrix inversion result in poor estimate of the DPD coefficients and cause a divergence in the algorithm for higher nonlinearity orders.

Chapter five shows the superior performance of orthogonal polynomials (OP) over MP is shown. The coefficients of OP are then extracted using LU decomposition and triangular matrix inversion. This technique converges even for high nonlinearity orders. This work was published in the International Journal of Electronics and Communications and IEEE CAMAD 2010 conference [P2], [P5].

The last issue concerning Co-DPD is that the estimation algorithm seems inefficient. In chapter five a new streamlined technique is proposed. This technique has the same performance in terms of the amount of distortions, while taking around half of the computational burden. This work was accepted for publication in the RWS 2014 conference [P6].

1.5 Research Focus

Due to the non-ideal behaviour of components in the transmission and reception chain, the overall system can be considered a nonlinear dynamic system. Consequently, to avoid signal degradation and maintain an acceptable signal quality, compensation techniques have to be considered and included in the chain.

This dissertation mainly proposes advanced statistical signal processing techniques to compensate the nonlinearities in modern communication systems. The main contributions of this thesis can be summarized as follows:

- A partitioned compensation technique is proposed to compensate the PA nonlinear AM-AM and AM-PM distortions. The compensation is distributed between transmitter and receiver. The phase and amplitude nonlinearities are compensated at the transmitter and the receiver sides respectively. The advantages of the technique can be considered as an improvement in efficiency and eliminating the need for training sequences. The proposed technique can easily be implemented in cell phones [P3].

- It is widely known that DPD increases the PAPR which subsequently limits the efficiency improvement. A new CFR technique is proposed in this thesis which is then combined with DPD, such that maximizes the efficiency subject to spectral mask constraints. This technique is useful for base stations [P4].

- One of the major problems in DPD design is the model orders. To overcome this problem, a new method is proposed in this thesis which is based on the Akaike and Bayesian Information Criteria (AIC and BIC) [P1].

- MIMO systems also suffer from nonlinearities of integrated PA chips. The other issue is the numerical problems in the DPD implementation in fixed point arithmetic. This thesis proposes the application of orthogonal polynomials and LU decomposition in DPD design [P2], [P5].

- The nonlinear cross-talk adds distortions to MIMO systems and has been shown to be compensated by Co-DPDs. This dissertation proposes a simplified estimation technique that can be used in Co-DPD estimation [P6].

Chapter Two: Linearization of mobile stations using partitioned compensation technique

2.1 Introduction

The performance of a communication system is notably limited by the amount of distortions produced by non-ideal behaviour of the devices in the transceiver chain. The nonlinear distortions caused by the transmitter, and in particular PA, can be considered as a major problem in signal transmission.

As far as the power efficiency is concerned, it is necessary to drive the PA into its compression region, which in turn introduces considerable distortions to the signal. Consequently, to maintain an acceptable BER at the receiver end, the use of linearization techniques seems to be inevitable for compensating these distortions while preserving high power efficiency [21]. DPD [10] and post-compensation at the receiver side [22], stand for the most commonly used linearization techniques. When no linearization technique is applied, it is mandatory to consider a power back-off at the transmitter for the BER considerations [23], assuming that the channel equalization is perfect at the receiver side and that the receiver is fully linear.

The DPD consists of applying a gain, with inverse amplitude and phase behaviour than the complex gain behaviour of the PA, to the signal which precedes the PA. The overall behaviour of the cascade of the predistortion function and the PA is, therefore, linear versus input power.

As a result of compensating for the gain compression of the PA near saturation, the signal PAPR increases [21]. Due to the limitation in the maximum power of the input signal, an increase in the PAPR forces a decrease in the mean input power (typically 2-4 dB). Therefore,

this power back-off degrades the drain efficiency of the PA [21], [24]. On the other hand, DPD performance degrades notably by the impairments in the up and down-conversion circuits (the feedback path components such as mixers, filters, quadrature modulator and demodulator). These impairments highly result in a poor estimation of the inverse function of the PA behaviour [25], [26].

In the case of techniques other than DPD, it is compulsory to consider training sequences to characterize the PA nonlinearity [22]. In one of these approaches, a training sequence is placed in each frame and is sent to the receiver. Firstly, the effect of the channel is removed by equalization techniques. The resulting signal is then used as the replica of the signal at the output of the PA for DPD estimation. Based on the training sequence data known at the receiver side, the DPD function is then extracted. The model is then passed to the transmitter for compensation [22].

The limitation in the efficiency improvement is the common disadvantage among all current predistortion techniques, since the signal predistortion precedes the power amplifier, which results in an increase in the PAPR of the signal.

This chapter proposes a new linearization technique that compensates for the transmitter amplitude nonlinearity in the receiver side in order to maintain high power efficiency in the PA. Cumulative distribution function (CDF) based algorithm is employed to estimate the PA AM-AM nonlinearity. The estimation for nonlinearity using CDF has been limited in previous works to predistortion in the transmitter side [27], [28] and amplitude-only compensation techniques [29], [30]. This chapter combines the transmitter phase-predistortion idea along with the receiver CDF-based amplitude post-compensation technique to maintain the improvement in overall efficiency and result in an acceptable BER performance of the transceiver in a radio link.

In this chapter first, the motivation of compensating for the amplitude nonlinearity in the transmitter and the concept of the new distributed distortion compensation approach is explained. Then the details of the theoretical analysis of the proposed approach are discussed. The simulation and measurement results and the performance of the proposed linearization technique are provided.

2.2 Distributed Distortion Compensation

In systems deploying standards such as WCDMA, LTE and WiMAX, the handset PAs are usually operated in quasi-linear region. They are mostly designed to be linear at specific back-off power [31].

To avoid nonlinear distortions in the transmission of high PAPR signals, PAs are usually biased at class A or class AB. Furthermore a large back-off from the saturation power should be considered for the PA to work linearly [12]. This large back-off causes degradation in the power efficiency improvement and hence worsens the heat dissipation.

Nonlinear distortion in the cell phone can be compensated using digital signal processing. The nominal power consumption of these processing modules may exceed 1 Watt [32]. Base stations on the other hand consume around 85% of the total energy of the communication network and their power consumption depending on the size, coverage and technology and ranges from 147 Watts to 10 KWatts [33],[34]. Hence implementing the PA nonlinearity compensation at the base station will rarely impact the overall efficiency of the network markedly. The gain expansion of DPD increases the PAPR and limits the efficiency achievement of the linearization. Conversely, the phase nonlinearity compensation by itself does not affect the PAPR, and therefore does not degrade the efficiency of mobile handset.

Having known the distribution function of the original OFDM signal [35], the amplitude nonlinearity of the PA can be estimated at the receiver side through a comparison between the distribution of the original signal before PA amplification and the received signal.

Doherty PAs show better power efficiency in back-off compared to their class AB counterparts [12]. Nevertheless, for high PAPR signals, they exhibit highly nonlinear behaviour [36].

2.3 Amplitude and Phase Nonlinearity Compensation in the Distributed Distortion Compensation

The main concept of the proposed partitioned distortion compensation technique consists of transmitting the signal through PA amplification after pre-distorting the phase nonlinearity. Provided reasonable amplitude nonlinearity, the transmitted signal should meet the spectral mask requirements.

After equalizing the channel effects at the receiver side, the baseband equalized signal is used to estimate the amplitude nonlinear distortion by estimating the empirical CDF of

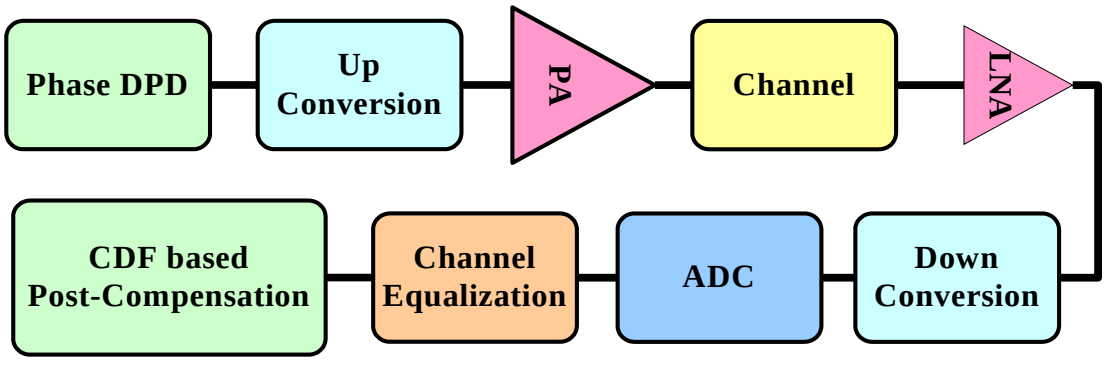

Figure 2-1- The proposed architecture

the received signal amplitude. Then the transmitter amplitude nonlinearity is compensated for at the receiver side. Figure 2-1 depicts the block diagram of the whole transceiver. The spectrum of the received signal with and without Phase-DPD and the spectral mask are shown in

Figure 2-2. It can be seen that in both cases, the transmitted signal spectrum passes the mask requirements.

2.1.1 Amplitude Nonlinearity Estimation

If *x* and *y* denote the complex input and output signals, respectively, of a PA and if r_x and r_y are their corresponding amplitudes, the static PA behaviour can be modeled as:

$$y = F(r_x)\, exp\big(jG(r_x)\big)\, x \tag{2-1}$$

where *F* and *G* are the amplitude and phase compression functions of the PA, respectively, which depend on the amplitude of the input signal.

If the input signal distribution is known, then by estimating the empirical CDF of the output signal, the amplitude nonlinearity can be obtained.

If F_1denotes the nonlinear function relating the amplitude of the output signal to the amplitude of the input, then:

$$r_y = rF(r) = F_1(r) \tag{2-2}$$

where, for simplicity, *r* is used instead of r_x.

If $F_y(.)$ and $F_x(.)$ denote the CDFs of the output and input signals respectively, It is established from [37] that the following relationship holds between the distribution of the output amplitude and the input signal distribution

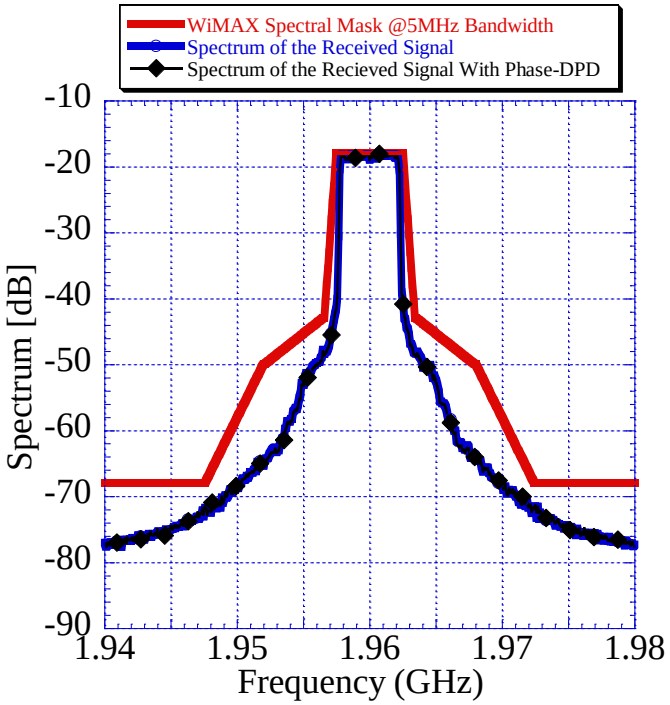

Figure 2-2- Received signal spectrum and spectral mask for WiMAX

$$F_y(r_y) = F_x(r) \tag{2-3}$$

Using (2-2) and (2-3), the nonlinearity can be estimated as follows:

$$r = F_x^{-1}(F_y(r_y)) \tag{2-4}$$

$$F_1{}^{-1}(r_y) = F_x^{-1}(F_y(r_y)) \tag{2-5}$$

In [35] it has been shown that most of OFDM based signals follow a Rayleigh, a generalized Rician, or Weibull distribution.

In the following, the equations for the Rayleigh, Rician and Weibull distributions are given [32, 34, 35]. The Rayleigh distribution is given by:

$$f(r|b_R) = \frac{r}{{b_R}^2} exp\left(\frac{-r^2}{2{b_R}^2}\right) \quad r > 0\,, b_R > 0 \tag{2-6}$$

in which $b_R$is the scale parameter of the distribution, The Rician distribution is given by:

$$f(r|a_I, b_I) = \frac{r}{{b_I}^2} exp\left(\frac{-(r^2 + {a_I}^2)}{2{b_I}^2}\right) I_0\left(\frac{r a_I}{{b_I}^2}\right)$$
$$r > 0\,, b_I > 0, a_I \geq 0 \tag{2-7}$$

where a_I accordingly, is an indication of the distance between the reference point and the center of the bivariate distribution and b_I is the scale parameter, and $I_0(.)$ denotes the Bessel function of the first kind.

The Weibull distribution is given by:

$$f(r|b_W, c_W) = \frac{c_W}{b_W}\left(\frac{r}{b_W}\right)^{c_W - 1} exp\left(-\left[\frac{r}{b_W}\right]^{c_W}\right)$$
$$r > 0\,, \qquad b_W > 0, \qquad c_W > 0 \tag{2-8}$$

Weibull distribution is a two parameter distribution with $c_W$as the shape and $b_W$as the scale parameters.

In the case of Rayleigh and Weibull assumptions for the input signal, the amplitude nonlinearity can be estimated using the following formulas as provided in [29]:

$$r = {F_1}^{-1}(r_y) = \sqrt{-2{b_R}^2 ln\left(1 - F_y(r_y)\right)} \tag{2-9}$$

$$r = {F_1}^{-1}(r_y) = b_W\left(-ln\left(1 - F_y(r_y)\right)\right)^{1/c_W} \tag{2-10}$$

In the case of Rician distribution, the inverse CDF is hard to find in closed form. Once, the distribution parameters have been estimated for the input signal, the inverse CDF can be evaluated at the signal values by interpolation.

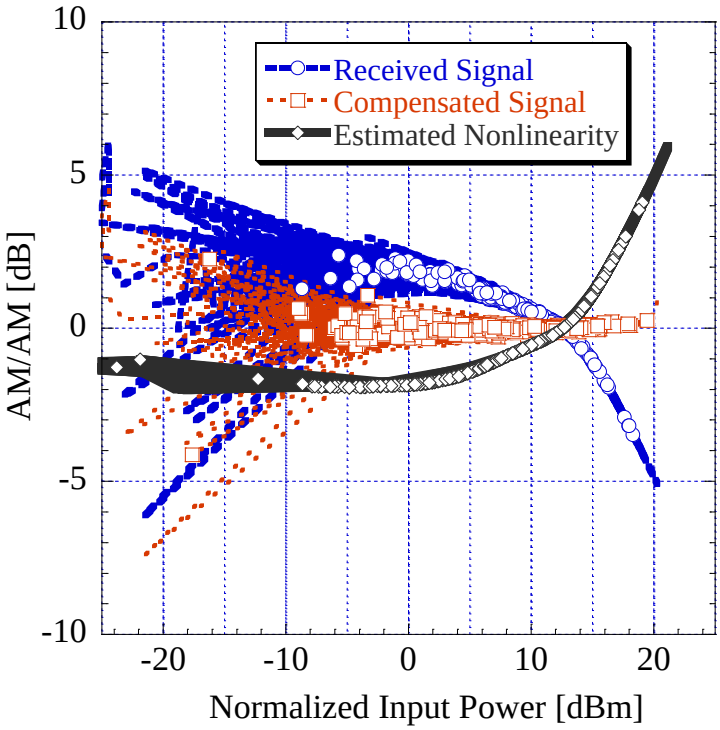

Figure 2-3- The AM/AM of the received and the compensated signal along with the estimated nonlinearity

From (2-4), it can be inferred that the nonlinearity of the amplitude can be estimated by obtaining the CDF of the received signal. The AM/AM of the estimated nonlinearity and the linearized system are shown in

Figure 2-3.

The first step in estimating the AM/AM is to fit a parametric CDF to the transmitted signal amplitude. The Maximum Likelihood estimation (MLE) of the parameters can be obtained through simple averaging of proper statistics. For instance in the case of Weibull distribution, the parameters can be estimated using the following equations. If it is assumed that $r_1, \dots, r_N$ are the

samples of the transmitted frame and $R_i = \ln(r_i)$, then the MLE estimated values (indicated by$\widehat{(.)}$) for the Weibull parameters can be obtained as below:

$$\hat{c}_W = \left[\frac{\sum_{i=1}^{N} r_i^{\hat{c}_W} R_i}{\sum_{i=1}^{N} r_i^{\hat{c}_W}} - \bar{R}\right]^{-1} \tag{2-11}$$

$$\hat{b}_W = \left(\frac{1}{N}\sum_{i=1}^{N} r_i^{\hat{c}_W}\right)^{1/\hat{c}_W} \tag{2-12}$$

where $\overline{(.)}$denotes the average value. The second step in estimating the AM/AM nonlinearity of PA, is to estimate the CDF of the received signal. Empirical methods should be employed to obtain the CDF. Among empirical CDF estimation techniques, Kaplan-Meier has been found simple and accurate [38]. In this approach the survival function is estimated, which is defined as:

$$P(r) = 1 - F(r) \tag{2-13}$$

where $F(r)$ denotes the CDF of random variable r.

To estimate the survival function, the scale is divided to N intervals namely: $(0, r_1) \cdots (r_{N-1}, r_N)$. Then $P(r)$ can be estimated as below:

$$P(r) = \prod_{r_i < r} \left(\frac{n_i - d_i}{n_i}\right) \tag{2-14}$$

where n_i is the number of samples greater than r_{i-1} and $d_i$represents the number of observations that are greater than r_{i-1} but smaller than r_i.

2.3.1 The Phase Nonlinearity problem

The estimation of the phase nonlinearity is not a trivial task. If θ_x and θ_y represent the phases of the input and output signals, respectively, and $G(r)$ is the phase distortion for given amplitude *r*, for the amplifier model represented by (2-1), then:

$$\theta_y = \theta_x + G(r) = \theta_x + \theta_d \tag{2-15}$$

If a uniform phase distribution is assumed for a given input signals level, the output signal phase also has a uniform distribution for which the maximum likelihood (ML) estimate will be zero.

As it has been shown in [39], a complex Gaussian process has a uniform distributed phase. This uniform phase distribution holds for any given amplitude. At constant signal amplitude, the phase distortion according to the complex gain model of the PA will be a constant value. Hence the output signal phase distribution given the amplitude will be uniform and there is no way to estimate the phase compression function through the CDF of phase.

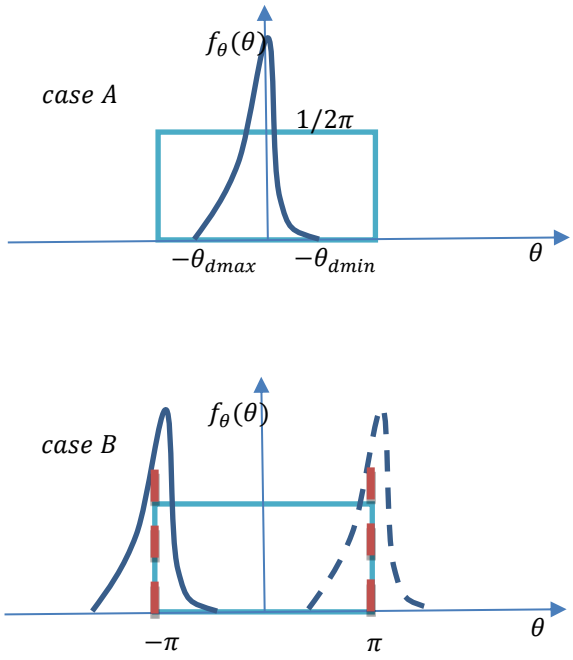

Figure 2-4- The convolution of input phase distribution with the PDF of phase distortion

The PDF of sum of two independent random variables is the convolution of their PDFs. If it is assumed that the phase distortion is distributed between θ_{dmin} and θ_{dmax}, then for the output phase distribution two cases are possible. In the first case which is shown in Figure 2-4 as case A it can be concluded that

$$f_{\theta_y}(\theta_y) = \frac{1}{2\pi}\int_{\theta_{dmin}}^{\theta_{dmax}} f_{\theta_d}(\theta_d)d\theta_d = \frac{1}{2\pi} \tag{2-16}$$

In case B the PDF is the summation of two parts:

$$\begin{aligned} f_{\theta_y}(\theta_y) = \frac{1}{2\pi}\int_{-\pi}^{\theta_y-\theta_{dmin}} & f_{\theta_d}(-\theta_d)d\theta_d \\ & + \frac{1}{2\pi}\int_{\pi+\theta_y-\theta_{dmax}}^{\pi} f_{\theta_d}(-\theta_d)d\theta_d \\ & = \frac{1}{2\pi}\int_{\theta_{dmin}}^{\theta_{dmax}} f_{\theta_d}(\theta_d)d\theta_d = \frac{1}{2\pi} \end{aligned} \tag{2-17}$$

According to the above mentioned reason, it was decided to proceed with phase distortion compensated at the transmitter. This phase DPD does not deteriorate the efficiency since the PAPR does not change and there is no need for a power back-off. Another advantage of the proposed method is that, in general, the phase nonlinearity is less sensitive to the PAPR of the signal as reported in [36]. The phase DPD can be implemented independently from the signal characteristics without being adaptive.

In this chapter the phase nonlinearity is compensated in the transmitter side using a simple LUT [40], [41] which can be estimated using a high PAPR signal and then plugged for other standards and with any PAPR.

As mentioned above, the phase-only DPD does not change the PAPR. According to the complex gain model of the PA which is described in (2-1), the Phase-pre-distorted signal can be represented as follows:

$$z = exp\big(-jG(r)\big)\,x \tag{2-18}$$

Then the PAPR of the pre-distorted signal (z)can be obtained as:

$$PAPR(z) = 10log_{10}\left(\frac{max(z^*z)}{E(z^*z)}\right) \tag{2-19}$$

where $(.)^*$stands for the complex conjugate operator.

Then based on the fact that

$$z^*z = \big(exp\big(-jG(r)\big)\,x\big)^*\big(exp\big(-jG(r)\big)\,x\big) = x^*x \tag{2-20}$$

the PAPR of the signal after phase pre-distortion does not change.

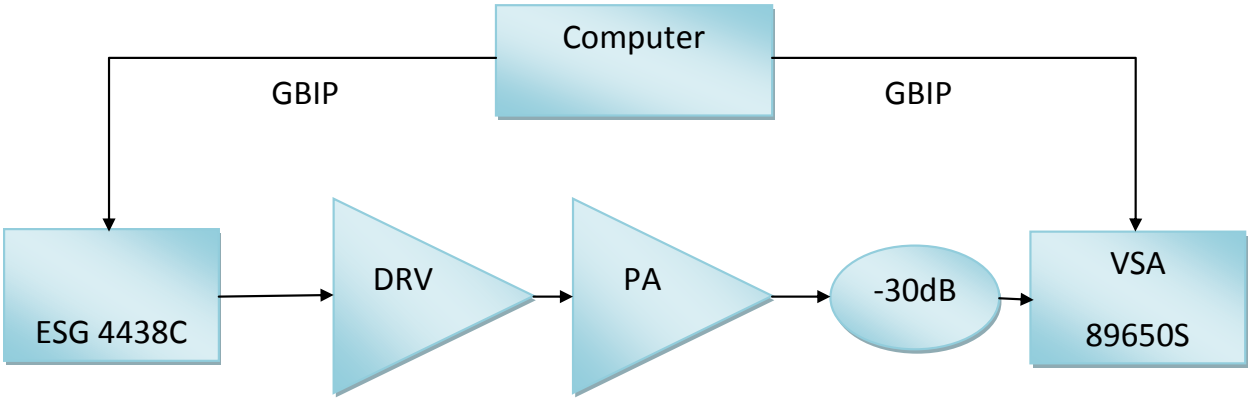

Figure 2-5- Measurement setup

2.4 Measurement Results

2.4.1 Measurement Setup

In this section, the measurement results for different linearization techniques, i.e. DPD, PBO and distributed distortion compensation, are discussed. The performances of these approaches are then compared in terms of DE, NMSE, EVM and ACPR.

The measurement setup, consisting of a vector signal generator, a driver amplifier, a PA, an attenuator and a vector signal analyzer (VSA), is shown in Figure 2-5.

The excitation signals are uploaded to the signal generator using a general-purpose interface bus (GPIB). The specifications of these excitation signals for system identification and evaluation are summarized in Table 2-1.

The first device-under-test (DUT) is composed of a cascade of a class-A driver and a class-AB PA, which is biased at a gate-to-source voltage (Vgs) of 10.3 V and a drain-to-source voltage (Vds) of 28 V. Both amplification blocks were designed using a TF10107 LDMOS transistor.

This PA works at carrier frequency of 1.96GHz. The second DUT includes a class-A driver and a Doherty PA. The Peaking PA was biased at Vgs of -5.5 V and the Carrier PA at -2.8 V. The Vds was set to 28 V. The Doherty PA was designed using CGH40010, 10W GaN HEMT transistors. The center frequency for this PA is 2.425GHz.

The third DUT is a 0.5 Watts PA, proper for mobile applications. The center frequency for this PA was set to 900MHz. The device is biased at Vgs of -1.2V and Vds of 8V. The quiescent current was obtained 118mA. 27.52dBm output power was obtained at a maximum drive of 11.5dBm. The maximum measured CW efficiency was 54.74%.

2.4.2 Results and Discussion

The simulation and measurement results are divided into three different cases. In section 2.4.2.1, the results for an ideal line of sight channel for high power PAs (i.e. the first and the second PAs introduced in 2.4.1) are provided. In section 2.4.2.2, measurement results for the ideal channel and a low power PA (i.e. the third PA) is discussed. In the last section, fading and AWGN noise is applied to the signal after PA and the simulation results of different techniques are presented.

2.4.2.1 Measurements: Ideal Channel case, high power PAs

The signals in Table 2-1 have been used in these measurements to evaluate the performance of the proposed CDF-based distributed distortion compensation technique. The performance of the proposed technique is then compared to DPD and back-off methods.

In this subsection the channel is assumed to be ideal and the DUT is the class AB PA. Five different cases are compared using six different signals.

Case 1: In the first case no linearization technique is applied.

Case 2: In the second case, the proposed distributed distortion compensation with look up table; In this case, a LUT has been built for each signal using its own measured data (DPD1) based phase pre-distortion along with the CDF-based amplitude compensation is applied.

Case 3: The third case consists of using the proposed distributed distortion compensation technique where the LUT is obtained using the highest PAPR signal (signal number 2) and applied for all six signals (DPD2).

Case 4: In the fourth case a memory polynomial DPD has been used to compensate for the phase and amplitude distortion in the transmitter (Conventional DPD).

Case 5: Case number five consists of backing off the signal to avoid the nonlinear distortions in the compression region of the PA.

Table 2-2 and Table 2-3 summarizes the measurement results for the above mentioned cases. It can be concluded that the distributed distortion compensation technique provides enough accuracy in the compensation (NMSEs are similar to the memory polynomial DPD and back-off techniques), however the efficiency is significantly higher in the proposed distributed distortion compensation technique (nearly the double).

Table 2-1- Parameters of the modulated signals

Signal Number	Standard	PAPR (dB)	Bandwidth (MHz)	Sampling Frequency(MHz)
1	WiMAX	10.2	5	92.16
2	WiMAX	12.7	5	92.16
3	LTE	10.4	5	61.44
4	LTE	12.4	5	61.44
5	WCDMA	9	3.84	92.16
6	WCDMA	10.9	3.84	92.16

Table 2-2- Measurement results for the first 3 cases

Signal Number	Case	DE(%)	NMSE(dB)	Distribution
1	Case1: No Compensation	17.2	-17.8	Ideal channel
	Case 2: DPD1	17.9	-38	Weibull
	Case 3:DPD2	18	-38	Weibull
2	Case1: No Compensation	13.2	-18.7	
	Case 2:DPD1	13.2	-35	Weibull
	Case 3:DPD2	13.2	-35	Weibull
3	Case1: No Compensation	19.6	-16.9	
	Case 2:DPD1	19.3	-42	Weibull
	Case 3:DPD2	19.6	-42	Weibull
4	Case1: No Compensation	15	-17.8	
	Case 2:DPD1	15	-37	Weibull
	Case 3:DPD2	15	-37	Weibull
5	Case1: No Compensation	22.3	-16.5	
	Case 2:DPD1	22.3	-37	Rice
	Case 3:DPD2	22.2	-37	Rice
6	Case1: No Compensation	18	-16.9	
	Case 2:DPD1	18	-37	Weibull
	Case 3:DPD2	17.9	-37	Weibull

Table 2-3- Measurement results for the cases 4 and 5

Signal Number	Case	DE(%)	NMSE(dB)
3	Case 4: MP DPD	11.8	-42
	Case 5: Back off	0.8	-37
4	Case 4: MP DPD	8.4	-42
	Case 5:Back off	0.8	-36
5	Case 4: MP DPD	13.9	-42
	Case 5: Back off	0.8	-35
6	Case 4: MP DPD	9.7	-39
	Case 5: Back off	0.8	-36

It can be observed as well that the phase DPD in the second and the third cases have similar DE and NMSE performance. Therefore, it can be concluded that the phase distortion is relatively insensitive to the signal characteristics and non-adaptive phase predistortion is sufficient to achieve good linearity.

In the second set of measurements, two PAs have been considered. The first PA is the class AB PA used in the previous set of measurement, while the second PA is a Doherty PA as described in the measurement setup. The performance of the proposed partitioned distortion compensation technique in terms of power efficiency and linearity is compared to the PBO technique in practical scenarios for uplink transmissions. Indeed, given that the proposed technique is targeted for uplink applications, where no adaptive DPD is assumed to be used in the transmitter side to simplify the implementation in mobile terminals. Therefore, the use of

Doherty PAs is not preferred, since the linearity requirements cannot be met without predistortion in this case.

Table 2-4 shows the DE and NMSE of the Doherty PA when using the partitioned distortion compensation technique and compares it to the DE and NMSE of the class AB PA when PBO is used. The DE of the Doherty PA using the proposed distributed distortion compensation technique is considerably higher compared to the conventional class AB PA with 3dB back-off, while the performance in terms of NMSE are comparable for both techniques.

Table 2-4- Measurement results for partitioned compensation and the conventional transceiver

Signal Number	Power Amplifier	DE(%)	NMSE(dB)
4	Case 2: Doherty-Distributed Distortion Compensation	18.7	-33
4	Case 5:Class AB@3dB B.O.	3.1	-34
5	Case 2: Doherty- Distributed Distortion Compensation	27.3	-26
5	Case 5:Class AB@3dB B.O.	6.2	-30

2.4.2.2 Measurements: Ideal Channel case, low power PA

To further clarify the effectiveness of the proposed technique, another set of measurements was performed using a low power PA which is proper for the targeted application.

Two standards were considered in this course of measurements: WiMAX and LTE. Table 2-5 summarizes the specifications of the above mentioned standards.

Table 2-5- Standard specifications

Standard	Bandwidth (MHz)	Occupied Bandwidth(MHz)	$ACPR_1$ dBc@5MHz	$ACPR_2$ dBc@10MHz	EVM (%)	Sampling Frequency (MHz)
WiMAX(Uplink)	5	4.75	30	44	3.2	92.16
LTE(Uplink)	5	4.5	33	36	3.2	61.44

To compare the performance of the proposed method, three other cases were considered. No-compensation, DPD and back-off are the three other cases that were studied. Tables 2-6 and 2-7 list the performances of these systems for WiMAX and LTE standards.

Table 2-6- Measurement results for the WiMAX signal

Technique for WiMAX	Signal PAPR (dB)	PAPR before PA(dB)	$ACPR_1$ (dBc)L,H	$ACPR_2$ (dBc)L,H	EVM (%)	Efficiency (%)
No Compensation	11.4	11.4	-37.7 -38.2	-57.1 -57.6	3.4	10.8
DPD	11.4	15.4	-46.5 -46.9	-56.9 -57.3	0.8	4.4
BO@3dB	11.4	11.4	-45.1 -45.6	-60 -60.2	1.3	5.6
PC	11.4	11.4	-38.7 -39.1	-57.5 -58.2	1.2	10.8

As can be seen from Table 2-6, it is clear that the Phase-only DPD does not change the PAPR of the signal, hence, does not affect the efficiency. All the cases, the spectrum as well as the EVM requirements are satisfied and the ACPRs are below the mask except for the no-compensation case. However, the efficiency shows a noticeable improvement in the case of the

partitioned compensation technique. The PDF of the input data was assumed to be Rician for this case.

Table 2-7 lists the performances for the LTE signal. To pass the mask in this case 1dB back-off was considered for the proposed technique. The input data distribution was assumed to be Weibull.

Table 2-7- Measurement results for the LTE signal

Technique for LTE	Signal PAPR (dB)	PAPR before PA(dB)	$ACPR_1$ (dBc)L,H	$ACPR_2$ (dBc)L,H	EVM (%)	Efficiency (%)
No Compensation	7.8	7.8	-30 -30.5	-44 -44.5	5.9	23.4
DPD	7.8	11.4	-42.5 -43.3	-51.5 -51.7	0.6	11
BO@2.5dB	7.8	7.8	-37.1 -37.8	-50.7 -51.1	1.7	14
PC@1dB BO	7.8	7.8	-33.8 -34.4	-48.9 -49.5	1.7	19.2

2.4.2.3 Standard channel model simulations

In the previous section it was assumed that the channel is ideal and its effect was not account for in simulations and measurements. This paragraph provides the simulation results for the previously mentioned structures in the presence of a non-ideal channel. For the proposed technique, DPD and back-off methods a LUT based model for the PA with around 4.5dB gain compression was applied.

The applied signals in these simulations were generated using Advanced Design System (ADS) 2009 Update 1. The length of the signal was considered to be 50ms (five farmes) with 3 MHz bandwidth, 64QAM modulation coding rate of 4/5 and an oversampling ratio of 8.

The applied channel is an ITU channel model which has been provided in ADS along with the additive white Gaussian noise (AWGN). The pedestrian, channel A with Doppler frequency of 0.1 Hz was chosen to model the fading effects of the wireless channel coupled with AWGN channel. The channel applied herein, has a delay spread of 410ns and Doppler spread

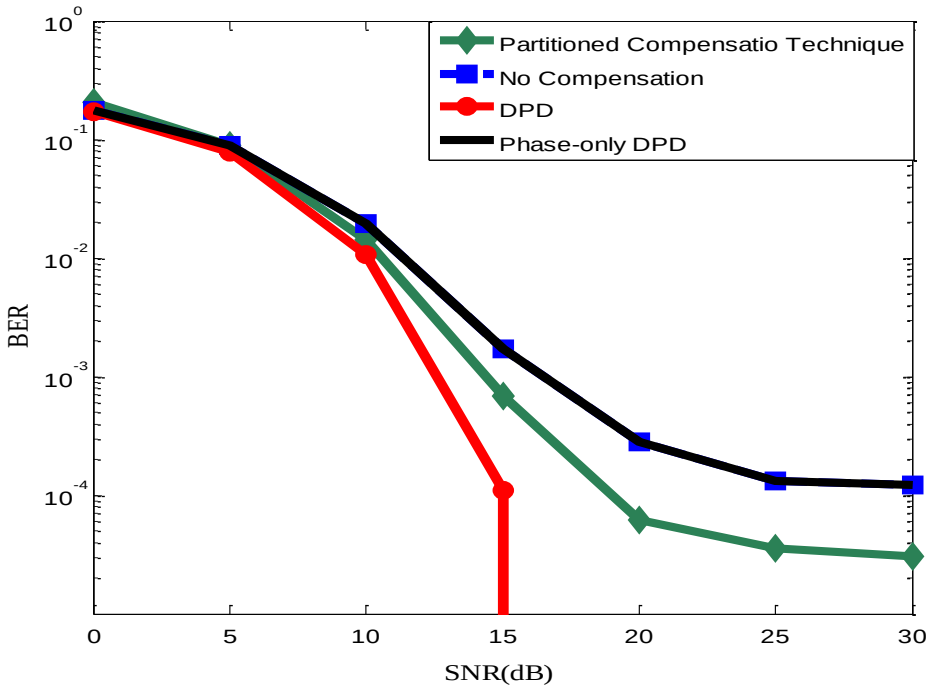

Figure 2-6- BER versus SNR for the proposed method for LUT-based PA with no compensation (black solid line), the conventional DPD (red circle), the proposed method (green diamond) and phase-only DPD (blue square)

of 0.1Hz, accordingly, the channel is a multipath, Doppler flat, frequency selective and stationary one. Among channel equalization techniques for the OFDM data, the minimum mean square estimation (MMSE) method which was implemented in ADS was applied in this work.

Finally, the CDF-based post compensation technique was implemented in MATLAB and was linked to the ADS simulation.

Figure 2-6 shows the receiver BER results versus SNR. As shown in Figure 2-5 the performance of the proposed method for a LUT-based PA model could reasonably compensate for the distortions at high SNRs. To obtain a BER of 10^{-4}, compared to the DPD and BO, 5 dB more SNR is required. The increase in the required SNR for an acceptable BER will reduce the maximum range of the transceiver. This in turn means that this technique trades the maximum range to get an improvement in efficiency. By adopting a better equalization technique such decision feedback equalization, it is anticipated that a significant improvement of the BER by 2-3 orders of magnitude will be achieved for the proposed partitioned compensated technique [42]. In such case, it is anticipated that the SNR corresponding to 10^{-4} will be reduced by 2-3 dB.

It can be concluded that the proposed distributed distortion compensation approach provides the same NMSE and reasonable BER with improved efficiency among all the compared techniques.

2.5 Conclusion

In this chapter, a new architecture for PA distortion compensation has been proposed for uplink applications. In the proposed distributed distortion compensation technique, the conventional class AB PAs was replaced with high efficiency Doherty PAs. The proposed configuration compensates for the phase distortions using non-adaptive LUT-based phase DPD at the transmitter of the mobile terminal. The amplitude nonlinearity was compensated at the receiver of the base station after channel equalization and it does not require any training sequence. It was shown that the distributed distortion compensation provides almost the same

linearity and reasonable BER performance as conventional techniques, but with considerably improved DE.

Chapter Three: Efficient combination of crest factor reduction, DPD and post distortion compensation techniques to linearize transmitters

3.1 Introduction

As previously discussed, modern wireless telecommunication standards such as WCDMA, WiMAX and LTE have been suggested to improve the data rate while conserving spectral efficiency. These achievements are obtained at the cost of a considerably high PAPR of the signal. PA efficiency decreases as the power back-off increases. In most cases the amount of required back-off, when driving wireless transmitters with modulated signals having a given PAPR, equals to the PAPR of the signal. Therefore, PAPR reduction techniques offer an effective way to improve the PA efficiency and hence minimize energy dissipation of transmitters.

Many PAPR reduction techniques have been proposed in the literature [43], [44]. These techniques include clipping, coding schemes, phase optimization, nonlinear companding transforms, tone reservation, tone injection, constellation shaping, partial transmission sequence and selective mapping [43], [44].

Among the above mentioned techniques, Clipping-and-Filtering has found many applications due to its simplicity [43]. Clipping is applied at the transmitter side and produces in-band and out-of-band distortions. However, these distortions can hardly be compensated for. Generally, clipping along with the out-of-band distortion filtering degrades the BER performance of the system. To mitigate this degradation in signal quality, some iterative methods have been proposed in the literature [15]; however, the computational burden of these methods seems high. On the other hand, soft clipping has the possibility to be compensated for at the receiver. This

compensation can be performed using Cumulative Distribution Function (CDF) based blind compensation technique [45].

The problem with clipping-and-filtering lies in the hard clipping part of the method. Despite soft clipping, which is invertible, it is difficult to compensate for the distortions hard clipping produces. As a result, two soft clipping methods have been proposed in [46],[47]. Both of the methods were built based on the Saleh model of PA which has only two parameters [46],[47]. The limited number of parameters of the function restricts the performance of the clipping.

First, in this chapter, the hard clipping problem is studied and a closed form expression for the clipping threshold is proposed. Then, a new soft clipping technique is suggested which is based on a polynomial model. The identification of the clipping function is cast as a constrained optimization problem. To minimize the PAPR while meeting the standard spectral mask requirements a closed form expression for the output spectrum is first derived based on some approximations on the signal spectrum. Following that, the ACPRs are obtained through the output spectrum. The PAPR is then minimized such that the ACPRs pass the standard.

3.2 Clipping-and-Filtering

The traditional crest factor reduction technique relies on two steps: hard clipping and filtering. Hard clipping reduces the PAPR and the filtering is required to regulate the out of band distortions such that it passes the standard mask. The clipping function generally produces in band and out of band distortions. Filtering can reduce the out-of-band distortion to maintain spectral efficiency, but in band distortion remains almost unaffected, even with filtering. This uncompensated distortion reduces signal quality and therefore increases the BER at the receiver. The Clipping-and-Filtering (CLF) technique improves the efficiency by reducing the PAPR of

the signal. However, this improvement is obtained at the cost of degradation in signal quality which can be regarded as information loss. Some remedial techniques have been suggested in the literature [15] to cope with the information loss in CLF, but the majority of them are iterative methods with high computational complexities.

The other problem in applying the CLF technique is the uncertainty in the PAPR of the clipped signal. Since hard clipping is followed by filtering, the PAPR obtained from clipping will change by filtering. Thus, to obtain a certain PAPR, one should iterate the CLF procedure.

Although the CLF technique is well-known, there is no systematic method to obtain the clipping threshold. Therefore, in this section, a method to obtain the desired PAPR by clipping is proposed. Filtering will change this PAPR, and the desired PAPR after filtering can be obtained by brute forcing the clipping level and the filter length.

The clipping function can be described as:

$$y = \begin{cases} x & |x| \le c \\ ce^{j\theta_x} & |x| > c \end{cases} \tag{3-1}$$

where c stands for the clipping threshold, x and y denote the complex input and output signals respectively, and θ_x represents the phase of the input signal.

To set the threshold, one can use the Probability Density Function (PDF) of the input signal to obtain a certain PAPR. As is widely known, modern communication signals can be regarded as circular complex processes and in most cases close to complex Gaussian random variables [49].

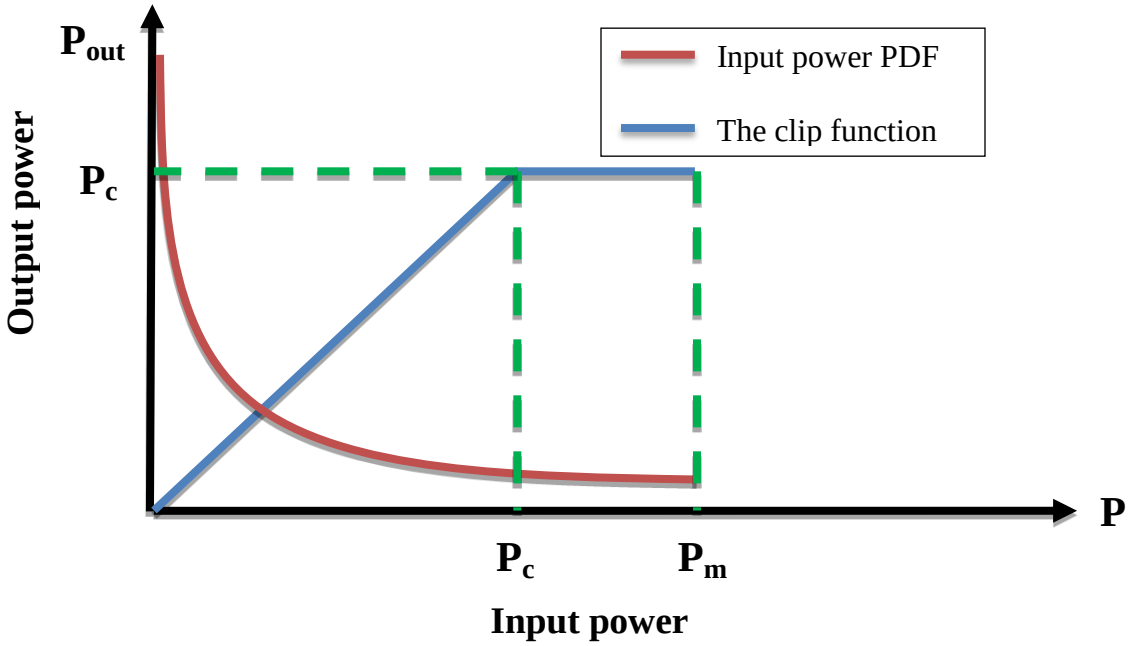

Figure 3-1- Threshold selection for the CLF method

Based on the Gaussian assumption for the distribution, the envelope and the power of the signal follow Rayleigh and exponential distributions respectively. To obtain the PAPR in terms of the clipping parameter c, the distribution of the power p should first be estimated. The exponential distribution with $\frac{1}{\lambda}$ as its mean can be described as:

$$f_{exp}(p) = \lambda exp(-\lambda p) \tag{3-2}$$

Figure 3-1shows the details of the clipping method. If the maximum input power is considered to be P_m, and the clipping power as P_c for a 50 ohm system, the average power can be obtained as follows:

$$P_c = \frac{c^2}{100} \tag{3-3}$$

$$P_{mean} = \int_0^{Pc} p\, f_{exp}(p)dp + \int_{Pc}^{Pm} p_c\, f_{exp}(p)dp \tag{3-4}$$

$$P_{mean} = \frac{1}{\lambda}(1 - exp(-\lambda P_c)) - P_c\, exp(-\lambda P_m) \tag{3-5}$$

Hence, the PAPR becomes

$$PAPR = \frac{P_c}{\frac{1}{\lambda}(1 - exp(-\lambda P_c)) - P_c\, exp(-\lambda P_m)} \tag{3-6}$$

To obtain a certain PAPR, PAPR_dB, in dB one can solve the following equation for P_c:

$$10^{\frac{PAPR\_dB}{10}}\left(\frac{1}{\lambda}(1 - exp(-\lambda P_c)) - P_c\, exp(-\lambda P_m)\right) - P_c = 0 \tag{3-7}$$

Thus, by solving this equation, one can find the clipping level P_c for a certain PAPR in the CLF technique.

3.3 Optimized soft clipping technique

In [46] and [47], two Soft Clipping Crest Factor Reduction (SCCFR) techniques have been proposed, however, little information was provided about the spectral performances of the proposed techniques. Also, there has been little research done on the PAPR of the resulting signal.

Therefore, the main objective of this chapter is to propose a CFR technique that minimizes the PAPR of the signal such that the output spectrum of the PA meets the ACPR requirements. To this end, it is assumed that the combination of the DPD and the PA behaves as a linear system. Hence, the nonlinear distortion of the transmitter output is only due to the CFR

function. To reduce the signal PAPR subject to the ACPR constraints, it is necessary to apply a nonlinear function to the input signal. Consequently, the next step is to evaluate the spectrum of the resulting signal and there are two ways to perform this step. In the first technique, one can apply the nonlinear function to the whole length of the signal and then estimate the PAPR, spectrum and ACPR. Through an optimization process, the optimized CFR function can then be obtained. The possible main critique to this technique is its computational complexity, as the nonlinearity should be applied to the whole data length.

The other possible technique is to analyze the problem mathematically, and find the desired variables, i.e. PAPR and ACPR, analytically as a function of CFR parameters. To approach this target, in this chapter the CFR function is first assumed to follow a memory-less MP model with real coefficients and only odd order terms. The real nature of the coefficients causes the CFR function to be an amplitude-only gain compression with no phase distortion. The next key modeling assumption of this chapter is that the spectrum of modern communication signals has been assumed to have a rectangular pulse shape. Then, based on the theorem relating the output of such a nonlinear function to the input spectrum, which is assumed herein to be a pulse in the frequency domain, the analytic form of the output spectrum and hence, the ACPR as a function of frequency and CFR parameters are obtained. Later, the PAPR is calculated as a function of CFR parameters. Having the PAPR and the ACPR of the function obtained, one can run the optimization much faster than the previous technique, as it is not necessary to apply the function to the whole data.

Figure 3-2 depicts the block diagrams of the transceivers that are compared. The combination of the DPD and the PA is considered herein as a linear system, and it is assumed

that the PA nonlinearity is fully compensated by the DPD. Since, the CFR technique reduces the PAPR of the signal, hence the DPD avalanche problem, as reported in [48], is not an issue.

Based on the above discussion, in this section a polynomial based soft clipping technique is proposed. The closed form expression of the spectrum of the soft clipped signal is then derived. Additionally, the PAPR of the signal is achieved in terms of the nonlinear soft clipping function parameters.

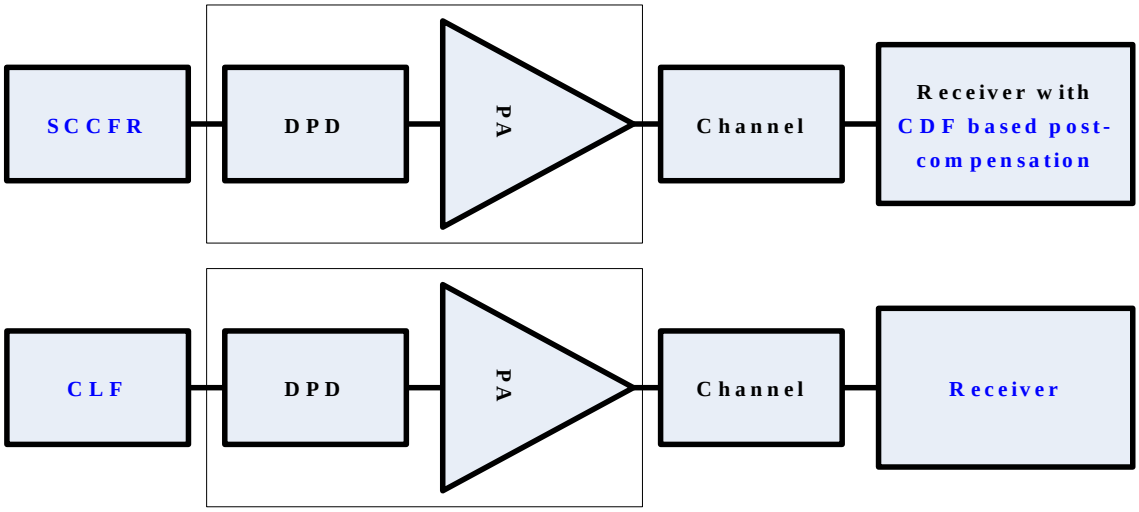

Figure 3-2- The block diagrams of the transceivers

Consequently, to minimize the PAPR, an optimization is performed. The major advantage of the soft clipping technique is that it can be easily compensated for at the receiver by blind CDF estimation. Since, the soft clipping function includes only amplitude compression, then CDF based techniques, which is effective for amplitude-only distortion, can be employed to recover the signal quality at the receiver side.

The proposed crest factor reduction technique relies on a polynomial based compressing function as below:

$$y(n) = x(n)\sum_{k=0}^{K} a_{2k+1}|x(n)|^{2k} \tag{3-8}$$

The soft clipping function has to be determined such that the resulting signal has the minimal PAPR subject to the ACPR constraints. As a result, the problem of minimizing the PAPR with a controlled ACPR can be cast as a nonlinear optimization, and can be formulated as follows:

$$\boldsymbol{a} = arg\min_{\boldsymbol{a}} PAPR\big(F(x,\boldsymbol{a})\big)\ subject\ to\ \{ACPR_{1,2}\} \tag{3-9}$$

in which $\boldsymbol{a} = [a_1, \dots, a_{2K+1}]$ denotes the coefficients vector and

$$F(x,\boldsymbol{a}) = x\sum_{k=0}^{K} a_{2k+1}|x|^{2k} \tag{3-10}$$

The first step in this optimization problem is to find the PAPR and $ACPR_1$ and $ACPR_2$ (adjacent and alternate channel power ratios for the adjacent and alternate channels) in terms of the function parameters.

In [49], the general spectrum and the covariance function of the output of a polynomial nonlinearity driven by a circular complex signal have been derived. It has been shown that the output covariance of a polynomial nonlinearity of type (3-8), $c_{2y}(\tau)$, can be related to the coefficients of the nonlinear function and the input covariance, $c_{2x}(\tau)$, as [37][49]:

$$c_{2y}(\tau) = \sum_{m=0}^{K} \alpha_{2m+1} c_{2x}(\tau)|c_{2x}(\tau)|^{2m} \tag{3-11}$$

where

$$\alpha_{2m+1} = \frac{1}{m+1}\left|\sum_{k=m}^{K} a_{2k+1}\binom{k}{m}(k+1)!\,c_{2x}(0)^{k-m}\right|^2 \tag{3-12}$$

and the output spectrum of the nonlinearity can be obtained as below [37]:

$$S_{2y}(f) = \sum_{m=0}^{K} \alpha_{2m+1} \underbrace{S_{2x}(f) * \cdots * S_{2x}(f)}_{m+1} * \underbrace{S_{2x}(-f) * \cdots * S_{2x}(-f)}_{m} \tag{3-13}$$

A useful approximation to obtain a closed form expression of the output spectrum is assuming a rectangular pulse shape for the input signal spectrum. As one of the main considerations in modern communication signal design is to maximize the spectral efficiency, the spectrum is localized and this approximation seems valid for most of the standards such as WCDMA and WiMAX. To consider the effect of noise, the whole spectrum is assumed to be the summation of two pulses as is shown in Figure 3-3.

To summarize the results of the analysis, just the final expressions for the spectral density is provided here, and the full mathematical analysis is explained in appendix A of this thesis.

As can be concluded from appendix A, the overall spectrum will take the following form:

$$S_{2y}(f) = \sum_{m=0}^{K} \alpha_{2m+1} \sum_{i=0}^{2m+1} \binom{2m+1}{k} S_{2x(i)}(f) * S_{2w(2m+1-i)}(f) \tag{3-14}$$

Based on the output spectrum the $ACPR_1$ and $ACPR_2$ can be calculated through the following equations:

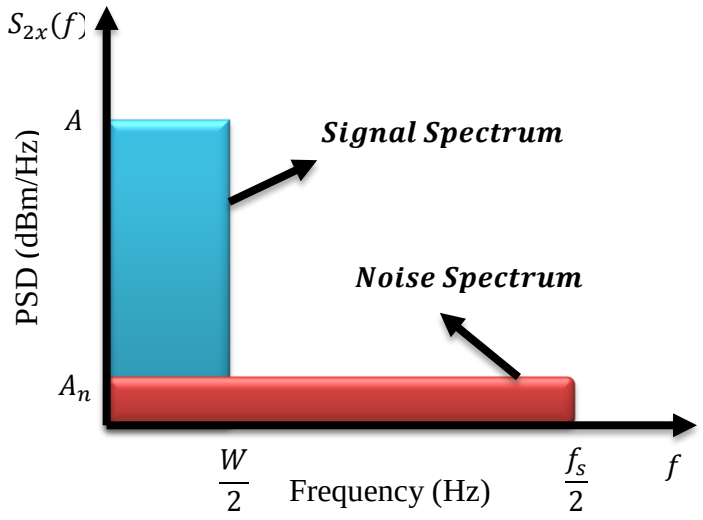

Figure 3-3- Pulse approximation of the input signal spectrum

$$ACPR_1 = 10log10\left(\frac{E\left(S_{2y}\left(\left[-\frac{W}{2}:\frac{W}{2}\right]\right)\right)}{E\left(S_{2y}\left(W_{offset1}+\left[-\frac{W}{2}:\frac{W}{2}\right]\right)\right)}\right) \quad (3\text{-}15)$$

$$ACPR_2 = 10log10\left(\frac{E\left(S_{2y}\left(\left[-\frac{W}{2}:\frac{W}{2}\right]\right)\right)}{E\left(S_{2y}\left(W_{offset2}+\left[-\frac{W}{2}:\frac{W}{2}\right]\right)\right)}\right) \quad (3\text{-}16)$$

where $E(.)$ denotes the expectation operator, $W$is the bandwidth of the signal and $W_{offset1}$and $W_{offset2}$ are the offset frequencies from the carrier frequency in which the ACPRs should be calculated, and depends on the standard.

Once the output spectrum has been obtained, the next step is to find the PAPR of the output signal. As can be concluded from (3-11) the mean power of the output signal can be obtained as [49]:

$$P_{meany} = E(|y|^2) = c_{2y}(0) = \sum_{m=0}^{K} \alpha_{2m+1} c_{2x}(0) |c_{2x}(0)|^{2m} \tag{3-17}$$

The next problem is to find the maximum power of the output signal. In the optimization problem a constraint can be applied on the derivatives of the nonlinearity such that the nonlinearity is always a monotonic function. In this thesis, however the maximum of the output signal has been checked at the extreme values of the output signal (roots of the first derivative of the power).

To obtain the maximum power of the output, first the derivative of the output power versus input power should be calculated. As we know from (3-8)

$$y = x \sum_{k=0}^{K} a_{2k+1} |x|^{2k} \tag{3-18}$$

subsequently, for the power of the output signal, the expression should be multiplied by its complex conjugate value:

$$yy^* = \left(x \sum_{k=0}^{K} a_{2k+1} |x|^{2k} \right) \left(x \sum_{k=0}^{K} a_{2k+1} |x|^{2k} \right)^* \tag{3-19}$$

$$|y|^2 = G(|x|, \boldsymbol{a}) = \sum_{k=0}^{K} \sum_{k=0}^{K} a_{2k+1} a_{2j+1}^* |x|^{2k+2j+2} \tag{3-20}$$

$$\begin{aligned} |y|^2 = & \sum_{n=1}^{K+1} |x|^{2n} \left(\sum_{m=1-n}^{n-1} a_{n+m} a_{n-m}^* \right) \\ & + \sum_{n=K+1}^{2K+1} |x|^{2n} \left(\sum_{m=n-(2K+1)}^{2K+1-n} a_{n+m} a_{n-m}^* \right) \end{aligned} \tag{3-21}$$

$$\frac{d|y|^2}{d|x|} = \sum_{n=1}^{K+1} |x|^{2n-1} \left(\sum_{m=1-n}^{n-1} 2n a_{n+m} a_{n-m}^{*} \right) + \sum_{n=K+1}^{2K+1} |x|^{2n-1} \left(\sum_{m=n-(2K+1)}^{2K+1-n} 2n a_{n+m} a_{n-m}^{*} \right) \tag{3-22}$$

Therefore, the next step is to find the roots of the derivative polynomial to find the maximum points. The real roots of this polynomial are called$R_{1:h}$, where h is the number of real positive roots. To make the algorithm more robust, the value of the nonlinearity at maximum input power has been calculated, and compared to the value of the function at extreme points, i.e. the roots of the derivative. The maximum power has been calculated as:

$$P_{maxy} = max(G(|x|, \boldsymbol{a})) = max\big(G(R_{1:h}, \boldsymbol{a}), G(|x|_{max}, \boldsymbol{a})\big) \tag{3-23}$$

Accordingly, the PAPR of the nonlinearity is:

$$PAPR\big(F(x, \boldsymbol{a})\big) = 10log10\left(\frac{P_{maxy}}{P_{meany}}\right) \tag{3-24}$$

Once all the optimization variables have been calculated in terms of the nonlinearity coefficients, the optimization problem can be performed. As shown in (3-9) the PAPR minimization subject to ACPR constraints can be formulated as below:

$$\boldsymbol{a} = arg\min_{\boldsymbol{a}} PAPR\big(F(x, \boldsymbol{a})\big)\ subject\ to\ \{ACPR_{1,2}\} \tag{3-25}$$

The reason behind the spectrum estimation algorithm is that this algorithm reduces the complexity of the ACPR evaluation, and hence decreases the running time of the optimization algorithm. One can apply the nonlinearity on the data and calculate the spectrum using spectrum estimation functions. Since spectrum estimation, convolution, polynomial root finding and polynomial evaluation were used in the algorithm, the complexity follows $O(Nlog(N))$where N

denotes the length of the data. In this technique, to evaluate the ACPR, the PSD expression has to be evaluated at around 1024 points of the baseband spectrum.

As previously discussed, one can apply the CFR function to the whole data length, and evaluate the ACPR by estimating the PSD of the resulting data. The complexity of the polynomial evaluation and spectrum estimation is also$O(Nlog(N))$. Subsequently, the only difference is that for this technique, the whole length of the data should be used which includes 200000 samples. This fact shows that the complexity of the proposed method is much lower than the traditional technique for spectrum estimation.

3.4 Post-compensation at the receiver

Based on the fact that the polynomial SCCFR function coefficients are real, the nonlinear function can be considered as an amplitude-only nonlinearity. To compensate for this distortion, CDF based nonlinearity compensation technique can be employed [29]. The theorem that best describes this technique was explained in section 2.1.1. By estimating the distribution of the received signal, one can compensate for the amplitude distortion, which is the case in the described SCCFR method.

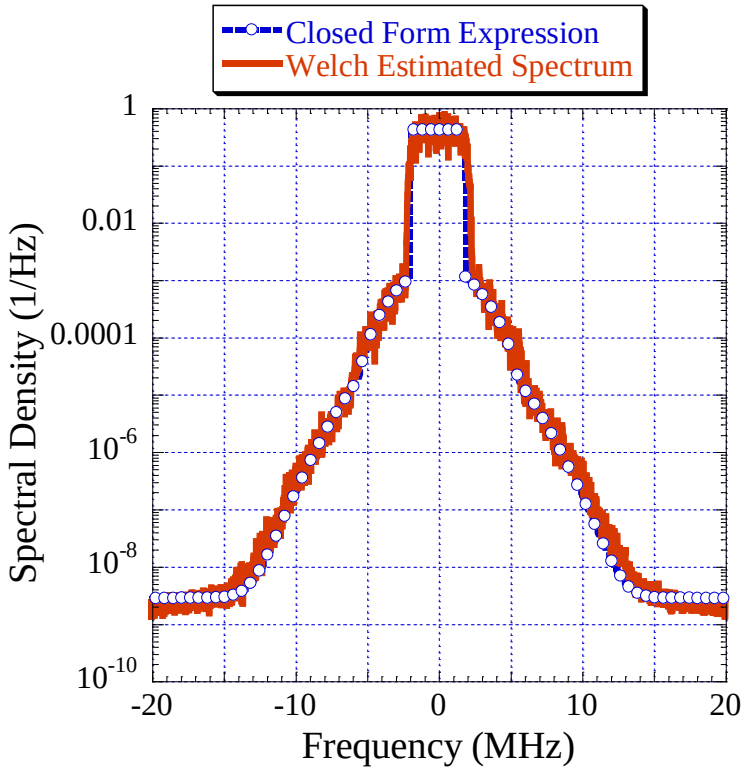

Figure 3-4- Estimated spectrums: Empirical estimate (red square) and approximation (blue circle)

3.5 Performance evaluation of SCCFR and CLF techniques

In this section, the results of the previously described algorithms i.e. SCCFR and CLF are compared. Simulations and measurements are performed to evaluate the performances of the two PAPR reduction techniques.

Before providing the results of the two CFR techniques, the estimated spectrum should be confirmed. To validate the performance of the closed-form expression for the output spectrum as described in (3-14), simulation result of the SCCFR method is shown in Figure 3-4. The output spectrum at the output of a polynomial nonlinearity following the expression (3-8), which was the optimized SCCFR polynomial, is compared to the exact spectrum. As shown in Figure 3-4, the estimated spectrum seems to fit the exact spectrum.

3.5.1 Simulation and measurement Setup

The excitation signals for system identification and evaluation are modulated Wideband Code Division Multiple Access (WCDMA) with signal bandwidth of 3.84 MHz. Waveform is 4-ms long (245883 samples) and is sampled at 61.44 MHz. Table 3-1 summarizes the signal specifications for the WCDMA signal. For modeling purposes 8000 points of the signals were used.

Table 3-1- Parameters of the modulated signal

WCDMA Parameters
Band width = 3.84 MHz
Chips per Slot = 2560
Modulation=16QAM
Samples per Chip = 16
PAPR=9.9 dB
3GPP Test Model 5

All the simulations were performed in MATLAB. The simulated ACPR results for both SCCFR and CLF methods were obtained assuming an ideal linear PA. This assumption can be realized in practice using a full DPD preceding the PA.

For the measurements, one class AB PA was used in the course of experimental evaluation. The PA block is a cascade of a class A driver and a class AB PA biased at Vgs=9V,Vds=28V. Both blocks are designed using the TF10107 LDMOS transistor from Ericsson Inc. The output signals were attenuated and captured by vector signal analyzer (VSA89650S from Agilent Inc).

The simulation and measurement set-ups follow the block diagram in Figure 2-5. The Digital Predistortion (DPD) is applied after PAPR reduction such that the overall combination of the DPD and the PA becomes a linear system. The channel herein is considered to be an ideal channel (i.e. line of sight).

3.5.2 Simulation and measurement results

As reported in [51] the standard specifications of WCDMA mobile user equipment can be summarized as Table 3-2.

Table 3-2- Standard requirements for WCDMA 16QAM

Standard	ACPR1(dBc)	ACPR2(dBc)	EVM (%)	$W_{offset1}$	$W_{offset2}$
WCDMA(16QAM)	33	43	12.5	5MHz	5MHz

Based on the standard requirements, the optimization was performed for the SCCFR method. The minimum PAPR acquired by the SCCFR, while passing all the standard constraints, is found to be 5.5 dB. This optimization could decrease the PAPR around 4.4 dB from an initial

PAPR of 9.9 dB. The resulting simulated ACPR1 and ACPR2 are reported in Table 3-3. To compare both methods, the Clipping-and-Filtering threshold was set to 5.5 dB. The resulting ACPRs passed the mask, hence, there was no need for filtering.

Table 3-3- Simulation results for SCCFR and CLF methods

Method	PAPR(dB)	ACPR1(dBc)	ACPR2(dBc)
SCCFR	5.5	33.3	58.5
CLF	5.5	33.8	49.1

To validate the concepts by measurements, memory polynomial DPDs with memory depth of 3 and nonlinearity degree of 12 were applied to the SCCFR and CLF signals. The signals were then passed through the PA.

Table 3-4 summarizes the measurement results. As shown in this table, the ACPRs pass the standard constraints in both methods, and the PAPRs are close to the simulated values.

Table 3-4- Measurement results for SCCFR and CLF methods

Method	PAPR DPD	PAPR PA	ACPR1L	ACPR1H	ACPR2L	ACPR2H
SCCFR	8.8	5.6	34	34.3	55.8	56.5
CLF	8.4	5.7	33.8	34.3	47.8	47.7

Table 3-5 reports the results of the efficiency and the EVM of both techniques. For the SCCFR method, the EVM reported herein is after post compensation, but in the case of CLF no post compensation can be employed theoretically.

It is evident from the table that with the same efficiency, SCCFR provides substantial and non-negligible improvement in the EVM.

Table 3-5- Efficiency and EVM for SCCFR and CLF methods

Method	Efficiency (%)	EVM (%)
SCFR	36.1	1.25
CLF	36	4.52

3.6 Performance evaluation of the cascade structure

In the second method of benchmarking the algorithms, the SCCFR method was combined with the CLF method to reduce the PAPR while satisfying the spectral requirements. It is shown in this chapter that by employing the CDF based compensation at the receiver side; the EVM requirement of the standard is passed. Meanwhile, the CLF method with a PAPR higher than what is obtained through combined structure does not pass the standard EVM. The simulation and measurement set-up remains the same as the section V, however, the

SCCFR is replaced with the combined structure namely SCCFR-CLF block.

Figure 3-5 illustrates the block diagram of the SCCFR-CLF technique. As previously mentioned, the DPD+PA cascade is considered to be fully linear.

To compare SCCFR-CLF with CLF-only technique, first the CLF was applied on the signal with a clipping threshold of 2.8 dB and a filter length of 13. The resulting PAPR, ACPR1 and ACPR2 were 3.65dB, 33.3 dBc and 76.2dBc respectively.

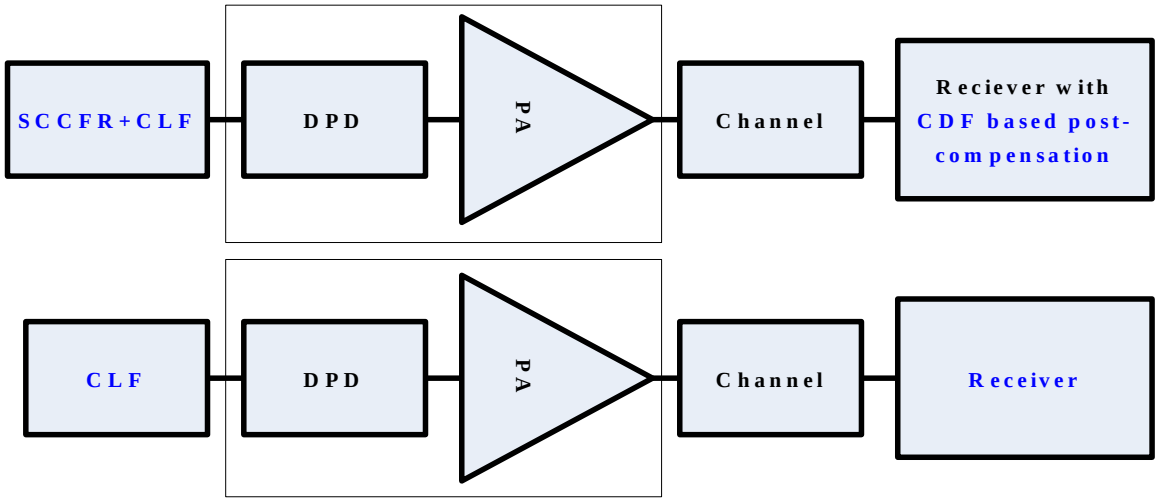

Figure 3-5- The transceiver block diagram for the combined CFR technique and CLF

To combine the CLF technique with SCCFR, the out of band distortions should be filtered such that the spectrum resembles a pulse. A FIR filter of length 800 and with 5 MHz bandwidth is applied on the CLF output.

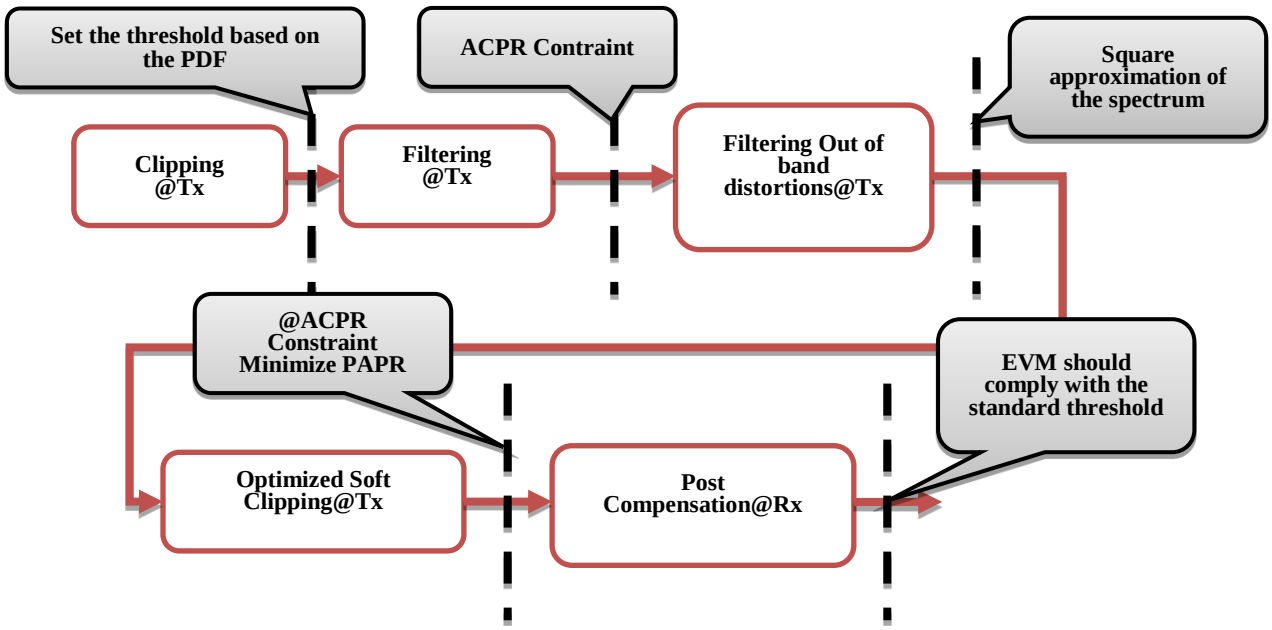

Figure 3-6- The general block diagram of SCCFR-CLF transceiver

The PAPR of the signal increased to 5dB. Then the optimization was carried out to reduce the PAPR using SCCFR technique. This technique could reduce the PAPR to 3.4 dB with 33.2 dBc ACPR1 and 58.9 dBc ACPR2. In the CLF method the clipping threshold was set to 3.6 dB with a filter length of 11, resulting in 4.2 dB PAPR, 33.8 dBc ACPR1 and 75.6 dBc ACPR2. Table 3-6 reports the simulation results. Figure 3-6 illustrates the SCCFR-CLF algorithm and the resulting transceiver architecture.

Table 3-6- -Simulation results for SCCFR-CLF and CLF methods

Method	PAPR(dB)	ACPR1(dBc)	ACPR2(dBc)
SCCFR-CLF	3.4	33.2	58.9
CLF	4.2	33.8	75.6

The measurement results are listed in Table 3-7. As can be seen the output spectrums meet the ACPR requirements. Table 3-8 reports the EVM and the efficiency for both structures.

Table 3-7- Measurement results for SCCFR-CLF and CLF methods

Method	PAPR DPD	PAPR PA	ACPR1L	ACPR1H	ACPR2L	ACPR2H
SCCFR-CLF	6.1	3.4	33.2	33.4	55.6	56.5
CLF	6.7	4.3	34	34.3	61	60

The reason why the CLF threshold was set to 3.6 dB is that, as can be seen from Table 3-8, this threshold cannot pass the EVM requirement. While meeting the standard EVM, the combined structure results in a remarkable improvement in efficiency. Since the part of the nonlinearity of the combined structure can be compensated, the CDF based post-distortion could reduce the EVM down to the standard requirement, while for the CLF technique no post compensation technique has been proposed in the literature.

Table 3-8- Efficiency and EVM for SCCFR-CLF and CLF methods

Method	Efficiency (%)	EVM (%)
SCCFR-CLF	45.4	11.56
CLF	42.3	12.7

3.7 Conclusion

In this chapter, a new soft clipping (spectrum constrained) technique using polynomial function was proposed. The resulting spectrum and PAPR were then calculated in closed form based on the polynomial parameters. The polynomial coefficient extraction was performed using PAPR minimization subject to spectral constraints. It has been shown that the proposed technique could result in improvements in efficiency and EVM. Using CDF based post-distortion compensation; the SCCFR method could improve the EVM while having the same efficiency as CLF. To further decrease the PAPR, SCCFR were combined with CLF which resulted in lower PAPR and higher efficiency. The most important issue is that the combined structure using CDF based post compensation can pass the EVM requirement. However the CLF method with a higher PAPR could not meet the EVM threshold.

Chapter Four: Optimization of model orders for DPD design and PA modeling purposes: SISO and MIMO

4.1 Introduction

Nonlinear distortion in the transmitter chain is mainly due to PA stage. Compensation of nonlinear distortion is of major concern in linear transmitter design. DPD has been shown to be a powerful tool among digital linearization techniques for which the memory polynomials (MP) successfully model PA behaviour [16]. A DPD is commonly designed based on a parametric model for the PA.

A range of parametric models have been proposed in the literature for PA. These models can be categorized into three classes: Volterra-based models, box-based models and neural networks. Recent studies show that MP is one of the most computationally efficient methods, which can model the dynamics of the PA accurately with least complexity [16],[17]. Another advantage of the MP model is that it is linear with respect to the model parameters, meaning that the MP model identification problem can be solved as a linear regression problem. When identifying a model the number of parameters or the model dimensions should be known a priori. Hence model order selection is the next step in DPD design and it is mostly built based on minimizing the NMSE [9],[10],[53]. The NMSE measure is monotonic versus the model dimensions. In [10], a threshold is defined for the NMSE and the order of the model is swept till the NMSE meets the threshold. This procedure is time-consuming since the model has to be identified for all the orders. In addition, it can only be done manually and there is no systematic approach for trading off small improvements in NMSE versus significant increases in computational complexity. To automate the model order selection process and to decrease the

number of iterations, an appropriate model order selection criterion, which considers the NMSE as well as the computational complexity of the system, has to be adopted.

The application of information theoretical criteria in PA modelling has been limited to memory depth estimation only. An embedding dimension estimation procedure has been developed in [9] for memory depth estimation based on the mutual information concept.

Akaike information criterion (AIC) and Bayesian information criterion (BIC) were proposed in [55] and [56] for model order selection of linear models. Indeed, AIC and BIC has found diverse applications in the literature for model order selection. In fact, most of the linear models, such as autoregressive (AR), moving average (MA) and autoregressive moving average (ARMA) models, have their orders selected based on AIC and BIC in most cases. However model order selection of nonlinear systems can be considered as a linear regression problem if the nonlinear system can be modelled using a parametric model linear in its parameters. As mentioned previously, the first step in system identification problems is to select the model orders or number of parameters. These integer-valued numbers, which determine the complexity of the model, should be set in the beginning of the statistical model identification procedure [57]. The main objective is to find the "best approximating" model for a given dataset. The selected model is chosen among some competing models with different number of parameters. To select the model there should be a criterion that includes the mean square error and the complexity of the model [58]. In [58] AIC has been used for polynomial regression. Order estimation before AIC was considered to be a hypothesis testing procedure. The null hypothesis was the lowest dimension model, which was compared against higher dimension models based on the model significance level. With the introduction of AIC the problem was then formulated as an estimation procedure. This estimation is done based on the maximum likelihood (ML) algorithm.

In reality likelihood function of the parameters is the most sensitive quantity to the deviations of the parameters with respect to the true values [55].

In spite of the fact that AIC is the optimal rule under ML estimation, always there is a probability of over fitting the model due to the fact that the complexity term is trivial when compared to the error term. BIC on the other hand maximizes the Maximum A Posteriori (MAP) probability of estimation and the probability of correct detection tends to unity when the number of samples is high [57].

In [59] the problems of over-fitting and under-fitting were addressed and the performance of AIC was evaluated based on the Eigen-values of signal and noise.

By using the AIC and BIC in determining the model order in PA modeling and predistortion, the order selection of the model is automated in comparison to previous works in this area. This automation was possible because the criterion has a global optimum in contrast to NMSE, which is monotonic. The optimum order can then be found by optimization algorithms such as gradient descent for modeling the SISO systems, and simulated annealing for modeling the MIMO systems, for which the model has higher number of parameters and the information criteria are generally a non-convex function.

4.1 SISO Model Order Selection Techniques

4.1.1 Power Amplifier Modeling

In a typical transmitter a PA represents the last active stage. To improve efficiencies in PAs it is generally accepted to drive them close to the saturation region. Accordingly PA can be viewed as a nonlinear system and be modeled by a nonlinear function as follows:

$$y = f(x) + e \tag{4-1}$$

where y and x represent the system output and input respectively and e denotes the residual uncorrelated error assumed herein having a Gaussian distribution and f is a complex nonlinear function, modeling the PA complex response.

In addition to the dependency of the output signal on the current input signal, there is a dependency from the previous input samples, which is known in the literature as a memory effect [10]. The most commonly used model for PA is the memory polynomial, which can be formulated as shown below:

$$f\big(x(n:n-m)\big) = \sum_{i=1}^{K}\sum_{m=0}^{Q}(a_{im}+jb_{im})x(n-m)|x(n-m)|^{i-1} \tag{4-2}$$

where $x(n:n-m)$is the $[x(n),\dots,x(n-m)]$vector, K, Q and are the nonlinearity order, the memory depth and $(a_{im}+jb_{im})$ are the complex parameters of the memory polynomial model, and $x(n)$ is the complex envelope of the signal. In literature, no automated method for choosing the optimal order (K,Q) has been proposed. In the following, Akaike information criterion and Bayesian information criterion (BIC) will be discussed.

4.1.2 AIC and BIC

AIC and BIC are obtained through the Kullback–Leibler (KL) distance, which measures the difference between two probability distributions. AIC and BIC belong to a family of model order selection rules in which the KL measure is minimized. KL measure is defined as below [57]:

$$I(p_0, PDF) = E_0\{ln(PDF(\boldsymbol{y})\} \tag{4-3}$$

where $\boldsymbol{y} = y(1:N)$ represents the output vector, N denotes the data length, $p_0$is the true PDF of the data, $PDF(\boldsymbol{y})$ is the PDF of the model and $E_0\{.\}$ is the expectation operator that is

calculated over the true PDF$p_0(\boldsymbol{y})$. Since the model likelihood (PDF) is not available, assumptions have to be made, which will contribute in determining the model order selection criterion [57].

The vector of model parameters is denoted by $\boldsymbol{\theta} = [\boldsymbol{a}; \boldsymbol{b}; \sigma^2]$**,** where σ, is the standard deviation of the error, e, and:

$$\boldsymbol{a} = \left[a_{10} \ \dots a_{1Q} a_{20} \dots \ a_{im} \dots \ a_{K0} \ \dots a_{KQ}\right]^T, \boldsymbol{b} = \left[b_{10} \ \dots b_{1Q} b_{20} \dots \ b_{im} \dots \ b_{K0} \ \dots b_{KQ}\right]^T$$

The likelihood function of the output vector $\boldsymbol{y}$ is denoted by $PDF(\boldsymbol{y}, \boldsymbol{\theta})$**,** which depends on the parameter vector $\boldsymbol{\theta}$. The ML estimate of the parameters will be [57]:

$$\widehat{\boldsymbol{\theta}} = arg\{\max_{\boldsymbol{\theta}} \{ln(PDF(\boldsymbol{y}, \boldsymbol{\theta}))\}\} \tag{4-4}$$

Based on the Gaussianity of the residuals and by considering the fact that for complex data the data length will be $2N$, the likelihood of the parameters for the nonlinear model follows the form below [57]:

$$PDF(\boldsymbol{y}, \boldsymbol{\theta}) = \frac{1}{(2\pi)^N \sigma^{2N}} exp\{-\left(\frac{\|\boldsymbol{y} - f(\boldsymbol{x})\|^2}{2\sigma^2}\right)\} \tag{4-5}$$

Therefore, at the ML estimate of the parameters $\widehat{\boldsymbol{\theta}}$, the following relation is satisfied [57]:

$$-2ln\left(PDF\left(\boldsymbol{y}, \widehat{\boldsymbol{\theta}}\right)\right) = constant + 2Nln(\hat{\sigma}^2) \tag{4-6}$$

In the definition of information criteria the Fisher information matrix plays a critical role. The fisher information matrix is defined as:

$$\boldsymbol{J} = -E\left\{\frac{\partial^2 ln(PDF(\boldsymbol{y}, \boldsymbol{\theta}))}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}\right\} \tag{4-7}$$

where $E\{.\}$ indicates the expectation. .Using this matrix, the AIC has been defined as below [57]:

$$AIC = -2ln\left(PDF\left(y,\hat{\theta}\right)\right) + trace\left(\boldsymbol{J}_y\left(\boldsymbol{J}^{-1} + \boldsymbol{J}_y^{-1}\right)\right) \tag{4-8}$$

In which $J$is the Fisher information matrix at the ML estimate of the parameters and equals to J_y. In fact J is the Fisher matrix for a fictitious data with the same length and the same PDF but independent from y, from which the ML estimate of the parameters are obtained. Hence AIC will take the following form [57]:

$$AIC = -2ln\left(PDF\left(\boldsymbol{y},\widehat{\boldsymbol{\theta}}\right)\right) + 2p \tag{4-9}$$

where $p = 2K(Q+1)+1$ is the number of parameters. Equation (8) can be reformulated as:

$$AIC = 2Nln(\hat{\sigma}^2) + 2p \tag{4-10}$$

The effect of finite sample size on AIC can be considered as reported in [57]:

$$AIC = 2Nln(\hat{\sigma}^2) + \frac{2N}{2N-p-1}2p \tag{4-11}$$

In [57] it was shown that the PDF of $\widehat{\boldsymbol{\theta}}$ under regularity conditions is as follows:

$$PDF\left(\widehat{\boldsymbol{\theta}}\right) = \frac{1}{(2\pi)^{p/2}|\boldsymbol{J}^{-1}|^{1/2}} exp\left(-\frac{1}{2}\left(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)^H \boldsymbol{J}\left(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)\right) \tag{4-12}$$

In the same work, it was also shown that the unconditional PDF for the output signal will be:

$$PDF(\boldsymbol{y}) \approx PDF\left(\boldsymbol{y},\widehat{\boldsymbol{\theta}}\right)PDF(\widehat{\boldsymbol{\theta}})(2\pi)^{p/2}|\boldsymbol{J}^{-1}|^{1/2} \tag{4-13}$$

and the Bayesian information criterion will be:

$$BIC = -2I \tag{4-14}$$

where:

$$I = ln\big(PDF(\boldsymbol{y})\big) = ln\left(PDF\left(\boldsymbol{y}, \widehat{\boldsymbol{\theta}}\right)\right) + ln(PDF(\widehat{\boldsymbol{\theta}})) + \frac{p}{2} ln(2\pi) - \frac{1}{2} ln(det(\boldsymbol{J})) \tag{4-15}$$

By considering that $PDF(\widehat{\boldsymbol{\theta}})$ and $\frac{p}{2}\ln(2\pi)$ are not functions of N, and by considering high values of N, these two terms can be neglected. The Bayesian information criterion can then be approximated to:

$$I \approx ln\left(PDF\left(\boldsymbol{y}, \widehat{\boldsymbol{\theta}}\right)\right) - \frac{1}{2} ln(det(\boldsymbol{J})) \tag{4-16}$$

The Fisher information matrix ($\boldsymbol{J}$) was approximated in [56] for large N as below:

$$ln(det(\boldsymbol{J})) = ln\left(det\left(N\frac{1}{N}\boldsymbol{J}\right)\right) \approx plnN + O(1) \tag{4-17}$$

By using (4-6), (4-15) and (4-16), the expression of the BIC is given in [57] can be written as follows:

$$BIC = 2N\, ln\, \hat{\sigma}^2 + p\, ln(N) \tag{4-18}$$

4.1.3 SISO system model order optimization

In the cases of AIC and BIC for SISO systems, optimization method such as Gradient descent can be employed to find the optimum dimensions of the model. If the criterion is called by $V_N(\theta)$for the data length of N, and $\theta = [K, Q]$ then the idea of Gradient descent is to move in the direction of$-\frac{\partial V_N(\theta)}{\partial \theta}$. However, the parameters (K, Q) are integer valued and the steps are discrete. $g(.)$ is defined as a monotonic function and α is a scale parameter. If we define the forward difference vector as below:

$$\boldsymbol{D} = [V_N(K+1) - V_N(K); V_N(Q+1) - V_N(Q)] \tag{4-19}$$

Then $g(.)$is defined as below:

$$g(\boldsymbol{D}) = tanh\left(\frac{|\boldsymbol{D}|}{\|\boldsymbol{D}\|}\right)\left(log_{10}(1+|\boldsymbol{D}|)\right) sign(\boldsymbol{D}) \tag{4-20}$$

wherein $\|.\|$is the norm function. The term $\frac{|\boldsymbol{D}|}{\|\boldsymbol{D}\|}$ generates a normalized positive displacement. This function makes suitable steps for large and small $\boldsymbol{D}$. For large $\boldsymbol{D}$ the logarithm term dominates and for small $\boldsymbol{D}$ the $tanh(.)$ term dominates.

The method used in this work consists of measuring the forward difference instead of the derivative and then rounding the value to the nearest integer to make the discrete step.

$$d\boldsymbol{\theta} = round(-\alpha g(V_N(\boldsymbol{\theta}+\mathbf{1}) - V_N(\boldsymbol{\theta}))) \tag{4-21}$$

Figure 4-1 shows the general algorithm of gradient descent. In the following section, the generalization of the criteria to the MIMO system will be studied.

4.2 Matrix Memory Polynomial Model Order Selection in MIMO Transmitters

4.2.1 MIMOAIC and MIMOBIC

The matrix memory polynomial (MMP) technique has been developed to accurately model MIMO transmitters. The MMP model considers the effect of RF crosstalk between the two branches of the MIMO transmitter in addition to the transmitter nonlinearities [20], [60]. Figure 4-2 shows the architecture of the MMP model, in which the output of each branch is a function of both input signals as follows:

$$y_i = f_{i1}(x_{in}^1) + f_{i2}(x_{in}^2) \;;\; i = 1,2 \tag{4-22}$$

where $f_{ij}(x)$ are nonlinear functions of the input signals, identified using both inputs and outputs of the MIMO transmitter. Using the modified Volterra series [61] for nonlinear functions, the expressions in (4-21) can be written as:

$$f_{ij}\big(x(n)\big) = \sum_{l=1}^{K_{i,j}} \sum_{m=0}^{Q_{i,j}} a_{lm,ij} x(n-m)|x(n-m)|^{l-1} \tag{4-23}$$

M is the maximum number of iteration

Initialize the model with $\boldsymbol{\theta}_1$

Measure the cost function *E*

For *k= 1 to M*

Consider the $d\boldsymbol{\theta}$ as the move vector

Update the parameters $\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + d\boldsymbol{\theta}(\boldsymbol{\theta}_k, \alpha)$

If $\Delta E < 0$ then accept the move and update parameters and increase α

If $\Delta E > 0$ then decrease α

If $\Delta E = 0$ check the neighborhood

End for loop

Figure 4-1- Gradient descent algorithm

where $K_{i,j}, Q_{i,j}$ are the nonlinearity order and the memory depth of the memory polynomial used in (21) and $a_{lm,ij}$ is the complex coefficient.

Considering the fisher information matrices of each branch of the MIMO transmitter as $\boldsymbol{J_1}$and $\boldsymbol{J_2}$ and the fisher information of the overall system as $\boldsymbol{J}$, it was shown in Appendix B that the following approximation is valid:

$$ln(det(\boldsymbol{J_i}) = ln(det(\boldsymbol{J}_{i,11})) + ln\left(det\left(\boldsymbol{J}_{i,22}\right)\right); for\ i = 1,2 \tag{4-24}$$

where $\boldsymbol{J}_{i,11}$ and $\boldsymbol{J}_{i,22}$ are matrices as defined in Appendix B.

Using (4-23), the BIC in (4-17) can be extended for the 2x2 MIMO system, and the new MIMOBIC is as follows:

$$MIMOBIC_i = 2N\ ln(\hat{\sigma}_i^2) + (p_1 + 1)ln(N) + p_2 ln(N) \tag{4-25}$$

In which $p_i = 2K_i(Q_i + 1)$

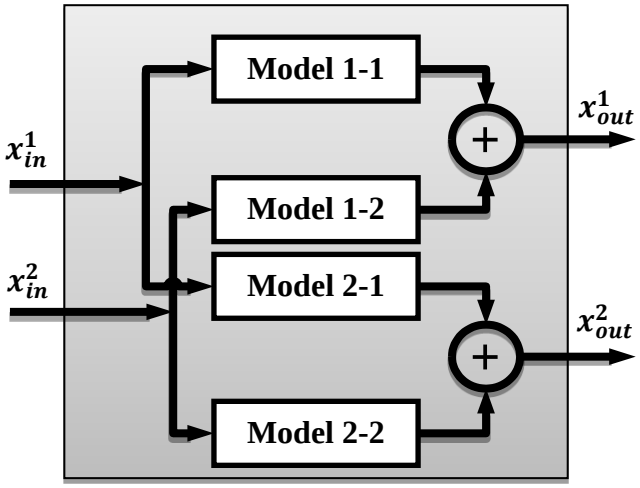

Figure 4-2- Matrix memory polynomial model for MIMO transmitters.

By similar analysis it can be deduced that for the general $M \times P$ MIMO system the MIMOBIC can be generalized:

$$MIMOBIC_i = 2N\, ln(\hat{\sigma}_i^2) + \left(1 + \sum_{j=1}^{M} p_j\right) ln(N) \tag{4-26}$$

And for the AIC we will have:

$$MIMOAIC_i = 2N ln(\hat{\sigma}_i^2) + \frac{2N}{2N - 1 - \sum_{j=1}^{M} p_j} 2\left(1 + \sum_{j=1}^{M} p_j\right) \tag{4-27}$$

4.2.2 MIMO system model order optimization

Obtaining the optimum degree for matrix memory polynomial models addressed in the paper necessitates minimizing MIMOAIC and MIMOBIC. Optimization of these integer variables $(K_{i,j}, Q_{i,j})$using a double-precision cost function can be cast as a mixed integer nonlinear programming problem.

In the case of MIMO system the situation is different. First, the problem may no longer be convex, since we have a larger number of parameters. As a result, alternative methods such as Branch and bound, cutting plane, dynamic programming, implicit enumeration, and Lagrangian relaxation and simulated annealing were proposed in the literature [61], [62] to solve this type of problem. Simulated annealing was shown to be a powerful method for optimizing mixed nonlinear integer programming. In this paper simulated annealing was used to optimize the MIMOAIC and MIMOBIC to obtain the optimum degrees and memory depths in the memory polynomial based models. Although the problem is not generally convex, the gradient descent algorithm was also applied. It was shown that the gradient descent could find a local optimum for which the performance of the estimated model is very close to the optimum value. This may come from the fact that the non-convexity of the problem is mostly due to the limitations arising

from the accuracy of the measurement set up. The gradient descent could find the local optimum in a few numbers of iterations and the results are provided to compare them with simulated annealing.

Simulated annealing first evaluates the cost function for an initial configuration. Then it perturbs the initial point with a random perturbation. Two cases will happen: the cost function will decrease which in this case, it will be considered as an acceptable movement. On the other hand if the cost function increases by ΔE then it will be accepted with probability $e^{(-\frac{\Delta E}{KT})}$. This property in simulated annealing avoids the algorithm to be trapped in local optimums. After each iteration, the temperature cools down with a scheme which limits uphill jumps at lower temperatures [62].The modified Simulated Annealing algorithm for integer parameters is explained in Figure 4-3.[62]

For our case where the variables are integers, each move should be integer too. Therefore, for the configuration update the following expression was used:

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + [\boldsymbol{e}_k T_k] \tag{4-28}$$

In which $\boldsymbol{e}_k$is a zero mean vector Gaussian random variable normalized by its norm. After checking the bounds on the parameter vector again, the integer part of the new parameters is used. The initial temperatures were set to T_0 = 100 and the cooling scheme was selected to be $T_k = T_0(0.95)^k$ [62].

4.3 MEASUREMENT RESULTS

4.3.1 Measurement Setup

The excitation signals for system identification and evaluation are modulated Wideband Code Division Multiple access (WCDMA) with signal bandwidth of 3.84 MHz. Four different signals with different number of carriers, and therefore different bandwidths, denoted as 1, 11,

111, 1111, have been used in the measurement validation. All these waveforms are 2-ms long and are sampled at the same frequency, 92.16 MHz. Table 4-1 summarizes the signal specifications for each carrier of these WCDMA signals. For modeling purposes 8000 points of the signals were used.

M is the maximum number of iteration

T_1 is the current temperature

Initialize the model with $\boldsymbol{\theta}_1$

Measure the cost function *E*

For *k= 1 to M*

Generate a zero mean Gaussian random vector $\boldsymbol{e}_k$

Consider the $[\boldsymbol{e}_k T_k]$ as the move vector

Update the parameters $\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + [\boldsymbol{e}_k T_k]$

If $\Delta E < 0$ then accept the move and update parameters and temperature

If $\Delta E > 0$ then accept the move with probability $e^{\left(-\frac{\Delta E}{KT}\right)}$ and update the parameters and temperature

End for loop

Figure 4-3-Simulated Annealing algorithm

Two different PAs, one class AB and one Doherty amplifier, are used in the course of measurement and performance evaluation.

Table 4-1- Modulated signals parameters

WCDMA Parameters
Band width = 3.84 MHz
Chips per Slot = 2560
Samples per Chip = 16
3GPP Test Model 1, As suggested in 3GPPTS 25.141 V3.9.0

The DUT in the class AB PA case is a cascade of a class A driver and a class AB PA biased at Vgs=9V,Vds=28V. Both blocks are designed using the TF10107 LDMOS transistor. The output signals were attenuated and captured by vector signal analyzer (VSA89650S).

The DUT in the Doherty case is a high power high efficiency Doherty PA designed by LDMOS transistors provided by Powerwave Technologies. It has a peak power of 300 W and its small signal gain is 61dB. The operating frequency is in the 2110–2170-MHz frequency band.

The measurement setup for the SISO and MIMO system are shown in Figure 2-5 and Figure 4-4, respectively. For the SISO system in Figure 2-5, it consists of a signal generator, a driver PA, PA, an attenuator and a vector signal analyzer (VSA) and for the MIMO case in Figure 4-4, it consists of two signal generators (ESG4438C), cascade of two amplifiers (Class A and Class AB) and a VSA (89650S). The two signals are downloaded to the signal generators using GPIB interface.

4.3.2 Results and discussion

The performance of AIC, BIC and the 2x2 MIMOAIC and MIMOBIC are examined and evaluated over single-branch and dual-branch PAs respectively.

Table 4-2 and Table 4-3 compares the performance of the memory polynomial model with the nonlinearity and memory depth order selected based on the AIC, BIC and maximum order of nonlinearity order of 16 and memory depth of 7 for class AB PA and Doherty PA. The maximum order is set to be greater than what was reported in [17].

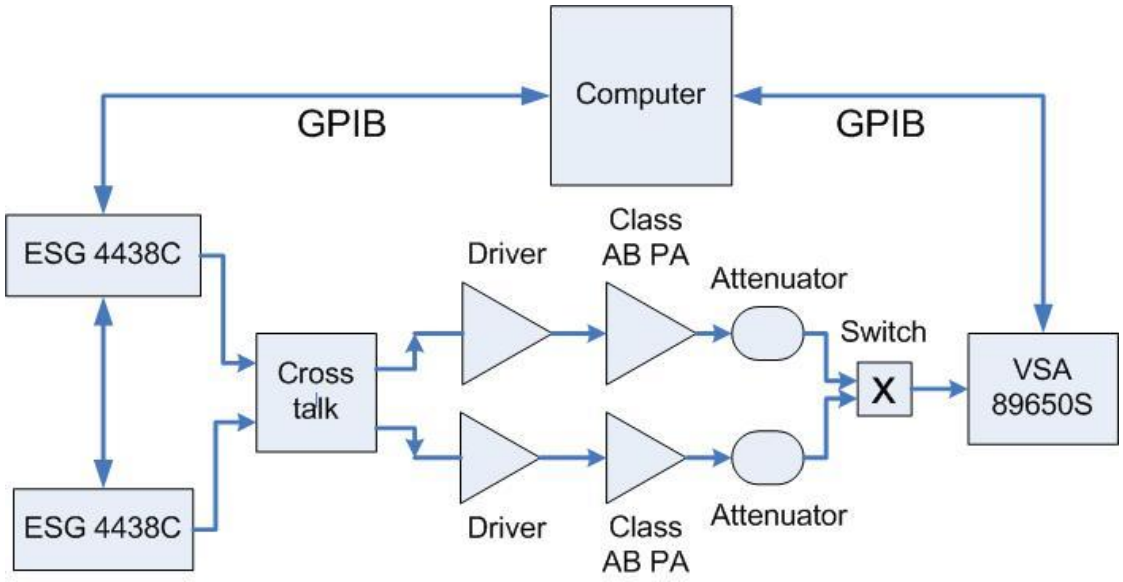

Figure 4-4- MIMO measurement setup

The calculated Normalized Mean Square Error (NMSE) and Adjacent Channel Error Power Ratio (ACEPR) metrics show that the performance of the model order selection based on BIC is very close to the maximum model order, but with less number of coefficients and processing complexity. The AIC criterion exhibits signs of over fitting. The ACEPRs are calculated based on the following formula:

$$Upper\ Channel\ ACEPR = \frac{\int_{f_c+B/2}^{f_c+3B/2} |X(f)|^2 df}{\int_{f_c-B/2}^{f_c+B/2} |X(f)|^2 df} \quad (4\text{-}29)$$

$$Lower\ channel\ ACEPR = \frac{\int_{f_c-3B/2}^{f_c-B/2} |X(f)|^2 df}{\int_{f_c-B/2}^{f_c+B/2} |X(f)|^2 df} \quad (4\text{-}30)$$

As can be seen from Table 4-2 and Table 4-3, when the number of carriers increases the order decreases but the memory effect increases. The models with optimized orders using the BIC criterion showed linearity results very close to the models with maximum order (K=16,Q=7) while reducing the complexity considerably. Compared to AIC, this method estimates more efficient orders. The AIC most of the time over-estimates the orders. As can be seen from Table 4-2 and Table 4-3 the same performance can be obtained by lower orders than what is estimated by BIC. This proves the validity of the new automated optimization approach.

Table 4-2- NMSE and ACEPR mesurements for the models with optimal orders compared to higher order models for class AB PA

PA Type		Class-AB			
WCDMA Signal Type		1	11	111	1111
Signal PAPR	[dB]	10.7	10.4	10.6	11.2
Compression	[dB]	4	4	3.5	4
AIC	(K, Q)	(11,6)	(11,5)	(12,6)	(11,6)
	# unknown	77	66	84	77
	NMSE	-46.7	-46.7	-46.1	-45.3
	ACEPR	-60.1	-52.0	-49.1	-47.8
BIC	(K,Q)	(11,2)	(11,4)	(11,5)	(10,6)
	# unknown	33	55	66	70
	NMSE	-46.6	-46.7	-46.0	-45.3
	ACEPR	-60	-52.0	-49.0	-47.8
Maximum (K=16,Q=7)	(K, Q)	(16,7)	(16,7)	(16,7)	(16,7)
	# unknown	128	128	128	128
	NMSE	-46.7	-46.7	-46.1	-45.4
	ACEPR	-60.1	-52.1	-49.1	-48.0

Table 4-3- NMSE and ACEPR mesurements of the the optimum models compared to the higher order models for Doherty PA

PA Type		Doherty			
WCDMA Signal Type		1	11	111	1111
Signal PAPR	[dB]	10.7	10.3	10.6	11.2
Compression	[dB]	4.5	4.5	4.5	4.5
AIC	(K, Q)	(10,3)	(9,5)	(9,5)	(10,6)
	# unknown	40	54	54	70
	NMSE	-39.5	-39.3	-38.3	-36.7
	ACEPR	-50.0	-44.9	-42.7	-40.9
BIC	(K, Q)	(10,3)	(9 ,4)	(9,4)	(10,5)
	# unknown	40	45	45	60
	NMSE	-39.5	-39.3	-38.2	-36.6
	ACEPR	-50.0	-44.9	-42.6	-40.9
Maximum (K=16,Q=7)	(K, Q)	(16,7)	(16,7)	(16,7)	(16,7)
	# unknown	128	128	128	128
	NMSE	-39.6	-39.3	-38.3	-36.7
	ACEPR	-50.0	-45.0	-42.7	-40.9

When the number of carriers increases, the higher bandwidth induces more memory effect. As can be seen form Table 4-2 the proposed optimization algorithm was able to predict

this difference and same nonlinearity order but higher memory depth (10,6) was obtained as the optimal model order for the four-carrier case.

Similarly, Table 4-4 summarizes the results for MIMOAIC and MIMOBIC with the cross-over model of the dual-branch transmitter in the presence of 15, 20 and 30 dB nonlinear crosstalk. The NMSE and ACEPR of the model selected based on MIMOAIC and MIMOBIC are compared with the maximum model order of nonlinearity order of 13 and memory depth of 6.

From the results presented, it is clear that the optimization using MIMOBIC was able to obtain very similar linearity performance compared to the maximum order (K=13,Q=6), while reducing the complexity of the models considerably. In Table 4-4 nonlinearity order 1 and memory depth 1 are the nonlinearity order and memory depth of the model 1-1 and nonlinearity order 2 and memory depth 2 are the nonlinearity order and memory depth of the model 1-2 in Figure 4-2.

When the cross-talk is low (-30dB), the model behaviour is close to a SISO system. For the -15 dB crosstalk case the memory effect is more noticeable. The reason for lower nonlinearity order comes from the fact that a smoother curve is enough to fit the data in the case of -15 dB crosstalk but more dispersion is generated by the PAs, which requires higher memory order. The amount of AM-AM compression which is the difference in dB between input and output peak amplitudes is reported in Table 4-4.

It is important to mention that not only the new optimization criteria is able to provide the user with an automated optimal order selection algorithm for any given amplifier, which guarantees a trade-off between linearity and complexity, it also allows to use an optimization algorithm to reduce considerably the number of iterations of finding this optimal value. Indeed,

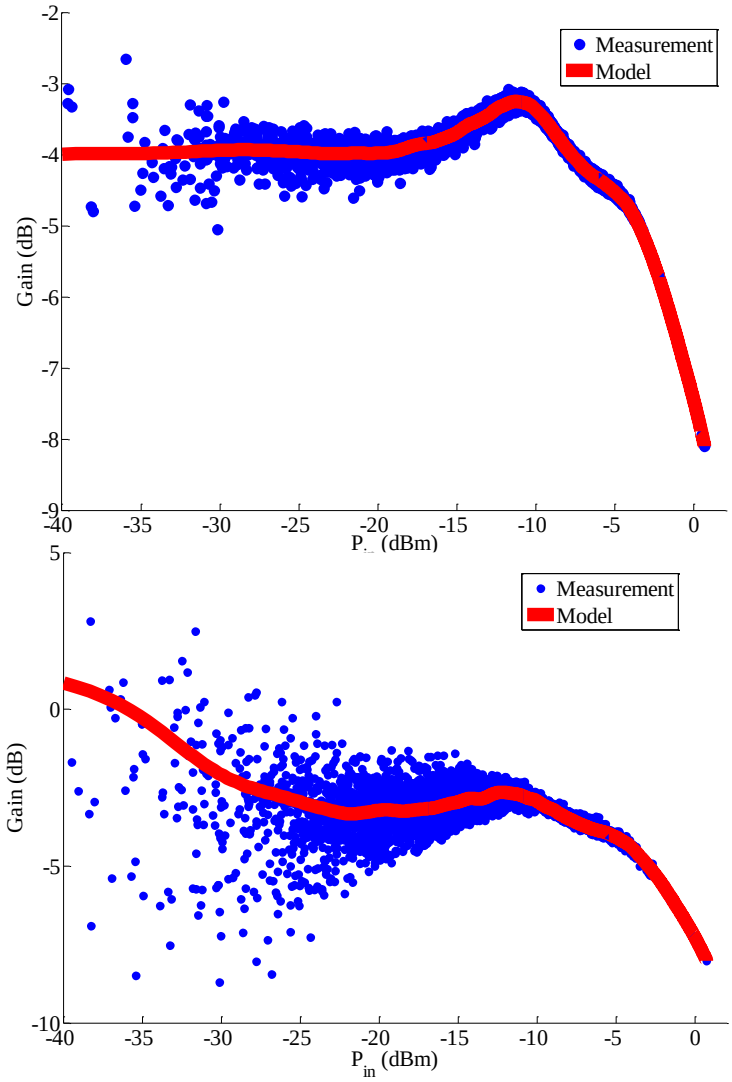

Figure 4-5- Doherty PA AM/AM for (top) one carrier and (bottom) four carrier signals

by using Gradient descent in the SISO case the number of iterations was reduced from 128 to maximum 30. In the case of MIMO system by using simulated annealing the number of iterations was reduced from 4096 to around 300 for the MIMO case. Using gradient descent for MIMO case shows that, although the global optimum could not be found, the performance of the local optimum is very close to the global optimum performance. The results of gradient descent

Table 4-4- NMSE of the optimum models compared to the higher order models for 2x2 MIMO systems

Crosstalk		-15dB	-20dB	-30dB
MIMOAIC	Order1	10	12	12
	Memory Depth1	9	5	8
	Order2	7	12	7
	Memory Depth2	9	7	7
	NMSE1	-36.0	-36.4	-43.5
	ACEPR	-47.6	-47.5	-56.3
MIMOBIC	Order1	10	8	11
	Memory Depth1	4	4	2
	Order2	7	6	2
	Memory Depth2	4	2	0
	NMSE1	-35.7	-36.1	-43.1
	ACEPR	-47.1	-47.3	-55.3
Maximum (K=13,Q=6)	Maximum NMSE1	-36.0	-36.4	-43.5
	Maximum ACEPR	-47.3	-47.5	-56.0
Signal PAPR	[dB]	10.00	10.7	10.7
PAPR with X-talk	[dB]	10.4	10.7	10.7
Compression	[dB]	2.5	2	4

Table 4-5- NMSE of the optimum models obtained using gradient descent for 2x2 MIMO systems

Crosstalk		-15dB	-20dB	-30dB
MIMOBIC	Order1	10	8	10
	Memory Depth1	6	2	2
	Order2	7	3	2
	Memory Depth2	5	2	0
	NMSE	-35.8	-35.9	-43.1
	ACEPR	-47.2	-47.0	-55.3
GD	# unknown	112	33	32
SA	# unknown	85	58	35

are compared to simulated annealing in Table 4-5. This optimum search algorithm reduces the number of iterations to 50 from 4096. Therefore, it is recommended to use the gradient descent even in the MIMO case.

4.4 Conclusion

In this chapter, the application of information criteria, namely AIC and BIC for model order selection in SISO and MIMO transmitters was presented. New criteria based on AIC and BIC, which includes model complexity as well as the NMSE metric, were proposed for MIMO dynamic nonlinear systems. To obtain the optimum model orders two optimization algorithms were used for SISO and MIMO systems. In the SISO case, a modified gradient descent algorithm was used. For the MIMO system since the problem is not generally convex and the number of

parameters was higher, the simulated annealing algorithm was applied first and its performance was compared to the gradient descent algorithm. It was shown that the gradient descent in MIMO was able to obtain a local minimum with performance very close to the global optimum. It was also shown that AIC based criteria over-estimates the orders and is not efficient. Also, it was shown that with a negligible difference in NMSE and ACEPR, the model complexity is considerably reduced using BIC based criteria. As a conclusion, using the proposed criteria and the gradient descent optimum search algorithm, the optimum model order can be extracted automatically without the need for a complete order sweep. The number of iterations is therefore reduced from 128 to 30 for the SISO systems and from 4096 to 300 for the MIMO systems.

Chapter Five: Estimation of Crossover DPD Using Orthogonal Polynomials in Fixed Point Arithmetic

5.1 Introduction

To guarantee high quality service and reliability of links for the customers, modern communication systems demand higher channel capacity. MIMO systems have been considered as one of the most promising solutions to increase the channel capacity and/or improve the link availability [63], [64], [65].

In MIMO systems multiple transmitters have to be integrated on a single chip. However, due to the imperfect isolation of integrated microwave components of the transmitters, unwanted effects occur which degrade system performance. Crosstalk is one of the major effects that limit the output spectral efficiency. The effects of crosstalk on the overall performance of the system and that of the linearization process of the 2x2 MIMO systems have been explained in [20], [66], [67].

PA nonlinearity increases the bit error rate of communication systems, due to both in-band and out-of-band distortions. DPD techniques have been proved to be powerful tool in the literature and have been used in commercial products to resolve the PA nonlinearity problem [9], [68], [69].

It was shown that the joint effect of radio frequency (RF) nonlinear crosstalk and PA nonlinear distortion can considerably degrade the performance of MIMO transmitters [20]. To address this problem, a Co-DPD architecture has been presented and developed [20]. This architecture can simultaneously compensate for the crosstalk and PA nonlinearities in the MIMO transmitter.

The crossover DPD scheme is established based on the memory polynomial model. However, it has been shown that the matrix inversion of memory polynomial based models suffer from an inherent instability issue [70], [71]. In this regard, orthogonal polynomials have been advised as the best solution to the rank deficient least squares (LS) estimation [70], [71]. The extension of orthogonal polynomials to the crossover DPD model was developed and tested for floating-point digital signal processors (DSPs) [60].

Fixed-point processors have been widely used in many digital signal processing applications and include a large portion of DSPs. Fixed-point processors have benefits comparing to their floating-point counterparts, in terms of cost, speed, power consumption, precision, volume and complexity of the mathematical operation [72].

This chapter investigates the implementation of crossover DPD based on conventional memory and orthogonal polynomial models in fixed-point processors. It is revealed that, in fixed-point arithmetic, the orthogonal polynomial resolves the numerical instability problem which occurs when the memory polynomial is used. To implement the matrix inversion in fixed point, LU decomposition was employed. Then using a triangular matrix inversion algorithm the inverse matrix was calculated [73], [74], [75].

The other issue regarding the Co-DPD is that the coefficient extraction method is not well-organized and it is feasible to simplify the algorithm. In this chapter, the cross-over DPD estimation algorithm is simplified such that the overall computational complexity is improved to the half of the conventional technique complexity. The performance of the new algorithm is then compared to cross-over DPD.

5.2 Modeling Nonlinearities in MIMO Systems

5.2.1 Fixed-point digital signal processors

Choosing the proper arithmetic for digital signal processing applications is not a trivial task. Several factors can help in selecting the optimal processor, among which the sampling rate, the complexity of the algorithm, the dynamic range and the production volume are the most effective ones.

It is well known that fixed-point arithmetic is optimum for simple algorithms with low sampling rates. On the other hand, more complicated algorithms with higher sampling rates performs better with floating-point representation. For higher dynamic range numbers, a floating-point DSP is the best choice; and, for a smaller chip size of the product, a fixed-point DSP should be used.

Some other metrics must be considered carefully when comparing the performance of different systems. Cost, speed and power consumption are the most important factors. Due to the more complex mathematical operations, floating-point processors are more expensive than their fixed-point counterparts. When the data can be handled within the acceptable limit of a fixed-point DSP, it is the best choice according to the speed; otherwise, the extra cost of a floating-point DSP should be paid, in order to achieve an acceptable speed. Due to simpler mathematical operations and less complicated hardware, the power consumption in fixed-point DSPs is less than floating-point DSPs, [72].

Based on the above discussion, fixed-point signal processing brings about a fairly high processing speed, lower cost, lower device volume, lower power consumption and likely satisfying precision when the data can be scaled to be within the range of a fixed-point DSP.

5.2.2 Power amplifier model identification

The most influential factor in the overall signal quality of a transmitter is the linearity of the PA. As mentioned earlier, DPD is an effective and practical way to compensate for nonlinear distortion of the PA. The DPD technique consists of three major stages: 1) proper selection of the nonlinear behaviour model for the nonlinear PA; 2) model coefficients' estimation and extraction based on captured data from the PA; and, finally, 3) predistortion of the input signal using the estimated inverse model.

5.2.3 Linear and nonlinear crosstalk

To take advantage of MIMO transceivers, some vendors provide integrated multiple transmitters on a single chip. The limited isolation between different ports of microwave components causes crosstalk in the transmitting signals. The crosstalk can cause in-band and out-of-band distortions. Linear crosstalk never passes through nonlinear components. On the other hand, nonlinear crosstalk is a coupling that passes through nonlinear components and introduces both in-band and out-of-band distortions. The effect of linear crosstalk can be canceled out by linear matrix inversion, and the nonlinear crosstalk effect can be effectively compensated using the crossover DPD model [19].

5.2.4 Orthogonal polynomials

There are two classes of methods in designing DPDs: recursive and estimate-plug methods. Recursive methods are based on recursive least squares (RLS) estimations and employ QR decomposition. Recursive methods are useful for non-stationary environments. When the system can be considered stationary, DPD can be designed in the same way as Wiener filters are designed. The estimation of the crossover DPD model can be obtained using the LS solution. However, the matrix inversion step inherent in the LS estimation can be numerically unstable for

crossover DPDs designed based on conventional memory polynomials [70], [71]. The solution for this problem is the orthogonal polynomial crossover DPD described in this chapter.

5.3 Modeling MIMO Systems using Orthogonal Polynomials

The effect of crosstalk at the output of 2x2 nonlinear MIMO transmitters with a coupling factor of α can be modeled as follows:

$$\begin{aligned} y_1 &= f_1(x_1 + \alpha x_2) \\ y_2 &= f_2(x_2 + \alpha x_1) \end{aligned} \tag{5-1}$$

where x_1 and x_2 are the baseband input signals to the MIMO transmitter, and y_1 and y_2 are the signals at the output of the MIMO transmitter.

The crossover DPD was introduced in [20] which can compensate for both the crosstalk and the nonlinear effects of each branch of the MIMO transmitter. The crossover DPD was extended to the orthogonal polynomial based model for floating-point DSPs in [60].

Considering both nonlinear crosstalk and PA nonlinearity, the crossover DPD, as shown in

Figure 5-1, can be implemented to simultaneously compensate for both sources of distortions. The crossover DPD consists of four sub-blocks. Each of these sub-blocks is a nonlinear function of one of the inputs.

In this case, the pre-distorted signals at the output of the crossover DPD are as follows:

$$\begin{aligned} y_1 &= f_{11}(x_1) + f_{12}(x_2) \\ y_2 &= f_{21}(x_1) + f_{22}(x_2) \end{aligned} \tag{5-2}$$

Based on our choice for the nonlinearity model, an orthogonal polynomial the general form is of $f_{ij}(x)$ can be expressed by

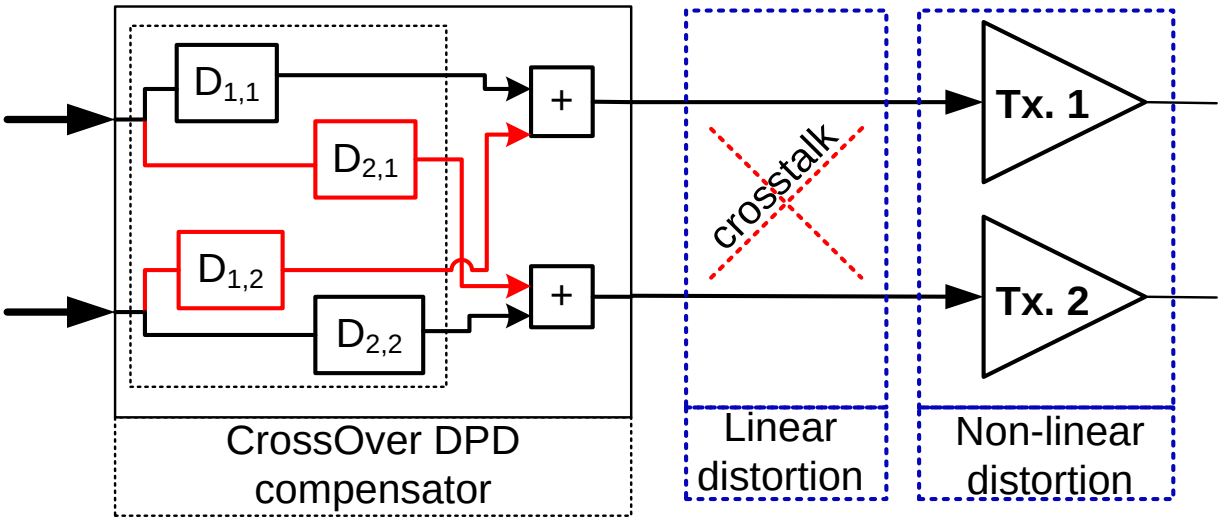

Figure 5-1- The architecture of the 2x2 nonlinear MIMO transmitters in the presence of nonlinear crosstalk

$$y(n) = \sum_{m=0}^{Q} \sum_{i=1}^{K} h_{im} \sum_{j=1}^{i} u(i,j) x(n-m) |x(n-m)|^{j-1} \qquad (5\text{-}3)$$

where $x(n)$ is the input signal, $y(n)$is the output complex signal, h_{im}s are the polynomial coefficients of the i^{th} degree with a memory depth of m, K denotes the maximum polynomial order, and Q denotes the maximum memory depth of the polynomial. The above representation is the general form of the polynomial model. It will be the orthogonal polynomial model when $u(k,i)$ is considered as [70]:

$$u(i,k) = (-1)^{k+i} \frac{(k+i)!}{(i-1)!\,(i+1)!\,(k-i)!} \qquad (5\text{-}4)$$

and the memory polynomial model when it is considered as:

$$u(i,k) = \delta(i,k) \tag{5-5}$$

where the $\delta(.,.)$ is the Dirac delta function.

In this thesis, the crossover DPD models based on both the memory polynomial and orthogonal polynomial models are compared; and, the advantages of the orthogonal polynomial over the memory polynomial in fixed-point analyses are explored.

The critical operations in the crossover DPD model identification are the coefficient extraction and matrix inversion. The maximum likelihood (ML) nonlinear model identification problem can be cast as a LS problem when the model error is considered as a Gaussian random process. Wideband modulated signals, such as WCDMA signals, can be assumed as Gaussian distributed signals; therefore, it is acceptable to treat the error as a Gaussian process. In this case, the LS estimation of the polynomial model would be an optimal solution in the ML sense.

For the 2x2 MIMO system, the estimation equations can be developed as follows:

$$[\boldsymbol{y}_1, \boldsymbol{y}_2] = [\Psi_{x1}, \Psi_{x2}] \begin{bmatrix} \boldsymbol{h}_{11} & \boldsymbol{h}_{12} \\ \boldsymbol{h}_{21} & \boldsymbol{h}_{22} \end{bmatrix} \tag{5-6}$$

where $\boldsymbol{y} = [y(Q+1) \cdots y(N)]^T$

$$\boldsymbol{h} = \left[h_{1Q}, h_{1Q-1}, \dots, h_{10}, \dots, h_{KQ}, \dots, h_{K0}\right]$$

$$\Psi_x = [S_x^0, \dots, S_x^K]$$

$$S_x^k = \begin{bmatrix} \psi_k(x(Q)) & \dots & \psi_k(x(0)) \\ \vdots & \ddots & \vdots \\ \psi_k(x(N)) & \dots & \psi_k(x(N-Q)) \end{bmatrix} \tag{5-7}$$

Where

$$\psi_k\big(x(n)\big) = \sum_{j=1}^{k} u(k,j)x(n)|x(n)|^{j-1} \tag{5-8}$$

5.4 LS Estimation Problems using the Conventional Polynomial Model

The LS estimation of the model parameters in (5-6) is an over-determined system of equations:

$$\boldsymbol{y} = \boldsymbol{\Psi h} \tag{5-9}$$

For the full rank $\boldsymbol{\Psi}$ matrix, the solution is a projection of vector **y** onto the vector space of the $\boldsymbol{\Psi}$ columns and the estimation is:

$$\boldsymbol{h} = (\boldsymbol{\Psi^H \Psi})^{-1}\boldsymbol{\Psi^H y} \tag{5-10}$$

In the case of orthogonal polynomials, the $\boldsymbol{\Psi}$ matrix is diagonal and is almost full rank [64, 65]. However, in the case of memory polynomials, it is a Hilbert matrix, which is a well-known rank deficient matrix.

Figure 5-2 compares the rank versus matrix size of the $\Delta = \boldsymbol{\Psi}^H\boldsymbol{\Psi}$ matrix for both memory and orthogonal polynomials. For the orthogonal polynomial based matrix, the rank of the $\Psi^H\Psi$ matrix increases linearly with the matrix size; whereas, for the memory polynomial case, the rank of the matrix is bounded to 30 for a matrix size greater than 30. These results indicate that the orthogonal polynomial based matrices are close to full rank and, in fact, invertible.

Based on these facts, it can be deduced that the orthogonal polynomial compensates for the numerical instability problems caused by practical applications of DPD design based on conventional memory polynomials. In fixed-point processors, there are a fixed number of bits to

represent numbers, so dealing with matrix inversions with high condition numbers is a critical issue in fixed-point arithmetic.

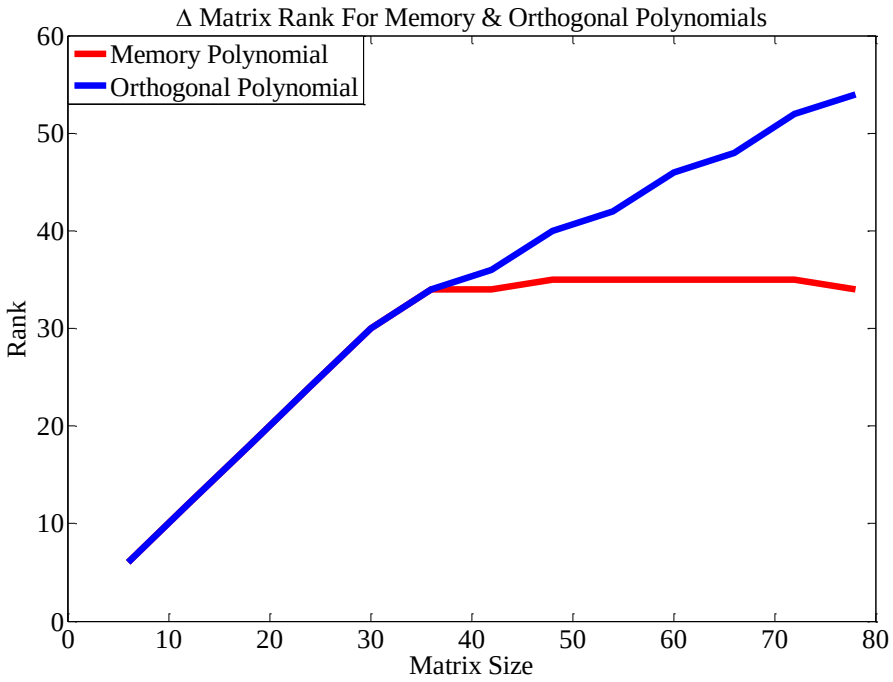

Figure 5-2- Δ matrix rank versus matrix size for memory and orthogonal polynomial models

5.5 Matrix inversion

Matrix inversion problem is the main issue in obtaining LS solution. A variety of matrix inversion techniques have been proposed in the literature which are based on matrix decompositions [73]. These decompositions include LU, QR and SVD. Among these methods, LU and QR seem to be the simplest decomposition algorithms [73].

LU decomposition results in lower and upper triangular matrices. For a rectangular matrix Δ with size n the LU decomposition algorithm [73] is shown in Figure 5-3. All the operations are done using fixed point arithmetic. Other factorization methods, such as QR factorization using Gram-Schmidt or House-Holder transforms [73], were also tested. However the results of the QR factorization based inversion were unsatisfactory. On the other hand LU decomposition could successfully perform the matrix inversion task.

In this chapter, the LU decomposition is utilized for the matrix inversion problem; and, by using an algorithm for triangular matrix inversion implemented in fixed point arithmetic, the coefficients of the orthogonal polynomial model were successfully extracted.

The matrix inversion is accomplished as follows:

$$\boldsymbol{\Delta = \Psi^H \Psi = LU}$$

$$\boldsymbol{f_{inv}(\Delta) = f_{tinv}(U) f_{tinv}(L)} \tag{5-11}$$

Where $\boldsymbol{f_{inv}(.)}$ denotes the fixed point matrix inversion operator and $\boldsymbol{f_{tinv}(.)}$ represents the fixed point triangular matrix inversion operator. The results indicate that the LU decomposition is the most robust solution.

5.6 Measurement Setup and Results

The measurement setup included two baseband time-aligned signal generators (ESG4438C), two sets of PAs and a spectrum analyzer (E4440A) and a vector signal analyzer (VSA89650S). Two sets of modulated 3.84 MHz 3GPP signals were used for the measurement. The signals' length was 100,000 symbols, in which 8,000 symbols around the maximum magnitude were employed in the model identification; and, the rest of the symbols were used in the measurement evaluation

The fixed-point toolbox of MATLAB was used for fixed-point processing of the DPD linearization technique.

$$for\ k = 1{:}n-1$$

$$rows = k+1{:}n$$

$$\Delta(rows,k) = \frac{\Delta(rows,k)}{\Delta(k,k)}$$

$$\Delta(rows,rows) = \Delta(rows,rows) - \Delta(rows,k)\Delta(k,rows)$$

$$end$$

$$T(i,j) = \Delta(i,j) \qquad for\ i > j$$

$$U(i,j) = \Delta(i,j) \qquad for\ i \leq j$$

$$L, I = I_n$$

$$for\ k = 1{:}n$$

$$L = L + T(:,k)I(k,:)$$

$$end$$

Figure 5-3- LU decomposition algorithm

The results shown in this section are based on 64-bit fixed-point processing with 32-bit fractional length. The normalized mean square error (NMSE) of the inverse model based on memory and orthogonal polynomials for different orders of nonlinearity are shown in Figure 5-4. It is clear from the NMSE results that, for memory polynomials, increasing the nonlinearity order led to divergence of the inverse modeling and unacceptable NMSE performance. However, the NMSE of the orthogonal polynomial model monotonically decreased when the nonlinearity order was increased.

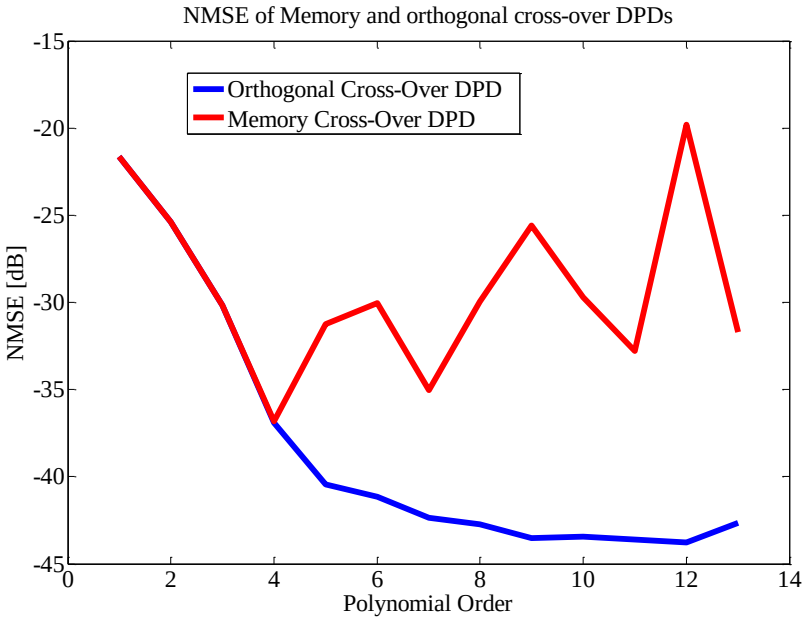

Figure 5-4- NMSE for inverse model in the cases of memory and orthogonal polynomial models in fixed-point arithmetic

The performance of the orthogonal polynomial model was verified for nonlinear distortion compensation of a dual-branch transmitter in the presence of -30 and -15 dB nonlinear crosstalk. Crossover DPD linearization based on both memory and orthogonal polynomials were compared for a nonlinearity order of 8 and 13 and a memory depth of 5.

Figure 5-5and Figure 5-6 show the power spectrum density (PSD) of one of the branches of the dual-branch MIMO transmitter. The PSD of the following scenarios have been plotted: 1) the input of the transmitter (green); 2) the output of the nonlinear transmitter (blue); 3) the output of the linearized transmitter using memory polynomial crossover DPD (black); and, 4) the output of the linearized transmitter using orthogonal polynomials (red). In Figure 5-5and Figure 5-6, the nonlinear crosstalk was -30 dB and -15 dB, respectively.

The results in Figure 5-5and Figure 5-6 demonstrate the fact that, in fixed-point processing, the rank deficiency and matrix instability of the memory polynomial model significantly reduced the linearization performance. Moreover, increasing the nonlinearity order for the memory polynomial model had double effects. It helped to get better estimates, but simultaneously accelerated the matrix instability issue. Some oscillations in the NMSE could be seen, indicating these two effects. However, this was not the case for crossover DPD linearization based on orthogonal polynomials, since for higher nonlinearity orders, the

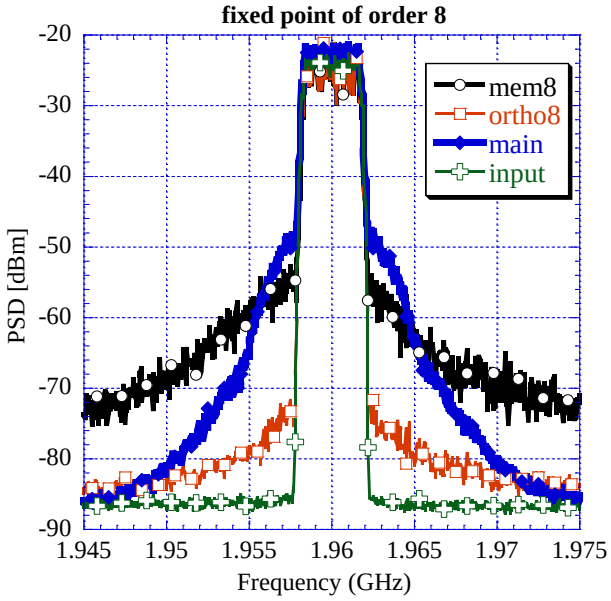

Figure 5-5- PSD of the amplifier with/without linearization for -30 dB crosstalk

linearization performance improved. Finally, Table 5-1 lists the measured ACPR of the output signal at 5 MHz frequency offset. In the table, the measured ACPR decreases from -42.28 dBc for the nonlinear transmitter to -51.62dBc and -52.03 dBc for the orthogonal polynomial model with nonlinearity orders of 8 and 13, respectively. In contrast, the measured ACPR rank deficiency of the

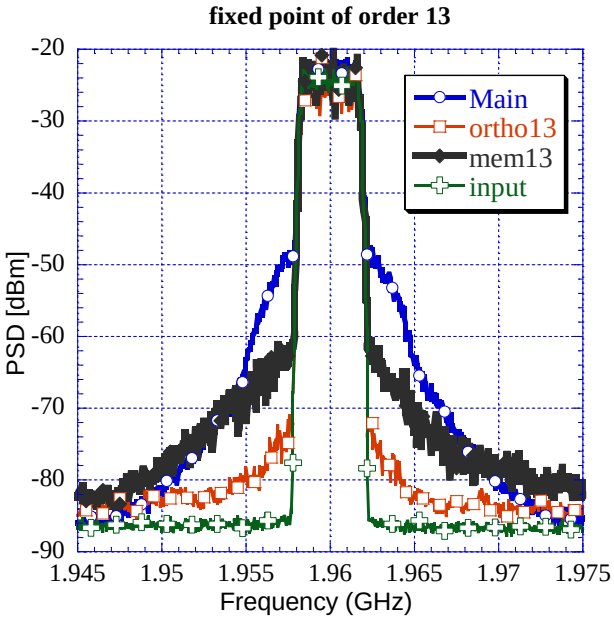

Figure 5-6- PSD of the amplifier with/without linearization for -15 dB crosstalk

increases to -33.03 dBc and -42.73 dBc for the memory polynomial model with nonlinearity orders of 8 and 13, respectively. The measured ACPR values show that the memory polynomial based crossover DPD model in fixed-point realization degraded the linearity of the system and also introduced additional out-of-band distortions.

Table 5-1-The measured ACPR at 5 MHz frequency offset for memory and orthogonal polynomial models with fixed-point arithmetic

ACPR [dBc]	No Linearization	Memory Polynomial		*Orthogonal Polynomial*	
		order 8	order 13	order 8	order 13
Lower Band	-42.28	-33.03	-42.73	-51.62	-52.03
Upper Band	-42.6	-35.84	-44.76	-52.83	-54.66

5.7 The Simplification of Co-DPD Parameter Extraction

Based on the feedback from the pilot study, it was concluded that the identification algorithm of the Co-DPD is not computationally efficient. To circumvent this computational complexity, the model was reorganized to take the advantage of possible simplifications. Let's denote the DPD output by z, the PA output by y and the input signal to the DPD by x, then for the output signals we have:

$$\begin{aligned} y_1 &= f_1(z_1 + \alpha z_2) \\ y_2 &= f_2(z_2 + \alpha z_1) \end{aligned} \tag{5-12}$$

After inverting the functions it can be written as follows:

$$\begin{aligned} z_1 + \alpha z_2 &= f_1^{-1}(y_1) \\ z_2 + \alpha z_1 &= f_2^{-1}(y_2) \end{aligned} \tag{5-13}$$

By solving the linear system it can be concluded that:

$$\begin{aligned} z_1 &= \frac{1}{1-\alpha^2}(f_1^{-1}(y_1) - \alpha f_2^{-1}(y_2)) \\ z_2 &= \frac{1}{1-\alpha^2}(f_2^{-1}(y_1) - \alpha f_1^{-1}(y_2)) \end{aligned} \tag{5-14}$$

When the DPD compensates the MIMO system distortions then the output will be similar to the input signal and we will have:

$$\begin{aligned} z_1 &= \frac{1}{1-\alpha^2}(f_1^{-1}(x_1) - \alpha f_2^{-1}(x_2)) \\ z_2 &= \frac{1}{1-\alpha^2}(f_2^{-1}(x_1) - \alpha f_1^{-1}(x_2)) \end{aligned} \tag{5-15}$$

This analysis show that the 2x2 DPD can be estimated by proper combinations of two 1x1 DPDs. The main issue about this method lies in this fact that the coupling factor should be assessed prior to 2D DPD estimation. To solve this problem, we suggested estimating the

coupling factor, from undistorted part of the output signals. At sufficient back off, PA can be considered linear, one can find the coupling factor through least squares.

$$\acute{y_1} = \acute{x_1} + \alpha \acute{x_2}$$
$$\acute{y_2} = \acute{x_2} + \alpha \acute{x_1} \quad (5\text{-}16)$$

where the prime represents the undistorted part of the signal.

To compare the computational cost of the proposed technique alongside the conventional technique, algorithm complexities should be calculated. As demonstrated in [73] the number of flops of Singular Value Decomposition (SVD) for $m \times n$ matrix is of order $O(mn^2)$. Considering the fact that the matrix size for the proposed technique is half of the conventional one, the number of flops reduces from $O(m(2n)^2)$or $O(4mn^2)$ to $O(2mn^2)$.

5.7.1 Simulation and Measurement Results

Many challenges appear in MIMO system integration. Among them nonlinear cross-talk along with the PA nonlinearity plays an integral role in performance degradation. To avoid this degradation, predistortion seems to be crucial. Among 2x2 DPDs, the Co-DPD proposed in [20] has been found a promising tool for compensating the above mentioned distortions. However, the identification process looks inefficient. In order to simplify the coefficient extraction method, the problem has been reformulated. It is shown that the 2x2 DPD can be built using 2 1x1 DPDs. The main advantage of this simplification can be sensed in the computation time of the DPD estimation process which is close to one fourth of the conventional technique. Although the proposed approach is faster but the accuracy of the new DPD is not compromised.

A. Simulation and measurement Setup

The measurement setup included two baseband time-aligned signal generators (ESG4438C), two sets of PAs and a spectrum analyzer (E4440A) along with a vector signal analyzer (VSA89650S). Two sets of modulated 3.84 MHz 3GPP signals were used for the measurement. The length of the signals was 100,000 symbols sampled at 61.44MHz, in which 8,000 symbols around the maximum magnitude were employed in the model identification; and, the rest of the symbols were used to evaluate the measurements. The DUTs are two class AB PAs designed by CGH40010 ,10W GaN transistors and biased at -2.7V, with a Vds of 28V. Figure 4-4 illustrates the measurement set up used for MIMO nonlinear system modeling and compensation.

B. Simulation and measurement results

Table 5-1 and Table 5-2 list the performances of the previously mentioned methods. The comparison has been made for two levels of cross-talk 20 dB and 10 dB. Table 5-2 compares the NMSEs after linearization. It is evident that the proposed technique could reasonably compensate for the distortions. The 3dB difference in the NMSE comparing to the conventional technique can be assigned to the error in the coupling factor estimation. The results of the ACPR are shown in Table 5-3.

The ACPR results show that the proposed technique could notably compensate for the out of band distortions and the results seem close to the conventional Co-DPD.

Overall, the proposed technique could show a performance close to the conventional Co-DPD; however by taking half of the computational resource.

Table 5-2-NMSE results for the new DPD and the conventional DPD

	New CoDPD	CoDPD
NMSE(dB) 20dB Xtalk	-40.4	-39.4
NMSE(dB) 10dB Xtalk	-29	-32

Table 5-3- ACPR results for the new DPD and the conventional DPD

	New CoDPD	CoDPD	No DPD
ACPR H(dBc) 20dB Xtalk	-62	-62	-42.5
ACPR L(dB) 20dB Xtalk	-61.5	-61.5	-41.9
ACPR H(dBc) 10dB Xtalk	-59.5	-60.7	-41.2
ACPR L(dB) 10dB Xtalk	-60.2	-61.4	-41.6

5.8 Conclusions

This chapter has explored the effect of fixed-point implementation on the performance of crossover DPD linearization. The memory and orthogonal polynomial based basis functions have been considered and compared. It was shown that, in fixed-point processors, the problem of estimating the inverse function for nonlinear distortion compensation can be achieved by using orthogonal polynomial basis functions. It was shown that the orthogonal polynomial model is close to full rank, but the memory polynomial model is rank deficient. The measured ACPR

decreased from -42.28 dBc to -52.03dBc for the orthogonal polynomial model, but increased to -33.03dBc for the memory polynomial model. The results indicate that, using fixed-point arithmetic, orthogonal polynomial crossover DPD had improved the linearity of the signal, while the memory polynomial model was unable to properly linearize the MIMO system.

Later in this chapter, the Co-DPD parameter extraction method was simplified. It was shown that the performance of the proposed technique is advantageous over the conventional one according to accuracy and computational cost.

Chapter Six: Conclusion

6.1 Summary and Final Remarks

This dissertation has presented and proposed several innovative signal processing techniques suitable for the estimation and compensation of different types of distortions and nonlinearities in SISO and MIMO wireless transceivers.

In chapter Two a new architecture for PA distortion compensation has been proposed for uplink applications. In the proposed distributed distortion compensation technique, the conventional class AB PA was substituted with high efficiency Doherty PA. The proposed configuration compensates for the phase distortions using non-adaptive LUT-based phase-only DPD at the transmitter. The amplitude nonlinearity was compensated at the receiver after channel equalization and it does not require any training sequence. It was shown that the distributed distortion compensation provides almost the same linearity and reasonable BER performance as conventional techniques, but with considerably improved DE.

Chapter Three presents a new soft clipping technique using polynomial function. This study has been shown that the proposed technique could result in improvements in both efficiency and EVM. The SCCFR method could improve the EVM while having the same efficiency as CLF. To further decrease the PAPR, two methods were combined which resulted in lower PAPR and higher efficiency. The most important finding is that the combined structure using CDF based post compensation could pass the EVM requirement. However, the CLF method with a higher PAPR could not meet the EVM threshold.

The techniques proposed in the second and third chapters to compensate for the nonlinear distortion and enhance the power efficiency, namely the partitioned distortion compensation technique between the transmitter and receivers for the uplink scenario and the soft clipping CFR technique for the transmitter in the downlink scenario can be implemented in existing networks without the need to upgrade the hardware of existing transceivers since both methods are fully software based. Since these techniques all fully implemented in the software on the I and Q base-band waveforms subsequent to the coding and modulation processing at the transmitter and/or preceding to the de-coding and de-modulation processing in the receiver, they can be practically implemented in the wireless radios regardless of the standard. Considering this fact, all the current and future standards such 3G, 4G and 5G can employ and benefit from the proposed techniques.

In chapter Four the application of information criteria, AIC and BIC for model order selection in SISO and MIMO transmitters was presented. New criteria based on AIC and BIC, which includes model complexity as well as the NMSE metric, were proposed for MIMO dynamic nonlinear systems. To obtain the optimum model orders two optimization algorithms were used for SISO and MIMO systems. Based on the convexity of the problem in the SISO case, a modified gradient descent algorithm was used. For the MIMO system since the problem is not generally convex and the number of parameters was higher, the simulated annealing algorithm was applied first and its performance was compared to the gradient descent algorithm. It was shown that the gradient descent in MIMO was able to obtain a local minimum with performance very close to the global optimum. It was also shown that AIC based criteria over-estimates the orders and is inefficient. Also, it was shown that with a negligible difference in NMSE and ACEPR, the model complexity is considerably reduced using BIC based criteria. As

a conclusion, using the proposed criteria and the gradient descent optimum search algorithm, the optimum model order can be extracted automatically without the need for a complete order sweep. The number of iterations is therefore reduced from 128 to 30 for the SISO systems and from 4096 to 300 for the MIMO systems.

Chapter Five has investigated the effect of fixed-point implementation on the performance of crossover DPD linearization. The memory and orthogonal polynomial based basis functions have been considered and compared. It was shown that, in fixed-point processors, the problem of estimating the inverse function for nonlinear distortion compensation can be accomplished by using orthogonal polynomial basis functions. It was shown that the orthogonal polynomial model is full rank, but the memory polynomial model is rank deficient. The measured ACPR decreased from -42.28 dBc to -52.03dBc for the orthogonal polynomial model, but increased to -33.03dBc for the memory polynomial model. The results indicate that using fixed-point arithmetic, orthogonal polynomial crossover DPD had a positive impact on the linearity of the signal, while the memory polynomial model was unable to properly linearize the MIMO system.

Later in chapter Five, a simplified Co-DPD extraction technique is proposed. Although the new coefficient extraction method has less computational cost the accuracy is not compromised. The performance of the new estimation technique is compared to the conventional approach.

6.2 Recommendations for Future Work

The main focus of this dissertation is to provide practical solutions to compensate for PA nonlinear distortions. The motivation of this dissertation was to improve the overall performance of the transceiver through digital signal processing techniques. Therefore the

present thesis can help researchers working in both academia and industry employ novel practical and effective tools for compensating the non-ideal behaviour of RF front ends. These signal processing techniques improve the quality of the signal while preserving power and spectral efficiencies in SISO and MIMO systems. The possible future directions of this thesis should focus on employing advanced DSP techniques to improve the performance of the proposed ideas.

The performance of the partitioned distortion compensation technique proposed in chapter one, highly depends on the channel equalization accuracy. The channel equalization technique discussed in section 2.4.2.3 can be replaced with a more accurate approach to enhance the overall performance of the method.

The optimized crest factor reduction method described in chapter Three, assumed an ideal line of sight channel. The extension of this idea to the fading plus AWGN channel will be promising.

In chapter Four new model order selection has been proposed which are based on information criteria. The error term associated with the criteria can be limited to in-band or out-of-band distortions. The other possibility is to apply it on subspace methods.

Chapter Five mainly deals with MIMO nonlinear systems. The fixed point performance of the Co-DPD algorithm showed a good indication for possible FPGA implementation.

APPENDIX A

As discussed in chapter three the spectrum of a modern communication signal is considered to be the summation of two pulses. Accordingly the problem of predicting the spectrum after nonlinearity will result in finding the *2m+1* times convolution of a pulse. To obtain the *2m+1*th order convolution of a pulse, two different theories were employed. The first expression originates from the Irwin-Hall distribution function. The Irwin-Hall distribution is actually the PDF of the sum of n independent uniform random variables, distributed uniformly between 0 and 1. Since the PDF of sum of random variables is the convolution of their PDFs [49], then the *2m+1*th order spectrum can be obtained from the Irwin-Hall distribution [50].

The Irwin-Hall theorem states that the PDF of a summation of n uniform random variables U_k with uniform distribution $U(0,1)$, will take the following form [50]:

$$u = \sum_{k=1}^{n} U_k \tag{A-1}$$

$$S(u,n) = PDF_u(u,n)$$

$$= \frac{1}{2(n-1)!}\sum_{k=0}^{n}(-1)^k \binom{n}{k}(u-k)^{n-1} sgn(u-k) \tag{A-2}$$

Hence the n^{th} order spectrum of a pulse approximated signal spectrum can be described as below:

$$S_{2u(n)}(f) = A^n \frac{f_s}{W}\left(\frac{W}{f_s}\right)^n S\left(\frac{f}{W}+\frac{n}{2}\right) = A^n \left(\frac{W}{f_s}\right)^{n-1} S\left(\frac{f}{W}+\frac{n}{2}\right) \tag{A-3}$$

where $S_{2u(n)}$ denotes the n^{th} order spectrum of u.

where A is the average amplitude of the spectrum of the input signal, measured in the occupied bandwidth, f_s is the sampling frequency of the input signal and W is the occupied bandwidth of the signal.

The overall spectrum can be obtained using the following expression:

$$S_{2y}(f) = \sum_{m=0}^{K} \alpha_{2m+1} \frac{A^{2m+1}\left(\frac{W}{f_s}\right)^{2m}}{2(2m)!} \sum_{k=0}^{2m+1} (-1)^k \binom{2m+1}{k} \left(\frac{f}{W} + \frac{2m+1}{2} - k\right)^{2m} sgn\left(\frac{f}{W} + \frac{2m+1}{2} - k\right) \quad \text{(A-4)}$$

The only problem in this spectrum is that it does not consider the noise level. To overcome this problem the spectrum of the signal has been considered as a summation of two pulses as shown in Figure 3-3.

For the noise model considered herein a simpler approximation has been used for the spectrum of the nth order nonlinearity of noise. This approximation relies on the Fourier transform of the nth power of a sinc function and is obtained based on random walk [51].

Since the noise spectrum has been considered to be a pulse then its Fourier transform will be a sinc function and the convolution of a pulse m times will be the inverse Fourier transform of the m^{th} power of a sinc function.

$$S_{2w(m)}(\omega) = \int_{-\infty}^{+\infty} \left(\frac{sin(x)}{x}\right)^m e^{j\omega x}\, dx \quad \text{(A-5)}$$

This expression has been approximated using central limit theorem in [51] as below:

$$S_{2w(m)}(\omega) \sim \sqrt{\frac{3\pi}{m}}\, e^{\frac{-3\omega^2}{4m}} \quad \text{(A-6)}$$

For a pulse with amplitude $A_n$and width of $f_s$the m^{th} order spectrum will take the following form:

$$S_{2w(m)}(f) = {A_n}^m \left(\frac{f_s}{f_s}\right)^{m-1} \sqrt{\frac{3\pi}{m}} exp\left(\frac{-3\left(\frac{2\pi f}{f_s}\right)^2}{4m}\right) \tag{A-7}$$

if we call the overall signal plus noise as v, the following equation is valid and is known as binomial expansion:

$$v^m = (x+w)^m = \sum_{k=0}^{m} \binom{m}{k} x^k w^{m-k} \tag{A-8}$$

On the other hand based on the independency of signal and noise, the spectrum of v can be described as below:

$$S_{2v}(f) = S_{2x}(f) + S_{2w}(f) \tag{A-9}$$

Then from (3-13):

$$S_{2y}(f) = \sum_{m=0}^{K} \alpha_{2m+1} \underbrace{S_{2v}(f) * \cdots * S_{2v}(f)}_{m+1} * \underbrace{S_{2v}(-f) * \cdots * S_{2v}(-f)}_{m} \tag{A-10}$$

From the rectangular approximation it can be inferred that:

$$S_{2x}(f) = S_{2x}(-f) \tag{A-11}$$

$$S_{2w}(f) = S_{2w}(-f) \tag{A-12}$$

Hence:

$$S_{2v}(f) = S_{2v}(-f) \tag{A-13}$$

So (A-10) can be reformulated as below:

$$S_{2y}(f) = \sum_{m=0}^{K} \alpha_{2m+1} \underbrace{S_{2v}(f) * \cdots * S_{2v}(f)}_{2m+1} \tag{A-14}$$

From (A-9) and (A-14) it can be concluded that

$$S_{2y}(f) = \sum_{m=0}^{K} \alpha_{2m+1} \underbrace{\left(S_{2x}(f) + S_{2w}(f)\right) * \cdots * \left(S_{2x}(f) + S_{2w}(f)\right)}_{2m+1} \tag{A-15}$$

By using (A-8) (the binomial theorem), (A-4) and (A-7) the overall spectrum will take the following form

$$S_{2y}(f) = \sum_{m=0}^{K} \alpha_{2m+1} \sum_{i=0}^{2m+1} \binom{2m+1}{k} S_{2x(i)}(f) * S_{2w(2m+1-i)}(f) \tag{A-16}$$

APPENDIX B

For the MIMO system the error will be as below:

$$e_i = y_i - f_{i1}(x_1) + f_{i2}(x_2) \tag{B-1}$$

For each branch:

$$PDF(\boldsymbol{y_i}) = PDF\left(\boldsymbol{y_i}, \widehat{\boldsymbol{\theta}}_{\boldsymbol{i}}\right) PDF\left(\widehat{\boldsymbol{\theta}}_{\boldsymbol{i}}\right) (2\pi)^{\frac{p}{2}} \left|\boldsymbol{J_i}^{-1}\right|^{\frac{1}{2}} \tag{B-2}$$

which results in:

$$I_i = ln\left(PDF\left(\boldsymbol{y_i}, \widehat{\boldsymbol{\theta}}_{\boldsymbol{i}}\right)\right) + ln(PDF\left(\widehat{\boldsymbol{\theta}}_{\boldsymbol{i}}\right)) + \frac{p}{2} ln(2\pi) - \frac{1}{2} ln(det(\boldsymbol{J_i})) \tag{B-3}$$

By approximating this equation the cost function will be:

$$I_1 \approx ln\left(PDF\left(\boldsymbol{y_1}, \widehat{\boldsymbol{\theta}}_{\boldsymbol{1}}\right)\right) - \frac{1}{2} ln(det(\boldsymbol{J_1})) \tag{B-4}$$

To obtain the Fisher information matrices $\boldsymbol{J_1}$and $\boldsymbol{J_2}$ four cases should be considered. If we classify the coefficients to (11), (12) for the first branch and (21) and (22) for the second branch, then each Fisher information matrix will have 4 blocks. Among them there are two distinct cases: $\frac{\partial^2 \ln(\mathrm{PDF}(\mathbf{y},\boldsymbol{\theta}))}{\partial \mathrm{a}_{11\mathrm{im}}\, \partial \mathrm{a}_{11\mathrm{jq}}}$, $\frac{\partial^2 \ln(\mathrm{PDF}(\mathbf{y},\boldsymbol{\theta}))}{\partial \mathrm{a}_{11\mathrm{im}}\, \partial \mathrm{a}_{12\mathrm{jq}}}$. The first one will be the $\boldsymbol{J}_{11}$which is the Fisher information matrix of the SISO system with input $x_1(n)$. The second one will be approximately zero because:

$$\frac{\partial^2 ln(PDF(\boldsymbol{y},\boldsymbol{\theta}))}{\partial a_{11im}\partial a_{12jq}}$$

$$= \frac{1}{2\sigma^2} E\left\{\sum_{k=1}^{N} 2x_1(k)\,|x_1(k)|^{i-1} x_2(k)^H |x_2(k)|^{j-1}\right\} \quad \text{(B-5)}$$

$$\approx 0$$

The overall Fisher information matrix for each branch will be as below:

$$\boldsymbol{J}_1 = \begin{pmatrix} \boldsymbol{J}_{11} & 0 \\ 0 & \boldsymbol{J}_{22} \end{pmatrix} \quad \text{(B-6)}$$

The determinant of the overall fisher information matrix will be:

$$det(\boldsymbol{J}_1) = det(\boldsymbol{J}_{11})\, det(\boldsymbol{J}_{22}) \quad \text{(B-7)}$$

By applying the logarithmic function to (B-7), we can write:

$$ln(det(\boldsymbol{J}_1)) = ln(det(\boldsymbol{J}_{11})) + ln(det(\boldsymbol{J}_{22})) \quad \text{(B-8)}$$

Therefore, $\boldsymbol{J}_2$will be as below:

$$ln(det(\boldsymbol{J}_2)) = ln(det(\boldsymbol{J}_{11})) + ln(det(\boldsymbol{J}_{22})) \quad \text{(B-9)}$$

List of Publications

Journal and conference papers:

[P1] M.V. Amiri, S.A. Bassam, M. Helaoui and F.M. Ghannouchi, New order selection technique using information criteria applied to SISO and MIMO systems predistortion. International Journal of Microwave and Wireless Technologies, 5, pp 123-131, March 2013.

[P2] M.V. Amiri, S.A. Bassam, M. Helaoui, Fadhel M. Ghannouchi, Estimation of crossover DPD using orthogonal polynomials in fixed point arithmetic, AEU - International Journal of Electronics and Communications, Volume 67, Issue 11, Pages 905-909, ISSN 1434-8411 November 2013.

[P3] M.V. Amiri, S.A. Bassam, M. Helaoui, Fadhel M. Ghannouchi," Efficient Partition of Distortion Mitigation in Radio Link at the Transmitter and the Receiver ," Submitted to the IEEE Transactions on Microwave Theory , August 2013.

[P4] M.V. Amiri, M. Helaoui, Fadhel M. Ghannouchi, "Optimized Soft Clipping Crest Factor Reduction Technique Using Polynomial," Submitted to the IEEE Transactions on Microwave Theory and Techniques, January 2013.

[P5] M.V. Amiri, S.A. Bassam, M. Helaoui, Fadhel M. Ghannouchi, "Matrix-based orthogonal polynomials for MIMO transmitter linearization," Computer Aided Modeling, Analysis and Design of Communication Links and Networks (CAMAD), 2010 15th IEEE International Workshop on , vol., no., pp.57-60, 3-4 Dec. 2010.

[P6] M.V. Amiri, M. Helaoui, Fadhel M. Ghannouchi, "Streamlined MIMO Cross-Over Digital Predistortion," IEEE Radio and Wireless Symposium (RWS), 2013, accepted.

References

[1] J. Gozalvez, "Green Radio Technologies [Mobile Radio]," Vehicular Technology Magazine, IEEE , vol.5, no.1, pp.9,14, March 2010.

[2] B. Sklar, "Digital Communications Fundamentals and Applications," second ed. Prentice Hall, 2001, 382-383,409.

[3] Congzheng Han, Harrold T., Armour, S., Krikidis, I., Videv, S., Grant, Peter M., Haas, H., Thompson, J.S., Ku, I., Cheng-Xiang Wang, Tuan Anh Le, Nakhai M.R., Jiayi Zhang, Hanzo L., "Green radio: radio techniques to enable energy-efficient wireless networks," Communications Magazine, IEEE , vol.49, no.6, pp.46,54, June 2011.

[4] Bagheri Rahim, Mirzaei A., Heidari M.E., Chehrazi S., Minjae Lee, Mikhemar M., Tang W.K., Abidi A.A., "Software-defined radio receiver: dream to reality," Communications Magazine, IEEE , vol.44, no.8, pp.111,118, Aug. 2006.

[5] Hongwei Yang, "A road to future broadband wireless access: MIMO-OFDM-Based air interface," Communications Magazine, IEEE, vol.43, no.1, pp.53, 60, Jan. 2005.

[6] Astely D., Dahlman E., Fodor G., Parkvall S., Sachs J., "LTE release 12 and beyond [Accepted From Open Call]," Communications Magazine, IEEE , vol.51, no.7, pp., July 2013.

[7] Janevski T., "5G Mobile Phone Concept," Consumer Communications and Networking Conference, 2009. CCNC 2009. 6th IEEE, vol., no., pp.1,2, 10-13 Jan. 2009.

[8] Bolcskei H., "MIMO-OFDM wireless systems: basics, perspectives, and challenges," Wireless Communications, IEEE, vol.13, no.4, pp.31, 37, Aug. 2006.

[9] Ghannouchi F.M., Hammi O., "Behavioural modeling and predistortion," Microwave Magazine, IEEE , vol.10, no.7, pp.52,64, Dec. 2009.

[10] Hammi O., Younes M., Ghannouchi F.M., "Metrics and Methods for Benchmarking of RF Transmitter Behavioural Models With Application to the Development of a Hybrid Memory Polynomial Model," Broadcasting, IEEE Transactions on , vol.56, no.3, pp.350,357, Sept. 2010.

[11] Bassam S. A. (2011), "Advanced signal processing techniques for impairments compensation and linearization of SISO and MIMO transmitters," (Order No. NR75271, University of Calgary (Canada)).

[12] Steve C. Cripps, "RF Power Amplifiers for Wireless Communications," Second Edition (Artech House Microwave Library (Hardcover)), Artech House, Inc., Norwood, MA, 2006.

[13] Bumman Kim, Jangheon Kim, Ildu Kim, Jeonghyeon Cha, "The Doherty power amplifier," Microwave Magazine, IEEE , vol.7, no.5, pp.42-50, Oct. 2006

[14] Joel Vuolevi, Timo Rahkonen, "Distortion in RF power amplifiers," Artech House microwave library 2003.

[15] S. Maas, "Nonlinear Microwave and RF Circuits," Artech House, Norwood, MA, 2003.

[16] Tehrani A.S., Haiying Cao, Afsardoost S., Eriksson T., Isaksson M., Fager C., "A Comparative Analysis of the Complexity/Accuracy Tradeoff in Power Amplifier

Behavioural Models," Microwave Theory and Techniques, IEEE Transactions on , vol.58, no.6, pp.1510-1520, June 2010.

[17] Isaksson M., Wisell D., Ronnow D., "A comparative analysis of behavioural models for RF power amplifiers," Microwave Theory and Techniques, IEEE Transactions on , vol.54, no.1,pp.348-359, Jan. 2006

[18] Bae H.G., Helaoui M., Seregin A., Boumaiza S., Ghannouchi F.M., "Blind Peak-to-Average Power Ratio Reduction Technique for WiMAX RF Front-end," 36th European Microwave Conference, 2006, pp. 149–152, 10-15 Sept. 2006.

[19] Wulich D., "Definition of efficient PAPR in OFDM," Communications Letters, IEEE , vol.9, no.9, pp.832,834, Sep 2005.

[20] S. A. Bassam, M. Helaoui, and F. M. Ghannouchi, "Crossover Digital Predistorter for the Compensation of Crosstalk and Nonlinearity in MIMO Transmitters", IEEE Trans. On Microwave Theory and Techniques, vol. 57, issue: 5, part: 1, May 2009, pp. 1119-1128.

[21] Bae H.G., Helaoui M., Seregin A., Boumaiza S., Ghannouchi F.M., "Blind Peak-to-Average Power Ratio Reduction Technique for WiMAX RF Front-end," 36th European Microwave Conference, 2006, pp. 149–152, 10-15 Sept. 2006.

[22] Zeleny J., Rosson P., Dehos C., Kaiser A., "Digital compensation of the power amplifier nonlinearities at relay station receivers in 802.16j very high data rate systems," 2010 IEEE Radio and Wireless Symposium (RWS), pp. 244–247, 10-14 Jan. 2010

[23] Gang Liu, Haldi P., Tsu-Jae King Liu, Niknejad A.M., "Fully Integrated CMOS Power Amplifier With Efficiency Enhancement at Power Back-Off," IEEE Journal of Solid-State Circuits, vol. 43, no. 3, pp. 600-609, March 2008

[24] Muruganathan S.D., Sesay A.B., "A QRD-RLS-Based Predistortion Scheme for High-Power Amplifier Linearization," IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 53, no. 10, pp. 1108–1112, Oct. 2006.

[25] Lei Ding, Zhengxiang Ma, Morgan D.R., Zierdt M., Tong Zhou G., "Compensation of Frequency-Dependent Gain/Phase Imbalance in Predistortion Linearization Systems," IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 55, no. 1, pp. 390–397, Feb. 2008.

[26] Koeppl H., Singerl P., "An Efficient Scheme for Nonlinear Modeling and Predistortion in Mixed-Signal Systems," IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 53, no. 12, pp. 1368–1372, Dec. 2006.

[27] Xinping Huang, Caron M., "Performance of a type-based digital predistorter for solid-state power amplifier linearization," 2005 IEEE International Symposium on Circuits and Systems (ISCAS 2005), vol. 2, pp. 1710–1713, 23-26 May 2005.

[28] Lou Jingyi, Gao Jun, Qu Xiaoxu, Chen Lin, "Improved CDF predistortion method and experimental results for HF power amplifier," 2010 International Conference on Wireless Communications and Signal Processing (WCSP), pp. 1–4, 21-23 Oct. 2010

[29] Congying Xia, Jacek Ilow F., "Blind Compensation of Memoryless Nonlinear Effects in OFDM Transmissions Using CDF", Communication Networks and Services Research Conference, 2003.

[30] Dongliang Huang, Xinping Huang, Henry Leung , "Nonlinear Compensation of High Power Amplifier Distortion for Communication Using a Histogram-Based Method," Signal Processing, IEEE Transactions on , vol.54, no.11, pp.4343-4351, Nov. 2006

[31] Franco M.J., "Mobile handset power amplifiers," Microwave Magazine, IEEE , vol.10, no.7, pp.16-19, Dec. 2009

[32] Linear Technology, "Digital Predistortion uModule Receiver Simplifies Base station Design," MILPITAS, CA, July 2009. [Online]. Available: http://cds.linear.com/docs/en/press-release/LTM9003.pdf

[33] Hikuma A., Fuke Y., Nakaminami N., Ohyane H. and Kobayashi H., "Radio Base Stations Equipments toward Economical Expansion of FOMA Coverage Areas, " NTT DoCoMo Technical Journal Vol. 6 No.1

[Online].Available:http://www.nttdocomo.co.jp/english/binary/pdf/corporate/technology/rd/technical_journal/bn/vol6_1/vol6_1_052en.pdf

[34] Maruyama S., Tanahashi K., Higuchi T. , "Base Transceiver Station for W-CDMA System," FUJITSU Sci. Tech. J., 38,2,pp.167-173, Dec. 2002.[Online].Available:http://www.fujitsu.com/downloads/MAG/vol38-2/paper07.pdf

[35] Y. Maghsoodi, "Exact distributions of envelopes of sums of stochastic sinusoids with general random amplitudes and phases, " Scinance Analytics, Nov. 2004.

[36] Hammi, O., Carichner, S., Vassilakis, B., Ghannouchi, F.M., "Synergetic Crest Factor Reduction and Baseband Digital Predistortion for Adaptive 3G Doherty Power Amplifier Linearizer Design," Microwave Theory and Techniques, IEEE Transactions on , vol.56, no.11, pp.2602-2608, Nov. 2008

[37] Papoulis A., "Probability, Random Variables and Stochastic Processes," Third edition, Mc Graw-Hill, 1991.

[38] Kaplan EL, Meier P., "Nonparametric estimation from incomplete observations," J. Am. Stat. Assoc. 1958;53:457-481.

[39] Krishnamoorthy K., "Handbook of Statistical Distributions with Applications," Chapman & Hall/CRC, 2006

[40] Wan-Jong Kim, Stapleton S.P., Jong Heon Kim, Edelman C., "Digital predistortion linearizes wireless power amplifiers," Microwave Magazine, IEEE , vol.6, no.3, pp. 54- 61, Sept. 2005

[41] Altera White Papers, " Implementing Digital IF & Digital Predistortion Linearizer Functions with Programmable Logic," May 2003, ver. 1.0.

[42] Ilic J., Strohmer T., "Sparsity Enhanced Decision Feedback Equalization," IEEE Transactions on Signal Processing, vol.60, no.5, pp.2422,2432, May 2012.

[43] Tao Jiang, Yiyan Wu, "An Overview: Peak-to-Average Power Ratio Reduction Techniques for OFDM Signals," Broadcasting, IEEE Transactions on , vol.54, no.2, pp.257-268, June 2008

[44] Seung Hee Han, Jae Hong Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," Wireless Communications, IEEE , vol.12, no.2, pp. 56- 65, April 2005.

[45] Dukhyun Kim, Stuber, G.L., "Clipping noise mitigation for OFDM by decision-aided reconstruction," Communications Letters, IEEE , vol.3, no.1, pp.4-6, Jan. 1999.

[46] Heung-Gyoon Ryu, Byoung-Ii Jin, In-Bae Kim , "PAPR reduction using soft clipping and ACI rejection in OFDM system ," Consumer Electronics, IEEE Transactions on , vol.48, no.1, pp.17-22, Feb 2002.

[47] Fujiwara, T., Tomisato S., Hata M., Fujii H., Asai T., Okumura Y., "An adaptive soft clipping method for spectrum sharing OFDMA systems," Personal, Indoor and Mobile Radio Communications, 2009 IEEE 20th International Symposium on , vol., no., pp.305-309, 13-16 Sept. 2009.

[48] Nader C., Landin P.N., Van Moer W., Bjorsell N., Handel P., Ronnow D., "Peak-Power Controlling Technique for Enhancing Digital Pre-Distortion of RF Power Amplifiers," Microwave Theory and Techniques, IEEE Transactions on , vol.60, no.11, pp.3571,3581, Nov. 2012.

[49] G. Tong Zhou and Raviv Raich., "Spectral analysis of polynomial nonlinearity with applications to RF power amplifiers, " EURASIP J. Appl. Signal Process., 1831-1840 January 2004.

[50] Jonhson N.L., Kotz S., Balakrishnan, "Continuous Univariate Distributions, " Volume 2, 2nd Edition, Wiley, 1995.

[51] R. E. Crandall, "Note on sinc-kernel sums and Poisson transformation, " (preprint, 15 June 2007). Available at http://www.reed.edu/~crandall.

[52] 3GPP TS 36.104 V9.13.0 (2012-09).

[53] Younes M., Hammi O., Kwan A., " Ghannouchi, F.M., "An Accurate Complexity-Reduced “PLUME” Model for Behavioural Modeling and Digital Predistortion of RF Power Amplifiers," Industrial Electronics, IEEE Transactions on , vol.58, no.4, pp.1397-1405, April 2011

[54] Wood J., LeFevre M., Runton D., Nanan J.-C., Noori B.H., Aaen P.H., "Envelope-domain time series (ET) behavioural model of a Doherty RF power amplifier for system design," Microwave Theory and Techniques, IEEE Transactions on , vol.54, no.8, pp.3163-3172, Aug. 2006

[55] Akaike, H., "A new look at the statistical model identification," Automatic Control, IEEE Transactions on , vol.19, no.6, pp. 716- 723, Dec 1974

[56] Schwarz G., "Estimating the Dimension of a Model," The Annals of Statistics,Vol. 6, No. 2 pp. 461-464 Mar.1978

[57] Stoica, P., Selen, Y., "Model-order selection: a review of information criterion rules," Signal Processing Magazine, IEEE , vol.21, no.4, pp. 36- 47, July 2004

[58] Bozdogan H., "Model Selection and Akaike's Information Criterion (AIC): The General Theory and Its Analytical Extensions, " Psychometrika, 52(3), 345-370, 1987.

[59] Liavas A.P., Regalia P.A., "On the behaviour of information theoretic criteria for model order selection," Signal Processing, IEEE Transactions on , vol.49, no.8, pp.1689-1695, Aug 2001

[60] Amiri M.V., Bassam S.A., Helaoui M., Ghannouchi F.M., "Matrix-based orthogonal polynomials for MIMO transmitter linearization," Computer Aided Modeling, Analysis and Design of Communication Links and Networks (CAMAD), 2010 15th IEEE International Workshop on , vol., no., pp.57-60, 3-4 Dec. 2010

[61] Arora J. S.,Huang M. W.,Hsieh C. C., "Methods for optimization of nonlinear problems with discrete variables: A review", Structural and Multidisciplinary Optimization, Volume 8, Numbers 2-3 (1994), 69-85

[62] Rutenbar, R.A., "Simulated annealing algorithms: an overview," Circuits and Devices Magazine, IEEE , vol.5, no.1, pp.19-26, Jan 1989

[63] J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bolcskei, "An overview of MIMO communications—A key to gigabit wireless," Proceeding of IEEE, vol. 92, no. 2, pp. 198–218, Feb. 2004.

[64] Y. Palaskas, A. Ravi, S. Pellerano, B. R. Carlton, M. A. Elmala, R.Bishop, G. Banerjee, R. B. Nicholls, S. K. Ling, N. Dinur, S. S. Taylor, and K. Soumyanath, "A 5-GHz 108-Mb/s 2x2 MIMO transceiver RFIC with fully integrated 20.5-dBm power amplifiers in 90-nmCMOS," IEEE Journal in Solid-State Circuits, vol. 41, no. 12, pp. 2746–2756, Dec. 2006.

[65] W.C. Hua, P.T. Lin, C.P. Lin, C.Y. Lin, H.L. Chang, C. W. Liu, T.Y. Yang, and G.K. Ma, "Coupling effects of dual SiGe power amplifiers for 802.11n MIMO applications, " in IEEE Radio Freq. Integr. Circuits Symposium, Jun. 11–13, 2006.

[66] Sulyman and M. Ibnkahla, "Performance of MIMO Systems with Antenna Selection over Nonlinear Fading channels," IEEE Journal on Selected Topics in Signal Processing, vol. 2, pp. 1–12, Apr. 2008.

[67] S. A. Bassam, M. Helaoui, F. M. Ghannouchi, "BER Performance Assessment of Linearized MIMO Transmitters in Presence of RF Crosstalk," IEEE Radio and Wireless Symposium (RWS'2010), pp. 33–36, New Orleans, LA, USA, Jan. 10–14. 2010.

[68] P. Midya, "Polynomial predistortion linearizing device, method, phone and base station, " U.S. Patent 6 236 837, July 30, 1998.

[69] P. Midya and J. Grosspietsch, "Scalar cost function based predistortion linearizing device, method, phone and base station, " U.S. Patent 6 240 278, July 30, 1998.

[70] R. Raich, H. Qian, and G. T. Zhou, "Orthogonal polynomials for power amplifier modeling and predistorter design," IEEE Transactions on Vehicular Technology., vol. 53, no. 5, pp. 1468–1479, Sep. 2004.

[71] R. Raich, H. Qian, and G. T. Zhou, "Digital baseband predistortion of nonlinear power amplifiers using orthogonal polynomials, " the 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '03), vol. 6, pp. VI-689-92, Apr. 6–10, 2003.

[72] Inacio and D.Ombres, "The DSP decision: fixed point or floating?, " IEEE Spectrum, vol. 33, no. 9, pp. 72–74, Sep. 1996.

[73] G. Golub and C. Van Loan, "Matrix Computations," 2nd ed. Johns Hopkins University Press, Baltimore, MD, 1989.

[74] Rajopadhye S.V., "Systolic arrays for LU decomposition," Circuits and Systems, 1988, IEEE International Symposium on, vol., no., pp.2513-2516 vol.3, 7-9 Jun 1988.

[75] B. Louka, M. Tchuenté, "Triangular matrix inversion on systolic arrays," Parallel Comput. 14 (2), 223_228, 1990.