In compliance with the
Canadian Privacy Legislation
some supporting forms
may have been removed from
this dissertation.

While these forms may be included
in the document page count,
their removal does not represent
any loss of content from the dissertation.

A Multiple-Polynomial Technique for HPA Linearization through RF Predistortion

by

Ahmad Zahran, B.Sc.

A thesis submitted to the
Faculty of Graduate Studies and Research
in partial fulfilment of the requirements for the degree of
Master of Applied Science in Electrical Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering
Department of Systems and Computer Engineering
Carleton University

May 2003

© Copyright 2003, Ahmad Zahran

National Library
of Canada

Bibliothèque nationale
du Canada

Acquisitions and
Bibliographic Services

Acquisitiions et
services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 0-612-88905-X

Our file Notre référence

ISBN: 0-612-88905-X

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

Canada

ABSTRACT

The increasing popularity of digital and wireless communication systems creates a need for efficient, distortion-free transmitters. Many of the transmission links in use today employ high power amplifiers (HPAs), which are inherently nonlinear. However, signal predistortion techniques provide an effective means to reduce the levels of signal distortion that could be caused by the HPA nonlinearities.

Throughout this thesis a new predistortion technique is studied. This technique divides the characteristics of an ideal predistorter into various sections in order to model each section with a low order polynomial. The predistorter obtained from this process is called a multipolynomial predistorter.

A detailed study of this predistorter is performed and numerous simulations are carried out to investigate the predistorter's performance on two types of HPAs. Also, a comparison is made with other predistorters proposed in literature to identify the predistorters' abilities to linearize a HPA.

ACKNOWLEDGEMENTS

I thank God first and foremost.

My warmest gratitude goes to my supervisor, Professor Mohamed El-Tanany. His guidance and wisdom gave me confidence to explore new grounds, taught me more than was required and extended my capabilities to new dimensions.

I am indebted to my family, especially to my parents, whose constant support and encouragement makes me want to strive further every day. My successes are because of you and for you.

TABLE OF CONTENTS

ABSTRACT	III
ACKNOWLEDGEMENTS	IV
TABLE OF CONTENTS	V
LIST OF FIGURES	VIII
LIST OF ACRONYMS	XI
LIST OF SYMBOLS	XIII
1 INTRODUCTION	1
1.1 Thesis Objectives	2
1.2 Thesis Contributions	2
1.3 Thesis Organization	4
2 BACKGROUND INFORMATION	6
2.1 The Nonlinearity Problem	6
2.2 Amplifier Characterization Techniques	9
2.2.1 Using Network Analyzers	10
2.2.2 The Two-Tone Method	11
2.2.3 The Unequal Three-Tone Method	15
2.2.4 The Unequal Two-Tone Method	18
2.2.5 Using Digital Modulation	21
2.3 Modeling Procedures	22
2.3.1 Saleh's Model	23
2.3.2 Ghorbani and Sheikhan's Model	25
2.3.3 The Polynomial Model	26

Table of Contents

2.4	Linearization Schemes	29
2.4.1	Feedback Linearization	29
2.4.2	Feedforward Linearization	31
2.4.3	Predistortion Linearization	33
2.5	Literature Review	36
3	SIMULATION MODEL	40
3.1	Transmission System Model	40
3.2	HPA Models	42
3.2.1	SSPA	43
3.2.2	TWTA	46
3.3	Overview of the Predistortion Technique	50
3.3.1	Ideal Predistorter Characteristics	50
3.3.2	Multiple-Polynomial Models	56
3.4	Performance Analysis	62
3.5	Model Verification	68
4	SIMULATION RESULTS	73
4.1	Effect of Sectioning and Polynomial Order	73
4.1.1	Effect of the Number of Sections	74
4.1.2	Effect of the Polynomial Order	79
4.1.3	Study of the PSD	84
4.2	Performance of the SSPA	88
4.3	Comparison Based on the TWTA Model	95
4.3.1	Performance Assessment Based on MSE Computations	97
4.3.2	Performance Assessment Based on BER Computations	101
4.3.3	Performance Assessment Based on the PSD	108
4.4	Effect of the Sampling Ratio	111
5	CONCLUSIONS AND RECOMMENDATIONS	117
5.1	Thesis Conclusions	117
5.2	Recommendations for Future Work	121

Table of Contents

APPENDIX A	NONLINEAR SSPA	123
REFERENCES		125

LIST OF FIGURES

Figure 2.1:	Input-Output Power Relationship of a Typical HPA	7
Figure 2.2:	Constellation Distortion due to Nonlinear Amplification	9
Figure 2.3:	Typical Spectral Output of the Two-Tone Test	15
Figure 2.4:	Input and Output Spectrum of the Unequal Three-Tone Method	17
Figure 2.5:	Input and Output Spectrum of the Unequal Two-Tone Method	20
Figure 2.6:	Block Diagram of the Feedback Linearization Method	30
Figure 2.7:	Block Diagram of the Feedforward Linearization Method	32
Figure 2.8:	Block Diagram of the Predistortion Linearization Method	34
Figure 3.1:	Block Diagram of the Transmission System Model	41
Figure 3.2:	Block Diagram of the Characterization Process of the ZHL-42 HPA	44
Figure 3.3:	AM/AM & AM/PM Characteristics of the SSPA at Various Frequencies	45
Figure 3.4a:	AM/AM & AM/PM Characteristics of the TWTA in Linear Scale	48
Figure 3.4b:	AM/AM & AM/PM Characteristics of the TWTA in Logarithmic Scale	49
Figure 3.5:	Three Approaches for HPA Linearization	52
Figure 3.6:	Ideal Predistorter Characteristics for the SSPA	54
Figure 3.7:	Ideal Predistorter Characteristics for the TWTA	55
Figure 3.8:	Amplitude Predistorter Characteristics divided into m Sections	57
Figure 3.9:	Implementation Model of the Predistorter	61

List of Figures

Figure 3.10:	Functional Block Diagram of the Predistorter	61
Figure 3.11:	Diagrams of the Assumed Transmission Systems	63
Figure 3.12:	MSE Variance at various amounts of Transmitted Symbols ...	67
Figure 3.13:	Amplification Process of the DSB-SC signal	71
Figure 4.1:	TD versus OBO for the MP-PD with $n=3$ at various values of m	75
Figure 4.2:	TD versus OBO for the MP-PD with $m=2, n=3$ using equal and unequal-length sections	78
Figure 4.3:	TD versus OBO for the MP-PD using various polynomial orders applying a 256-QAM signal	80
Figure 4.4:	TD versus OBO for the MP-PD with $n=2$ at various values of m	82
Figure 4.5:	PSD resulting from various MP-PDs on the TWTA using a 64-QAM signal with	86
Figure 4.6:	PSD resulting from various MP-PDs on the TWTA using a 256-QAM signal with	87
Figure 4.7:	Performance of various MP-PDs on the SSPA using 64 and 256-QAM signals with $\beta=0.5$	90
Figure 4.8:	Performance of various MP-PDs on the SSPA using 64 and 256-QAM signals with $\beta=0.25$	91
Figure 4.9:	Performance of various MP-PDs on the SSPA using 64 and 256-QAM signals with $\beta=0.125$	92
Figure 4.10:	PSD resulting from various MP-PDs on the SSPA	94
Figure 4.11:	Performance of various MP-PDs on the TWTA using 64 and 256-QAM signals with $\beta=0.5$	98
Figure 4.12:	Performance of various MP-PDs on the TWTA using 64 and 256-QAM signals with $\beta=0.25$	99
Figure 4.13:	Performance of various MP-PDs on the TWTA using 64 and 256-QAM signals with $\beta=0.125$	100
Figure 4.14:	Block Diagram of the Transmission System for the Semi-Analytic Approach	102
Figure 4.15:	BER versus E_b / N_o for a 64-QAM Signal using the MP-PD-[3,2] at various OBOs	103
Figure 4.16:	Histograms of the Error Signal	105
Figure 4.17:	Semi-Analytic Analysis of various MP-PDs on the TWTA using 64 and 256-QAM signals with $\beta=0.25$	107

List of Figures

Figure 4.18: PSD resulting from various MP-PDs on the TWTA with $\beta=0.25$	109
Figure 4.19: Block Diagram of the Sampling Process	112
Figure 4.20: TD Results from various Sampling Rates	113
Figure 4.21: Minimum TD for various Sampling Rates	114
Figure 4.22: TD Results for various Sampling Rates using the Semi-Analytic Approach	115
Figure 4.23: Minimum TD for various Sampling Rates using the Semi-Analytic Approach	116
Figure A.1: Performance of the Non-Linear TWTA using 64-QAM and 256-QAM Signals	124

LIST OF ACRONYMS

A&P-PD	Amplitude and Phase Predistorter
AWGN	Additive White Gaussian Noise
AM	Amplitude Modulation
AM/AM	Amplitude Modulation-to-Amplitude Modulation Conversion
AM/PM	Amplitude Modulation-to-Phase Modulation Conversion
BEM	Bandwidth-Efficient Modulation
BER	Bit Error Rate
BPSK	Binary Phase Shift Keying
CAD	Computer Aided Design
CDMA	Code-Division Multiple Access
dB	Decibel
DSB-SC	Double Sideband Suppressed Carrier
DSP	Digital Signal Processing
FDMA	Frequency-Division Multiple Access
HF	High Frequency
HPA	High Power Amplifier
IF	Intermediate Frequency
IM	Intermodulation
IM3	3 rd Order Intermodulation Product
IM5	5 th Order Intermodulation Product
IM7	7 th Order Intermodulation Product
ISI	Intersymbol Interference
K&S-PD	Karam and Sari Predistorter
LS	Least Squares
MDP-PM	Minimum Distortion Power Polynomial Model

List of Acronyms

MP-PD	Multipolynomial Predistorter
MSE	Mean Square Error
OBO	Output Backoff
PD	Predistorter
PM	Phase Modulation
QAM	Quadrature Amplitude Modulation
RF	Radio Frequency
SNR	Signal to Noise Ratio
SNDR	Signal to Noise plus Distortion Ratio
SRRC	Square Root Raised Cosine
SSPA	Solid State Power Amplifier
TD	Total Degradation
TDMA	Time-Division Multiple Access
TWTA	Travelling Wave Tube Amplifier
VHF	Very High Frequency

LIST OF SYMBOLS

A	signal amplitude
A_o	average amplitude of an AM signal
$A(\bullet)$	AM/AM characteristics of the predistorter
a	polynomial coefficients of the amplitude predistorter
a_i	in-phase component of the i th transmitted symbol
$\hat{a}_i$	in-phase component of the i th received symbol
α	linear voltage gain
α_o	amplitude index of an AM signal
b	polynomial coefficients of the phase predistorter
b_i	quadrature component of the i th transmitted symbol
$\hat{b}_i$	quadrature component of the i th received symbol
β	filter rolloff factor
β_o	phase deviation constant of a PM signal
c	AM/AM compression factor
c_i	complex component of the i th transmitted symbol
$\hat{c}_i$	complex component of the i th received symbol
$\hat{c}(t)$	filtered version of the received signal $\hat{x}(t)$
Δ	degradation
$E(\bullet)$	expectation operation
E_b	energy per bit
e	error signal
F_i	frequency of the in-phase components
F_q	frequency of the quadrature components

List of Symbols

f	frequency (Hz)
G	gain
G_R	receiver filter
G_T	transmitter filter
I	in-phase components of the error signal
K	optimal sampling rate
k	number of symbols transmitted
k_p	AM/PM conversion coefficient
ℓ	spacing factor of QAM signal components
M_{QAM}	constellation size of the transmitted QAM signal
$M(\bullet)$	AM/AM characteristics of the HPA
m	number of sections in multipolynomial predistorter
N_Q	QAM constellation length
N	average energy of the noise
N_o	noise PSD
n	polynomial order
$n(t)$	AWGN signal
ω	angular frequency (radians)
ω_c	carrier frequency (radians)
P_{in}	input power
P_{out}	output power
P_{sat}	input saturation point
$\Phi(\bullet)$	AM/PM characteristics of the HPA
$\Psi(\bullet)$	AM/PM characteristics of the predistorter
ϕ	amount of phase insertion
Q	quadrature components of the error signal
R	upsampling/downsampling rate
ρ	instantaneous amplitude
ρ_i	maximum input voltage of the i -th section
ρ_{max}	maximum possible input voltage that the predistorter will linearize
S	average energy of the signal

List of Symbols

T	time spacing between transmitted symbols
$TD_{\min}$	minimum measured of total degradation
θ	instantaneous phase
θ_o	initial phase value of a PM signal
v	signal amplitude
$v_i(t), v_{in}(t)$	input waveform
$v_o(t)$	output waveform
$x(t)$	transmitted signal
$\hat{x}(t)$	attenuated received signal
$Y_{\max}$	maximum amplitude of the predistorted signal
$y(t)$	predistorted signal
$z(t)$	amplified signal

Chapter 1

INTRODUCTION

In today's society there is a real need to transport all types of information over short or long distances at high speeds, all at the touch of a button. This creates a necessity for highly efficient communication systems. More so than in the past, with the revolution and growing demand for wireless communications, the need for bandwidth-efficient modulation (BEM) is growing on a daily basis. Whether it is by using frequency-division multiple access (FDMA), code-division multiple access (CDMA) or time-division multiple access (TDMA) schemes, all communication systems need to create a highly linear, bandwidth-efficient transmission signal, all while observing efficient power consumption. To do so, one requires a highly linear, highly efficient high power amplifier (HPA). However, HPAs are generally known to be considerably nonlinear elements. If they are operated at high output power levels they tend to heavily distort the transmitted signal and increase the out-of-band spectral emissions.

1.1 Thesis Objectives

HPAs need to be linearized in order to cause them to perform in a linear yet power-efficient manner. There are various linearization schemes that can be used to achieve this, each with advantages and disadvantages. For signals with complex envelopes or wide bandwidths, predistortion is one of the most commonly used methods.

The aim of this thesis is to analyze several of the predistortion schemes available in literature, and through an in-depth study of them, be able to evaluate their linearizing abilities. To do this, the most common and some of the most linear predistorters (PDs) are investigated. However, the main focus is on a new predistortion approach. This new predistortion technique is thoroughly tested and compared to other predistorters. The ability of the various predistorters to linearize a HPA is measured and compared on a quantitative basis.

1.2 Thesis Contributions

The main purpose of this work is to evaluate the performance of a new predistortion technique. This predistortion technique arises as an innovation of already existing predistorters. It exploits the idea of modeling the ideal characteristics of a predistorter by using polynomials and build on it by creating a model which employs more than one polynomial. This predistortion technique

is named the multiple-polynomial predistortion technique, or multipolynomial predistorter (MP-PD) for short.

In line with the objectives of this study, the multipolynomial technique is exhaustively tested to analyze several of its aspects:

- The effect of changing the number of sections that compose the multipolynomial predistorter is studied and it is found that an increase in the number of sections creates an improvement on the linearization ability of the predistorter.
- The effect of changing the order of the polynomials that make up the predistorter is also examined. As expected, a larger polynomial order produces a more effective linearizer. However, it is also observed that the polynomial order can be decreased to as low as a 2nd order, given there are enough sections in the predistorter. Hence an increase in the number of sections can allow the polynomial order to decrease without affecting the predistorter's performance. This creates the possibility of implementing lower order polynomials, which in turn creates a simpler system.
- The predistorter's ability to linearize HPAs is tested on both a solid state power amplifier (SSPA) and a travelling wave tube amplifier (TWTA). The amount of degradation caused by both HPAs is measured at various output backoff power levels. It is found that a SSPA could be linearized with a lower order polynomial or fewer sections than a TWTA. This is due to the nature of SSPAs, which generally tend to be less nonlinear than TWTAs.

- The multipolynomial predistorter's performance is then compared to various predistorters found in literature, namely in [7], [9] and [14]. By measuring the degradation of these predistorters at various output backoff power levels, a quantitative analysis of the predistorters' performance is obtained. It is observed that the multipolynomial predistorter approaches the performance of an ideal predistorter as the number of sections and the polynomial order are increased.
- Finally, the effect of the sampling rate on the predistorter's performance is studied. If the predistorter is to be implemented on a DSP system, this would be an important factor to consider. Results show that a sampling rate of 2.5 samples per symbol or higher will guarantee a sufficiently accurate performance.

It is found that the multipolynomial predistorter creates a highly flexible and accurate system, able to adjust to complex as well as simple HPA characteristics while maintaining a high degree of accuracy and a low level of implementation complexity.

1.3 Thesis Organisation

This thesis is divided into five main chapters. Chapter 2 contains the background information necessary to understand the linearity problem in HPAs. It also reviews some of the most common methods available in literature that deal with the characterization process as well as the modelling and linearization

schemes used on a HPA. An up-to-date literature review is then given on the various predistortion techniques available that have steered the way to the development of the multipolynomial predistorter.

Chapter 3 starts with a review of the simulation model which is used to study the performance of a predistorter. Also in this chapter, the characteristics of the various HPAs used in the simulation are described, together with the characteristics of the ideal predistorters necessary to linearize them. Subsequently, the multipolynomial predistortion technique is described in detail. After the method used to measure the performance of the predistorter is explained, a test is run to verify that the simulation is functioning correctly.

The results obtained from numerous simulations are presented in chapter 4. These results aim at giving an insight into the multipolynomial predistorter's performance through a comparison to some of the predistorters in literature.

The conclusions drawn from this thesis are detailed in chapter 5 along with recommendations for future work.

BACKGROUND INFORMATION

This chapter explains the basic linearity problem of HPAs and the necessity to linearize them. It also provides an analysis of the three basic steps required to obtain a linear system: characterization, modeling and linearization. Each one of these steps can be performed through various methods and a literature review of the most commonly used ones is presented here.

2.1 The Nonlinearity Problem

With the greater demand for satellite communications, broadband wireless transmissions and wireless terrestrial services such as digital radio, wireless internet services and cellular telephony to name a few, the need for optimal spectrum utilization is vital. Efficient transmission of information is necessary to ensure an effective communication system. Hence, transmission equipment needs to be highly linear in order to avoid distortions and spectral inefficiency.

HPAs are required when increasing the power of a signal to a level high enough for transmission. However, they tend to lose linearity at high input power levels [20]. Figure 2.1 illustrates the input-output power relationship of a typical HPA. It can be clearly seen that as the input power P_{in} increases, the loss in linearity ΔP_{out} also increases. As this happens, a point will be reached at which an increase in the input power will no longer create an increase in the output power. At such a point, the maximum possible output power is attained. This is termed the saturation point P_{sat} and its knowledge is fundamental in order to linearize the HPA.

Figure 2.1: Input-Output Power Relationship of a Typical HPA

As the loss of linearity increases, another phenomenon takes place: a phase shift is inserted into the amplified signal. The amount of phase which is inserted varies according to the input power level. It is customary to refer to this phenomenon as “phase distortion” or “phase insertion” [11], [18]. The latter term will be used throughout the rest of this thesis.

Loss in linearity and phase insertions create two types of distortion in the amplified signal. Primarily, undesired energy at new frequencies is created in the spectrum of the signal. This produces a larger bandwidth, which results in a less spectrally-efficient signal [23]. The other type of distortion takes place in the in-phase and quadrature components of the signal constellation. Amplitude and phase changes in the signal force the constellation points to change position and spread into clusters. Also, large phase insertions will cause the constellation to rotate. Such changes unavoidably generate a higher bit error rate (BER) in the received signal. These effects can be better visualized by looking at figure 2.2.

In order to avoid any distortion, the HPA can be operated at a state of high back-off from its saturation point. This results in a more linear amplification but yields a less efficient use of the HPA [15]. To employ the HPA in a more efficient manner, various methods can be exploited to linearize the input-output power relationship and eliminate phase insertions. These methods generally provide a better HPA efficiency along with lower levels of signal distortion. In most cases however, in order to linearize a HPA one must first characterize and model the device. There are various methods to perform these tasks. The most common ones are reviewed in the following sections.

Figure 2.2: Constellation Distortion due to Nonlinear Amplification

2.2 Amplifier Characterization Techniques

The term “amplifier characterization” is used to describe any process which is used to extract the AM/AM and AM/PM characteristics of a HPA [18]. Such a process must include all the testing and calculations required to find the input-output power relationship and phase insertion created by the HPA over a specific range of input power. The application of a method or equation that

would model these characteristics is outside the scope of the characterization process. However, some characterization techniques provide a method which as well as characterizing also create such a model.

The following subsections present a discussion of some of the most useful and relevant techniques.

2.2.1 Using Network Analyzers

Up to date, network analyzers have been the most trusted and widely used systems for characterizing HPAs. These multipurpose pieces of equipment have been used to extract the AM/AM and AM/PM characteristics of amplifiers for many years. Their operation is simple and requires no need for calculations, but entails extensive calibration. These devices are also quite expensive and very delicate.

Network analyzers have the ability to perform various functions, one of which is a continuous wave power sweep [1]. After the HPA is connected to the network analyzer and the system is calibrated appropriately, a power sweep is performed. This power sweep is carried out at a single frequency in order to measure the gain of the HPA over a range of input power levels. The results are presented on the screen of the network analyzer in the form of a graph displaying gain versus input power. From knowledge of the gain at any input power level, an input-output power relationship can be obtained. Such a

relationship represents the AM/AM characteristics of the HPA at the frequency of interest.

Through the same power sweep function, network analyzers can also measure a device's phase over the same set of input power levels. This is used on a HPA to find its AM/PM characteristics. By prompting the system to perform a single frequency sweep along with the power sweep, the network analyzer will be able to display a graph of measured phase versus input power. By eliminating the phase shift created by the amplifier from the measurements, one is left with the AM/PM characteristics of the HPA.

By using network analyzers, the need for complex setups and time-consuming calculations vanishes. This equipment requires only that the HPA is connected to it and that it is well calibrated. In a matter of seconds the AM/AM and AM/PM characteristics of the HPA can be found. However, because these devices are expensive, they are not the type of equipment that can be commonly found in all laboratories. For this reason, alternative methods need to be explored.

2.2.2 The Two-Tone Method

This is a less costly, more complex scheme for characterizing HPA nonlinearities. It tries to identify the behaviour of a HPA by analyzing the power

spectrum produced at its output when its input is a signal composed of two cosine waves of equal amplitude:

$$v_i(t) = v \cos(\omega_1 t) + v \cos(\omega_2 t) \quad (2.1)$$

These two tones are designed in a way that the frequency spacing between them is much smaller than the RF frequency of either tone.

The two-tone method assumes that the input-output relationship of a HPA can be defined by the following power series:

$$v_o = a_1 v_i + a_2 v_i^2 + a_3 v_i^3 + a_4 v_i^4 + a_5 v_i^5 + \dots (etc.) \quad (2.2)$$

where the dependency on time has been dropped for simplicity. This can also be used as a method of modeling and is further discussed in section 2.3.3. The power series is presented in the form of a relationship between the HPA's output RF signal, v_o , and input RF signal, v_i . For more details see [26].

By substituting the two tone input defined by (2.1) into the characteristic equation of the HPA defined by (2.2), the output signal will therefore be:

$$\begin{aligned} v_o = & a_1 v [\cos(\omega_1 t) + \cos(\omega_2 t)] \\ & + a_2 v^2 [\cos(\omega_1 t) + \cos(\omega_2 t)]^2 \\ & + a_3 v^3 [\cos(\omega_1 t) + \cos(\omega_2 t)]^3 \\ & + a_4 v^4 [\cos(\omega_1 t) + \cos(\omega_2 t)]^4 \\ & + a_5 v^5 [\cos(\omega_1 t) + \cos(\omega_2 t)]^5 + \dots (etc.) \end{aligned} \quad (2.3)$$

With the aid of basic trigonometric relationships, (2.3) can be expanded into individual tones of different orders and magnitudes. This expansion is long and tedious and will be omitted here but can be found in [5]. However, the result from this expansion is important so a basic description of it is necessary.

Each order of the polynomial will produce a number of tones equal to twice that order (i.e. equal to $2n$, where n is the order of the polynomial), and each tone will be scaled by a factor of $x_i a_n v^n$, where x_i is a specific fraction that scales the tone and has been obtained from the expansion. The tones from each polynomial degree are centered at frequencies that are products of that degree. For example, the products of the third order of the polynomial are 6 tones centered at $3\omega_1$, $3\omega_2$, $2\omega_1 \pm \omega_2$ and $2\omega_2 \pm \omega_1$, while the fourth order of the polynomial produces 8 tones centered at $4\omega_1$, $4\omega_2$, $3\omega_1 \pm \omega_2$, $3\omega_2 \pm \omega_1$ and $2\omega_2 \pm 2\omega_1$. In general, the tones produced by the even-orders are centered at frequencies distant from the main band at which the original two tones are located, and are therefore not as important as the products from the odd-orders of the polynomial.

The most important products are those which produce the greatest effects on the original signal and are therefore those that are nearest to the signal. These are termed the intermodulation (IM) products. The third-order intermodulation (IM3) products are those which mainly emerge from the third-order of the polynomial and are centered at $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$. Both tones are scaled by a factor of $(3/4)a_3 v^3$; however, at these frequencies there are also components from higher odd-order products. The fifth-order of the polynomial creates two fifth-order intermodulation (IM5) products centered at $3\omega_1 - 2\omega_2$ and $3\omega_2 - 2\omega_1$ which are

scaled by an amplitude of $(5/8)a_5v^5$, but also contributes the same amplitude to the IM3 products. In this same fashion, each odd order polynomial will add to the output spectrum of the HPA. Theoretically odd-order intermodulation products never cease to contribute to the output spectrum. In practice however, contributions from high-order IM products are usually neglected. It is a common practice to take into consideration only IM3 products, but sometimes IM5 and even IM7 products have been used for more precise and complex calculations. A diagram representing a typical output spectrum produced by a two-tone input is shown in figure 2.3.

In practice, the power of the IM products can be obtained through testing and measurements. From the results of these measurements and by using the expanded version of (2.3), the exact values of the a_n coefficients of the odd-orders of the polynomial can be found. If higher-order polynomials are necessary to characterize the HPA's nonlinearity, then more mathematical analysis will be needed in order to obtain the coefficients.

This method both characterizes and models the AM/AM characteristics of HPAs and has therefore been commonly used in the past. However, it creates a way of modeling amplifiers without the use of even-order polynomials. This, as well as the fact that the two-tone method cannot do any AM/PM characterizations, has lead researchers to develop more complete methods of characterizing HPAs.

Figure 2.3: Typical Spectral Output of the Two-Tone Test

2.2.3 The Unequal Three-Tone Method

This is a more novel technique for characterizing AM/AM and AM/PM distortion in HPAs by using only power measurements [10]. It was derived from an older technique created in 1956 [18]. The method consists of inserting an unequal three-tone signal composed of a low-index AM signal superposed on a low-phase deviation PM signal. This is represented as:

$$v_{in}(t) = A_o [1 + \alpha_o \cos(\Delta\omega t)] \cos[\omega t + \beta_o \sin(\Delta\omega t) + \theta_o] \quad (2.4)$$

where α_o and β_o are the amplitude index of the AM signal and the phase deviation constant of the PM signal respectively. A_o and θ_o are respectively, the average amplitude of the AM signal and the initial phase value of the PM signal.

This is better viewed as an unequal three-tone signal by expanding (2.4) to the following form:

$$\begin{aligned} v_{in}(t) = & A_o \cos(\omega t) + [A_o(\alpha_o + \beta_o)/2] \cos[(\omega + \Delta\omega)t] \\ & + [A_o(\alpha_o - \beta_o)/2] \cos[(\omega - \Delta\omega)t] \end{aligned} \quad (2.5)$$

When this unequal three-tone signal is used to drive a HPA, the output will be another three tones of different amplitudes as it can be seen in figure 2.4. The method relies on the fact that all HPAs are characterized by G , c and k_p which respectively represent the gain of the HPA at any given input, its AM/AM compression factor and its AM/PM conversion coefficient (described in [18]). These can be found by setting the power levels of the three input tones to certain values and then measuring the power of the three output tones. Hereafter, through some mathematical manipulations, the AM/AM and AM/PM characteristics can be found. A_o^{in} , A_r^{in} , A_l^{in} , A_o^{out} , A_r^{out} and A_l^{out} are the power levels of the three input tones and the three output tones in that order.

In practice, A_o^{in} must be changed to test discrete input power levels. At the same time, A_r^{in} must always be maintained around 30 dB below A_o^{in} . Another requirement is that A_l^{in} must be set at around 3 dB below A_r^{in} . At every change of the input power, A_o^{out} , A_r^{out} and A_l^{out} must be measured and recorded.

The values of c and k_p have to be calculated at every value of A_o . This is done by using the following equations [10]:

Figure 2.4: Input and Output Spectrum of the Unequal Three-Tone Method

$$c = \frac{R_2^2 - R_1^2 + (x_2 + x_1)(x_2 - x_1)}{2(x_1 - x_2)} \quad (2.6)$$

$$k_p = \pm \sqrt{R_1^2 - (c - x_1)^2} \quad (2.7)$$

where,

$$x_1 = \frac{2A_r^{\text{in}}}{A_r^{\text{in}} + A_l^{\text{in}}}, \quad x_2 = \frac{2A_l^{\text{in}}}{A_r^{\text{in}} + A_l^{\text{in}}},$$

$$R_1 = \frac{2A_r^{\text{out}}}{G(A_r^{\text{in}} + A_l^{\text{in}})}, \quad R_2 = \frac{2A_l^{\text{out}}}{G(A_r^{\text{in}} + A_l^{\text{in}})},$$

$$G = \frac{A_o^{\text{out}}}{A_o^{\text{in}}} \quad (2.8)$$

By finding the gain, G , at the various input power levels, a discrete version of the AM/AM characteristics of the HPA can be found. The phase

insertion created by the HPA can be found at any specific input power level by integrating k_p from negative infinity up to that power level:

$$\phi(P_{drive}) = \int_{-\infty}^{P_{drive}} k_p dP_{in} \quad (2.9)$$

The sign of k_p cannot be found and must be known beforehand.

By finding the phase shift at every input power value, one can approximate the AM/PM characteristics of the HPA from the discrete measurements.

This method has been proven very effective for characterizing HPAs. All that is needed is a signal generator that can create up to three tones and a spectrum analyzer. The need for an expensive network analyzer vanishes. However, the method leaves researchers with a large number of calculations. These can be further reduced by creating similar but simpler methods.

2.2.4 The Unequal Two-Tone Method

This technique is a more general case of the two tone method, and a more specific case of the unequal three-tone method. Like the unequal three-tone method, it can be used to find both AM/AM and AM/PM characteristics of a HPA. However, instead of applying a signal formed of three tones of unequal power to the input of the HPA, only two tones are used for the analysis. Assuming that a low index modulation is used ($\alpha \ll 1$) in order to allow for

small angle approximations and neglect high order terms, these two tones can be extracted from a signal of the form [3]:

$$\begin{aligned} V_i &= A[1 + \alpha \cos(\omega_m t)] \cos[\omega t + \alpha \sin(\omega_m t)] \\ &\approx A[\cos \omega t + \alpha \cos(\omega + \omega_m)t] \end{aligned} \quad (2.10)$$

A signal with such a low modulation index is achieved by adjusting one of the carrier tones to be 20 to 30 dB below the other. If this signal is then applied to the input of the HPA, the output will be of the form:

$$\begin{aligned} v_o &= GAm(t) \cos[\omega t + \theta(t)] \\ m(t) &= 1 + (1 - c)\alpha \cos \omega_m t \\ \theta(t) &= \alpha \sin \omega_m t + k_p \alpha \cos \omega_m t \end{aligned} \quad (2.11)$$

where G , c and k_p are defined as before.

The amplification process creates distortion in the output signal, as illustrated in figure 2.5. One will notice the creation of a third tone. With some mathematical analysis, (2.11) can be used to achieve two separate equations that define Δ_{out}^{LSB} and Δ_{out}^{USB} . By solving these equations simultaneously, the AM/AM compression factor and AM/PM conversion coefficient can be obtained at any desired input power:

$$c = 1 + 10^{\Delta_{out}^{USB}/10} \left(10^{\Delta_{out}^{LSB}/10} - 10^{\Delta_{out}^{USB}/10} \right) \quad (2.12)$$

$$k_p (rad) = \pm 2 \sqrt{10^{(\Delta_{out}^{USB} - \Delta_{out}^{LSB})/10} - c^2 / 4} \quad (2.13)$$

This method requires fewer calculations than the unequal three-tone method. From a simple creation of two tones, two equations are used to obtain c

Figure 2.5: Input and Output Spectrum of the Unequal Two-Tone Method

and k_p through the use of two variables. This creates a more speedy and efficient process. One can perform the same procedure to find the values of c and k_p at any input power value. Therefore, in the same way as it can be done for the unequal three-tone method, the AM/AM characteristics can be found by calculating the gain at discrete input power values.

Phase insertion created by the amplifier can also be measured at any discrete input power level by integrating k_p in the following manner:

$$\phi(P_{drive}) = 6.6 \int_{-\infty}^{P_{drive}} k_p(rad) dP_{in} \quad (\text{deg}) \quad (2.14)$$

where the factor of 6.6 is due to the conversion from $k_p(rad)$ to $k_p(deg)$. By collecting the values of the phase insertion at every input power level, a discrete version of the AM/PM characteristics of the HPA can be obtained.

This method gives users the same end result as the unequal three-tone method while allowing similar precision. Its advantage over the unequal three-tone method is that it requires the use of a more common device, a two-tone signal generator, and that it entails fewer calculations. Nevertheless both methods prove to be just as reliable.

2.2.5 Using Digital Modulation

Digital modulation is a more novel method of characterizing HPAs and one which is growing in popularity. The technique uses an approach similar to that of the two tone method but provides measurements of the AM/PM characteristics. This is done by creating a digitally modulated binary phase shift keying (BPSK) signal with a 1010... data pattern. Unlike other multi-tone methods, one single set of measurements are sufficient to find a HPA's characteristics [12].

The method is quite simple and requires less testing than the previous methods. A 1010... data pattern at a specific data rate is used to create a BPSK signal. The spectrum of the BPSK signals will consist of two main tones which cover a bandwidth equal to twice its data rate and are centered at zero, as well as smaller tones which occur at odd multiples of the data rate. A synthesizer is then used to shift the signal to the carrier frequency at which the HPA is to be tested. At this stage, the signal has become somewhat similar to the one used for the two tone test. The BPSK signal is then passed through the HPA.

By measuring the magnitude and phase of the input and output signals, one will see that a recurring pattern is formed. This pattern is due to the 1010... data. From it, a comparison can be drawn between the input and output signals at discrete intervals of time. At such intervals, the instantaneous input power can be found, together with the gain and phase shift in the output signal. From a single set of input and output data measurements, this process can be used to obtain the gain and phase shift created by the HPA at discrete input power levels. These discrete measurements are then used to find the AM/AM and AM/PM characteristics of the HPA.

This method is faster and requires less testing than the other three methods that employ multiple tones. It is also effective for characterizing HPAs which will be used to amplify RF digital signals. Because the characteristics of some HPAs vary according to the data rate, this method can be used to determine these characteristics.

2.3 Modeling Procedures

PA characteristics and nonlinearities need to be described and represented with great accuracy. This is because through these representations one can fully understand the behaviour of any HPA and can therefore use the information to find and rectify any errors that are caused by these nonlinearities. It is common to term these representations that describe the behaviour of a HPA as “models.”

There are two ways of modeling a HPA: either by representing its AM/AM and AM/PM characteristics, or by illustrating the behaviour of its in-phase and quadrature components. Since modeling the AM/AM and AM/PM characteristics of a HPA is more common and indeed a logical step following the characterization methods described in the previous section, these types of models will be the focus of this section.

To model the AM/AM and AM/PM characteristics of a HPA it is necessary to represent the data obtained from the characterization process in a continuous form. Such forms are typically mathematical models, which can be represented as graphs for illustration purposes. There are various models that have been used in the past, as well as some that are more recent. Each model varies in complexity and accuracy and each has been used for different purposes.

2.3.1 Saleh's Model

In 1982, Adel A. M. Saleh created a method for modeling TWTAs [24]. This method was created to describe discrete readings obtained from one of the characterization processes. Saleh's model has become very popular because of its simplicity and because it can be used to model either the AM/AM and AM/PM characteristics or the in-phase and quadrature components. The model is still used today for many modeling purposes as well as a reference of comparison for newly developed models.

Saleh developed two simple formulas to model the AM/AM and AM/PM characteristics of a TWTA:

$$M(\rho) = \frac{\alpha_a \rho}{1 + \beta_a \rho^2} \quad (2.15)$$

$$\Phi(\rho) = \frac{\alpha_\phi \rho^2}{1 + \beta_\phi \rho^2} \quad (2.16)$$

where $M(\rho)$ and $\Phi(\rho)$ are respectively the AM/AM and AM/PM characteristics of the amplifier being modeled, with reference to ρ , the amplitude of the input waveform. The constants α_a , β_a , α_ϕ and β_ϕ are uniquely defined for each amplifier and are obtained by employing standard minimum mean-square-error curve-fitting procedures on the data obtained from the characterization process.

It is found that (2.15) and (2.16) take on the general form:

$$z(\rho) = \frac{\alpha_z \rho^n}{1 + \beta_z \rho^2} \quad (2.17)$$

where n is equal to either 1 or 2.

If m pairs of measurements are made throughout the characterization process, where the measured pairs take the general form (z_i, ρ_i) , and $i = 1, 2, \dots, m$, then optimum values of α_z and β_z can be found for (2.17) to best fit the data. This can be done by using standard minimum mean-square-error curve-fitting procedures to create two formulas that define the following parameters (details in [24]):

$$\alpha_z = \frac{(\sum \rho_i^2)^2 - m \sum \rho_i^4}{(\sum \rho_i^2)(\sum w_i \rho_i^2) - (\sum \rho_i^4)(\sum w_i)} \quad (2.18)$$

$$\beta_z = \frac{(\sum \rho_i^2)(\sum w_i) - m \sum w_i \rho_i^2}{(\sum \rho_i^2)(\sum w_i \rho_i^2) - (\sum \rho_i^4)(\sum w_i)} \quad (2.19)$$

where w_i is defined as:

$$w_i = \left(\frac{z_i}{\rho_i^2} \right)^{-1} \quad (2.20)$$

for $i = 1, 2, \dots, m$.

Saleh's model has proven to be accurate when modeling many TWTAs. It also provides a simple and quick method of modeling amplifiers by using the data collected when characterising them. However, since solid state power amplifiers (SSPAs) have different characteristics than TWTAs, this method might not always be a reliable one.

2.3.2 Ghorbani and Sheikhan's Model

Based on the model developed by Saleh for TWTAs, in 1991 A. Ghorbani and M. Sheikhan created one with characteristics suitable for SSPAs [11]. The model uses a similar method to that used by Saleh but takes into account the various factors that distinguish a SSPA from a TWTAs.

This model is solely based on two equations that represent the AM/AM and AM/PM characteristics of a SSPA:

$$M(\rho) = \frac{x_1 \rho^{x_2}}{1 + x_3 \rho^{x_2}} + x_4 \rho \quad (2.21)$$

$$\Phi(\rho) = \frac{y_1 \rho^{y_2}}{1 + \rho_3 \rho^{y_2}} + y_4 \rho \quad (2.22)$$

where $M(\rho)$ and $\Phi(\rho)$ are the AM/AM and AM/PM characteristics of an SSPA respectively with reference to ρ , the amplitude of the input waveform. The parameters $x_1, x_2, x_3, x_4, y_1, y_2, y_3$ and y_4 are constants that uniquely model the characteristics of individual HPAs. These constants can be obtained through various curve-fitting techniques.

Ghorbani and Sheikhan's model is more accurate than Saleh's model for SSPAs. However, it requires finding eight different parameters to fully characterize a HPA, twice as many as it is needed in Saleh's model. Nevertheless, this model is preferred to Saleh's when the amplifier to be modeled is a SSPA.

2.3.3 The Polynomial Model

Polynomials have been used to model HPAs extensively in many applications throughout the past and are still a popular method of doing so at present. The polynomial model, sometimes referred to as a power series representation, creates a simple and highly effective technique for modeling

HPAs. It is especially useful for HPA linearization purposes by means of predistortion.

A HPA's AM/AM and AM/PM polynomial models can be expressed as power series representations of the following form:

$$M(\rho) = a_0 + a_1\rho + a_2\rho^2 + a_3\rho^3 + a_4\rho^4 + a_5\rho^5 + \dots(etc.) \quad (2.23)$$

$$\Phi(\rho) = b_0 + b_1\rho + b_2\rho^2 + b_3\rho^3 + b_4\rho^4 + b_5\rho^5 + \dots(etc.) \quad (2.24)$$

where the AM/AM characteristics $M(\rho)$ and the AM/PM characteristics $\Phi(\rho)$ are directly related to the amplitude of the input waveform ρ using a power series with coefficients a and b respectively.

Equations (2.23) and (2.24) take on a general form which can be used to model any type of nonlinear behaviour including that of HPAs [5], [15]. With the help of polynomial curve-fitting tools, typically computer aided design (CAD) software, the appropriate values of the a and b coefficients can be found so that equations (2.23) and (2.24) best fit the data obtained from the characterization process. It must be noticed however, that since (2.23) and (2.24) are infinite series, they must be truncated at some point in order for the CAD software to compute the algorithm. Customarily this truncation is chosen with the purpose of creating a polynomial of odd-order, typically of 3rd, 5th or 7th order. In general, a polynomial of larger order will produce a more accurate model of the HPA but will require more resources and processing power to compute. It is customary that both the AM/AM and AM/PM polynomial models are of the same order.

The polynomial model creates an easy and quick method of modeling HPAs. It also provides freedom of design in choosing a higher-degree polynomial for greater accuracy, or a lower-degree polynomial for less complexity.

A closely linked technique is the minimum distortion power polynomial model (MDP-PM), which is created as a variation of the polynomial model. This method also tries to create a polynomial to represent the characteristics of a HPA; however, in this method the coefficients that represent every power of the polynomial are not found by using the least-squares criteria. As its name indicates, the coefficients are found by means of minimizing the distortion power of the error signal between the HPA's actual characteristics and the model itself [16], [17]. The process of developing the polynomial is a complex one and involves many calculations, but the model which is created is an effective one.

Another popular variation of the polynomial modelling technique is the Volterra series. This method creates a polynomial designed to model the nonlinearities of a system through a set of linear functions. The Volterra series is sometimes considered a modeling technique of its own. However, because it uses polynomials for the modelling process, it can be thought of as a more explicit type of polynomial model. Like the Volterra series, there are several other modeling techniques that can be seen as variations of the polynomial model, most of which are still viewed as polynomials. Their main difference however, is that they use unconventional techniques to calculate the various coefficients that compose these polynomials.

2.4 Linearization Schemes

Linearization is the process of altering the characteristics of the HPA in order to allow a more linear input-output power relationship. This allows the HPA to be operated at a more efficient power level. Generally, various physical components are added to the HPA in order to create this change in its behaviour. Many different linearization schemes have been developed yet, feedback, feedforward and predistortion are the most commonly used methods today. Alternate methods are currently being investigated, but none of these have been widely used in wireless or microwave systems [19], [23].

2.4.1 Feedback Linearization

Although feedback linearization can be categorized into various types, two main methods are common: RF feedback and Cartesian feedback [23]. The first of these applies linearization on the RF signal itself. The main idea behind this linearizer is simple and can be seen in figure 2.6. Initially, an error signal is found by subtracting the attenuated output signal of the HPA from the input signal. This error signal is then used to control the gain of the HPA through the use of a variable attenuator. Fast and accurate control of the variable attenuator can provide good amplitude linearity. However, the system's linearity can be further improved by correcting any phase errors created by the HPA. To do this, the linearizer is equipped with a phase detector which measures the difference

Figure 2.6: Block Diagram of the Feedback Linearization Method

between the output and input phase and inserts a phase correction after the variable attenuator. Because delays in the feedback must be very small in order to maintain the system stable, RF feedback can usually only be used for HF and low VHF frequencies.

Cartesian feedback linearizes the input signal by separating it into its in-phase and quadrature components [15]. To do this, two identical feedback systems are used. Each one of the systems behaves in the same manner as an RF feedback linearizer, but only deals with one of the signal's components. This creates the need for up-converters. These up-converters reconstruct the input signal to a RF form after it is passed through the variable attenuators, and down-converters to separate the amplified signal into its in-phase and quadrature components so it can be fed back into the system. Because the Cartesian feedback linearizes the Cartesian components of the signal independently, the phase of the

signal is linearized in the process. Cartesian feedback is generally used with systems that require DSP implementation.

Feedback linearization is cheap and easy to implement and is robust against imprecise characterizations of the HPA. It also adjusts to changes in the characteristics of the HPA due to aging or temperature variations. However, this system is greatly limited by its incapacity to deal with wide bandwidth signals. Because such signals generate large delays in the feedback loop, the system can become unstable.

2.4.2 Feedforward Linearization

When the input signal to the HPA has a very wide bandwidth, delays in the feedback loop become length and therefore feedback linearization becomes ineffective. For such signals, feedforward linearization might be the answer. However, it must be noted that such linearizers are typically rather complex to implement and cannot always be added easily to a HPA [15]. A block diagram of the feedforward linearization technique can be seen in figure 2.7.

Initially the input signal is split up into two branches. The first branch allows the signal to pass through the HPA. The result is an amplified input which contains some distortion due to HPA nonlinearities. This signal is then split up into two branches once again. One branch of the amplified signal is passed through an attenuator to bring it back to the level of the initial input

Figure 2.7: Block Diagram of the Feedforward Linearization Method

signal. The second branch of the input signal is then subtracted from the amplified-attenuated signal to obtain the error signal. This error signal is essentially the distortion which is added to the input signal during the amplification process. Finally the error signal is amplified using a low-power, highly-linear amplifier and then subtracted from the amplified signal (this can be done in practice by using a simple RF combiner). Ideally, this creates a distortion-free amplification of the input signal.

In order to compensate for phase distortion, phase detectors can be inserted to reveal the phase of the input and amplified signals. The information from these detectors can be used to change the phase of the error signal to eliminate phase nonlinearities. Also, delay lines might be used in various branches of the linearizer if the signal's bandwidth creates considerable delays.

Feedforward linearization gives high linearity improvements and allows the use of wideband signals. However, to avoid distortion of the error signal, the error amplifier has to be operated at a high backoff. Such an amplification process is deemed highly inefficient, therefore bringing the efficiency of the whole system down, and making this linearizer's power consumption highly inefficient when compared to other linearizers [23]. One further fact that must be considered when using this technique is that it is very sensitive to changes in the characteristics of its components. Because the linearizer is formed from an open-loop process, changes due to time, temperature and signal level, can produce a mismatch in amplitude and phase measurements and therefore impede precise distortion cancellation. Components in the feedforward techniques must be constantly monitored and the open loop must be measured and balanced for optimal performance.

2.4.3 Predistortion Linearization

By knowing the characteristics of the HPA which is used, the input signal can be distorted before amplification so that once it passes through the HPA, the cascade process will result in lower distortion of the amplified signal [15], [19], [25]. A block diagram of the predistortion process can be seen in figure 2.8. The predistorter is created in such a way that its characteristics will reduce the nonlinear characteristics of the HPA, ideally producing a linear predistorter-HPA cascade.

Figure 2.8: Block Diagram of the Predistortion Linearization Method

Predistortion can be categorized into two main types: RF predistortion (also called analog predistortion) and digital predistortion [25]. The main difference between the two is the method which is used to implement the predistorter. Nevertheless, both require accurate knowledge of the AM/AM and AM/PM characteristics of the HPA. A widespread method for creating an RF predistorter in the past has been to use some sort of mathematical equation to model the characteristics of the HPA, very often a polynomial equation. From such a model one would be able to derive an inverse model for the predistorter, which when used in cascade with the HPA would create a more linear system. Such predistorters generally rely on the ability of the model to accurately describe the behaviour of the HPA. This means that an inaccurate model of the AM/AM and AM/PM characteristics of the HPA will typically result in an ineffective linearizer.

Digital predistortion on the other hand, does not rely on a polynomial model of the HPA's characteristics, but on knowledge of its behaviour at discrete instants. By taking measurements of the characteristics of a HPA at various input power values, these characteristics can be reconstructed. The samples representing the HPA characteristics can then be used, in the same way as for an RF predistorter, to obtain another set of points that shape the characteristics of a predistorter that would produce a linear PD-HPA cascade. The predistorter is therefore formed of two lookup tables that predistort either the amplitude and phase or the in-phase and quadrature components of the input signal. These lookup tables use basic interpolation techniques that allow any input signal to be predistorted. Such signals can be digital or analog and either at baseband, IF or RF frequencies. Predistortion can even be done on the constellation symbols themselves. This predistortion method is referred to as pre-warping.

Both RF and digital predistortion can be used for signals with wide bandwidths, generally wider than those which a feedforward system is able to process, and will obtain large linearization improvements. However, they require very accurate knowledge of the HPA's characteristics. This means that variations in these characteristics due to time or temperature must be monitored and any changes must trigger adjustments to the predistorter in order to maintain linearity. This can be done by using adaptive predistortion schemes. An adaptive system compares the output signal of the HPA to the input signal and looks for differences. These differences can be processed to account for changes in the behaviour of the HPA and used to update the predistorter. This ensures a constantly linear output.

2.5 Literature Review

RF predistortion has generally relied on polynomials to create an appropriate model for the predistorter. It is through various manipulations of these polynomials that a predistorter's characteristics are derived. In 1987, Pupolin and Greenstein created one of the first and most popular 3rd order polynomial predistorters for a TWTA in [22]. It was named the cubic predistorter and was designed to create a more linear PD-HPA cascade by eliminating the cubic nonlinearity term from the output signal. The idea was that because the cubic term was the main contributor to the nonlinearity effect, if it could be eliminated then the output would become more linear.

Their method for achieving this predistorter relied on the creation of a 3rd order polynomial that would model the AM/AM and AM/PM characteristics of the HPA. The polynomial would be created through curve-fitting techniques to best represent these characteristics. Thereafter, a predistorter which would combine with the HPA to eliminate cubic nonlinearities was mathematically derived from the polynomial.

Because even order nonlinearity effects occur at frequencies away from the main bandwidth of the signal, their cancellation was not an issue. However, higher odd-order nonlinearities still had an effect on the output signal. Nonetheless, the results proved encouraging and the idea of eliminating odd-order nonlinearity effects became attractive to researchers.

Soon after in 1989, Karam and Sari made the next significant contribution by creating the predistorter described in [14]. This predistorter also aimed at eliminating the odd order nonlinearity effects, but took the 5th order effects into consideration. In the same manner as in [22], curve-fitting procedures were used to create a 5th order polynomial that would model the AM/AM and AM/PM characteristics of the HPA. Naturally this predistorter created a more linear amplification process since both 3rd and 5th order nonlinearity effects were eliminated from the output signal.

The general idea used to create the predistorters presented in [14] and [22] spread, and soon after, various similar predistorters were created. Some aimed at eliminating up to 7th and 9th order nonlinearities through the same technique. However, higher order polynomial models meant that the predistorter would be more complex, and therefore harder to implement. Soon the focus shifted to creating curve-fitting techniques that would model the characteristics of a HPA more accurately. Achieving a more accurate model of the HPA would inevitably lead to creating a more effective predistorter.

In 1999 Lai and Bar-Ness used a novel curve-fitting technique described in [17] to model the characteristics of a TWTA. This curve-fitting technique was named the minimum distortion power polynomial model (MDP-PM). It focused on minimizing the distortion power of the error signal between the actual output of the HPA and the output of the model itself. The MDP-PM was used to model the characteristics of a TWTA using 1st, 3rd, 5th and 7th order polynomials. From these models, four different predistorters were fashioned and a general trend

was confirmed that the more accurately the characteristics were modeled, the better the predistorter's performance would be.

Nevertheless, in 1995 D'Andrea et al. devised a more effective way of creating a predistorter. In [8], this group of researchers studied the idea of creating a polynomial to model the ideal characteristics of a predistorter instead of the characteristics of the HPA itself. To do this, they first established that by knowing the AM/AM and AM/PM characteristics of any HPA, one could obtain an inverse relationship for these characteristics, which would guarantee a perfectly linear cascade over a certain input power range. These inverse relationships would be what an ideal predistorter would desirably look like. It would be these ideal characteristics that would be modeled through polynomial curve-fitting techniques to create two models that would linearize the amplitude and phase separately. This method was later named the amplitude and phase predistortion technique.

Two sets of predistorters were created for a TWTA using this method: a 3rd and a 5th order. Their results showed that both sets of predistorters showed vast improvements when compared to the predistorters created previously. They both performed better than Pupolin and Greenstein's 3rd order predistorter or Karam and Sari's 5th order predistorter. In [9], D'Andrea et al. further examined the amplitude and phase predistortion technique by studying its performance when linearizing 64-QAM and 256-QAM signals.

Using the same process, in 2001 Lai and Bar-Ness applied their MDP-PM to create a model of the ideal characteristics of a TWTA in [16]. From such a model they created 3rd and 5th order predistorters. Such predistorters proved to offer comparable results to the amplitude and phase predistorters. It was confirmed that a RF predistorter would have to model the characteristics of an ideal predistorter as flawlessly as possible in order to obtain the best possible performance.

However, digital predistorters are not limited by the accuracy of a model. In 1997 D'Andrea et al. created a digital predistorter that was in line with their previous work [6], [7]. This new predistorter followed the same steps to arrive at the characteristics of an ideal predistorter, but did not need to model these. By simply creating a set of two lookup tables, one for the ideal amplitude and one for the ideal phase characteristics, the ideal predistorter was implemented. The performance obtained by such a predistorter is ultimate.

Because of the nature of digital predistorters, their implementation is done using DSP technology. This removes the need for complex polynomials, which therefore eliminates nonlinearities in the output signal due to model inaccuracies. However, signals processed through digital predistorters are subject to the same limitations that apply to any DSP system, such as processing speed and accuracy. At many times it is more convenient to use a RF predistorter than a digital one.

SIMULATION MODEL

This chapter describes the model which is used to simulate a transmission system to test the linearity of a HPA. Furthermore, proper functioning of such a system is verified. The two different HPA models which are used for the analysis as well as a novel predistortion technique are studied in detail. The method used to measure the performance of the transmission system is also examined.

3.1 Transmission System Model

It is only usual that any linearizer which is used in conjunction with a HPA is embedded into a communication system of some sort. Therefore a model of such a transmission system needs to be created. The baseband-equivalent of the transmission system which was simulated to test the predistorter is shown in figure 3.1. The input of this transmission system is a square $N_Q \times N_Q$ - QAM signal, where N_Q is a power of two and N_Q^2 is the constellation size, M_{QAM} . The input signal consists of an array of complex-valued data symbols with a spacing of T-seconds between each other. They take the following form:

Figure 3.1: Block Diagram of the Transmission System Model

$$c_i = a_i + jb_i \quad (3.1)$$

where a_i and b_i are the in phase and quadrature components of one of the M_{QAM} symbols in the QAM constellation. These components are independent, identically distributed random variables which take their values from the set $\{\pm 1, \pm 3, \dots, \pm(M-1)\}$.

The blocks G_T and G_R represent the shaping filters at the transmitting and receiving ends respectively. These shaping filters are square-root raised cosine (SRRC) filters with a rolloff factor β . The first one of these filters, G_T , allows the stream of symbols to be converted to an analog waveform $x(t)$. This waveform is then passed through the predistorter and the HPA, ideally resulting in an amplified, low-distortion version of the original waveform.

In order to view and compare the amplified signal to the original signal, the new waveform must first be scaled down by means of an attenuator. This attenuator is preset to reduce the power of the new signal by the same amount as that by which it has been linearly amplified through the PD-HPA cascade. The resulting signal $\hat{x}(t)$ is the same as the originally created signal $x(t)$ but contains any distortion due to nonlinearities in the PD-HPA cascade.

To further compare the original data to the amplified data, $\hat{x}(t)$ must be converted back into a QAM constellation. To do this, the waveform is passed through another SRRC filter, G_R , and then sampled at the optimal phase which will eliminate any ISI. The result is an array of symbols which take the form:

$$\hat{c}_i = \hat{a}_i + j\hat{b}_i \quad (3.2)$$

The array of received symbols $\hat{c}_i$ are therefore comparable to the array of transmitted symbols c_i in the same way as the received analog waveform $\hat{x}(t)$ is comparable to the original waveform $x(t)$.

3.2 HPA Models

In order to test the predistorter, two different models for two different types of HPAs were created. The first one of these models was for a SSPA, while the second model was for a TWTA. These are the two main types of HPAs available in the market. Each type has a different set of characteristics and it is therefore important to test a predistorter's performance with both types of HPAs.

3.2.1 SSPA

SSPAs are generally less nonlinear TWTAs. They are also more durable with an average life expectancy of ten to twelve years, versus the average six years of life a TWTA is expected to have. Still, failures tend to occur more frequently in SSPAs. They can both amplify signals of wide bandwidths and the linear gain of both varies from one HPA to another. However, TWTAs are normally more efficient; the average efficiency of a TWTA is of around thirty percent, while that of a SSPA is of around fifteen to twenty percent.

In order to obtain the model of a SSPA that could be used in the simulation, the AM/AM and AM/PM characteristics of an actual SSPA had to be extracted. Due to the availability of HPAs in the laboratory, the SSPA that was chosen for characterization was model ZHL-42 manufactured by Mini-Circuits.

The method chosen to characterize this SSPA was through the use of a network analyzer. The network analyzer used to perform this characterization was the 8720ES model manufactured by Agilent. After correct calibration of the equipment, the ZHL-42 HPA was connected to the network analyzer as shown in figure 3.2. A DC power supply was used to feed the SSPA with the 15 volts necessary to allow it to function correctly. A 20 dB attenuator was connected to the output of the SSPA in order to reduce its output power and not overload the network analyzer [2]. Furthermore, the presence of the attenuator was programmed into the network analyzer to enable accurate measurements.

Figure 3.2: Block Diagram of the Characterization Process for the ZHL-42 HPA

Various power sweeps were performed on the ZHL-42 SSPA in order to obtain its AM/AM and AM/PM characteristics at various frequencies ranging from 995 MHz to 4005 MHz. The data describing these characteristics between an input power ranging from -20 dBm to 10 dBm in steps of 0.15 dB were obtained and stored directly from the network analyzer onto a floppy disk. This data was later converted from power (P) into a voltage (V) format through the conversion $V = 10^{\frac{P}{20}}$. This was used to create a set of lookup tables to simulate the behaviour of the HPA. By linear interpolation of the data, a continuous model of the AM/AM and AM/PM characteristics of the ZHL-42 SSPA was obtained. In order to represent the characteristics of a SSPA in the simulation process, the extracted characteristics at a frequency of 1 GHz were chosen. These characteristics, together with the characteristics at 0.995 MHz and 1.005 MHz are illustrated in voltage format in figure 3.3.

Figure 3.3: AM/AM & AM/PM Characteristics of the SSPA at Various Frequencies

At a frequency of 1 GHz, the linear power gain of the ZHL-42 HPA is of 33.82 dB. This means that the amplitude (in volts) of the input signal will be amplified by a factor of 49.09 when this SSPA is used in its linear region. The saturation point (P_{sat}) of a HPA is defined as the input power at which maximum output power is achieved. The saturation point of this SSPA occurs at an input power of 6.1 dBm (0.064 V), at which the output power is of 32.1 dBm (1.274 V). Hence the power gain at this point is of 26 dB. The saturation point of the SSPA cannot be seen in figure 3.3 due to resolution issues.

3.2.2 TWTA

In order to test the predistortion method on more nonlinear HPAs, a TWTA model was needed. The chosen model was one which was repetitively used through literature to test various methods of predistortion [6], [7], [8], [9], [14], [16], [17], [22]. By choosing such a model, we create the possibility of comparing the results obtained to those in literature.

The TWTA model used was that described by Saleh in [24] and reviewed in section 2.3.1. The essential equations that form the characteristics of this model are given by the general equations (2.15) and (2.16). With the aim of being consistent with the literature, the exact characteristic models of the TWTA were defined by the following equations:

$$M(\rho) = \frac{2\rho}{1+\rho^2} \quad (3.3)$$

$$\Phi(\rho) = \Phi_o \frac{2\rho^2}{1+\rho^2}, \quad \Phi_o = \frac{\pi}{6} \quad (3.4)$$

where $M(\bullet)$ and $\Phi(\bullet)$ represent the AM/AM and AM/PM characteristics of the TWTA with reference to the instantaneous input amplitude ρ . In fact, the instantaneous input signal ρ has a dependency on time, but this dependency has been dropped for simplicity purposes.

By using (3.3) and (3.4), a single equation can be derived to define the TWTA model:

$$z(\rho) = \frac{2\rho}{1+\rho^2} \exp\left(j \Phi_o \frac{2\rho^2}{1+\rho^2}\right) \quad (3.5)$$

where again, the dependency on time has been dropped.

The models of the AM/AM and AM/PM characteristics of the TWTA obtained from (3.5) can be seen in figure 3.4 (in linear and logarithmic scale). By comparing these to figure 3.3, it can be seen how the TWTA has greater amplitude nonlinearity and larger amounts of phase insertion than the SSPA.

The linear power gain of this TWTA is calculated to be 6.02 dB, which means that the power of the input signal would be amplified by a factor of 4 when the HPA is used in its linear region. The saturation point of the TWTA occurs at an input of 0 dBW (1 V). The maximum possible output voltage at such an input voltage is 1 V.

Figure 3.4a: AM/AM & AM/PM Characteristics of the TWTA in Linear Scale

Figure 3.4b: AM/AM & AM/PM Characteristics of the TWTA in Logarithmic Scale

3.3 Overview of the Predistortion Technique

In order to eliminate amplitude nonlinearities and phase insertions in the amplified signal, a predistorter for both the amplitude and the phase has to be created. Such a predistorter, which would create a perfectly linear power gain and a zero phase insertion at the output of the HPA, can only be achieved theoretically. In practice however, due to physical limitations at the time of implementation, only an approximation of that ideal predistorter can be created. Such approximations are modeled through various techniques, some of which have been reviewed in section 2.2.

3.3.1 Ideal Predistorter Characteristics

Ideal predistorters can therefore be created for the SSPA and the TWTA characteristics described in figures 3.3 and 3.4. This is done by creating two separate predistorters, one for the amplitude and one for the phase, whose cascade with the HPA will create a linear input-output relationship as well as a zero-phase insertion.

If the input signal to the amplifier is of the form:

$$x(t) = \rho(t)\cos(\omega_c t + \theta(t)) \quad (3.6)$$

where $\rho(t)$ and $\theta(t)$ respectively represent the instantaneous amplitude and phase of $x(t)$ at a carrier frequency ω_c , then the output of the power amplifier can be written as:

$$z(t) = M(\rho(t))\cos(\omega_c t + \theta(t) + \Phi(\rho(t))) \quad (3.7)$$

where $M(\bullet)$ and $\Phi(\bullet)$ respectively represent the AM/AM and AM/PM characteristics of the HPA.

In the same way, if $x(t)$ was the input to the predistorter, the output of this would take the form:

$$y(t) = A(\rho(t))\cos(\omega_c t + \theta(t) + \Psi(\rho(t))) \quad (3.8)$$

where the amplitude and phase characteristics functions of the predistorter $A(\bullet)$ and $\Psi(\bullet)$ are created to eliminate all HPA nonlinearity and create a PD-HPA cascade of the form:

$$M(A(\rho)) = \begin{cases} \alpha \rho, & 0 \leq \rho \leq \rho_{\max} \\ \alpha \rho_{\max}, & \rho > \rho_{\max} \end{cases} \quad (3.9)$$

$$\Psi(\rho) + \Phi(A(\rho)) = 0 \quad (3.10)$$

where $\rho_{\max}$ is the value of the input voltage at the saturation point of the HPA, α is the voltage gain of the PD-HPA cascade and the dependency on time has been dropped for simplicity.

As a HPA is linearized closer to its saturation point, the gain linear gain of the PD-HPA cascade is reduced. However, this will allow the efficiency of the amplifier to be maximized. This is better explained by means of an illustrative example such as the one in figure 3.5. From this figure, it can be seen how three different linearization approaches create three different gains. Maximum gain is achieved through gain 3, where the system is linearized up to a input of P_{in1} . This

Figure 3.5: Three Approaches for HPA Linearization

reduces the maximum power of the input signal into the system, which as a result, does not allow the HPA to be used to its full potential and therefore reduces its efficiency. Maximum efficiency is achieved when the amplifier is linearized at P_{sat} . This creates a gain which is the lowest of the three. Linearization at P_{in2} creates an intermediate gain with intermediate efficiency.

Two different linearization requirements exist for the SSPA and the TWTA. It is generally accepted that trying to linearize a SSPA up to its saturation point is not useful because around that region the gain of the HPA falls drastically [4]. Therefore, the ideal predistorter of the SSPA was created to linearize it at its 3 dB compression point. This point is described as the input power at which the gain of the HPA falls by 3 dB from its linear gain (33.8 dB), i.e. the input power at which the gain becomes 30.8 dB. The characteristics of this HPA indicate that the 3 dB compression point occurs at an input power of 0.58

dBm (an input voltage of 0.0338 V), at which the output power is 31.38 dBm (the output voltage is 1.172 V). Hence, the PD-HPA characteristics will be described by (3.9) and (3.10) where $\alpha = 34.67$ and $\rho_{\max} = 0.0338$ V.

The ideal predistorter characteristics that would linearize the SSPA as desired were created by means of simple lookup tables that performed the conversions described in (3.9) and (3.10). These can be seen in figure 3.6.

For the TWTA, the PD-HPA cascade was created to enable a linear input-output relationship, up until the compression point of the TWTA. As defined previously, this occurs at an input power of 0 dBW (an input voltage of 1V). Hence the characteristics of the PD-HPA cascade were obtained as described in (3.9) and (3.10) with $\alpha = 1$ and $\rho_{\max} = 1$ V. Thereafter, by substituting (3.3) and (3.4) into (3.9) and (3.10), the corresponding equations for the amplitude and phase characteristics of the ideal predistorter of the TWTA were found to be:

$$A(\rho) = \begin{cases} \frac{1 - \sqrt{1 - \rho^2}}{\rho}, & 0 < \rho \leq 1 \\ 1, & \rho > 1 \end{cases} \quad (3.11)$$

$$\Psi(\rho) = -\Phi(A(\rho)) = -\Phi_o \frac{2A^2(\rho)}{1 + A^2(\rho)} \quad (3.12)$$

The ideal amplitude and phase predistorter characteristics for both the SSPA and the TWTA are illustrated in figures 3.6 and 3.7 respectively. These are the characteristics to which the implemented predistorters must approximate in order to create a linear amplification process.

Figure 3.6: Ideal Predistorter Characteristics for the SSPA

Figure 3.7: Ideal Predistorter Characteristics for the TWTA

3.3.2 Multiple-Polynomial Models

One popular method for creating a predistorter is by using polynomial models. The higher the order of these polynomials, the closer the model will be to the ideal predistorter. However, an increase in the order of these polynomials will also mean an increase in the complexity of the predistorter structure as well as an increase in the processing time of the signals. Therefore, when using polynomial models to create the predistorter, this trade-off between performance and complexity must be taken into account. A polynomial order that will optimize the predistorter's performance must be chosen to suit the application where the predistorter will be used.

Further exploration of polynomial modelling has lead to the creation of a new predistorter structure. This new structure is unlike that of a typical polynomially-modeled predistorter in that it does not use only one polynomial to model the entire characteristics of the amplitude and the phase predistorter. By splitting the ideal predistorter structure into various sections, a polynomial model can be found for each section.

Figure 3.8 shows an illustrative example of this method. The characteristics of the ideal amplitude predistorter for the TWTA are used in this example. Here, the predistorter's characteristics have been divided into m^1 sections. These sections may or may not be of the same length depending on

¹ Section $m+1$ is not considered a section that would be modeled by a polynomial

Figure 3.8: Amplitude Predistorter Characteristics divided into m Sections

requirements and capabilities. This aspect of the predistorter model will be further studied in chapter 4. After sectioning the characteristics, it is necessary to create a model for each section that will correctly define it. By using curve-fitting methods, a polynomial of n^{th} order can be created for each section. The curve-fitting technique used, finds the coefficients of a polynomial to best fit some defined points of a given section in the least-square (LS) sense.

The LS technique evaluates the $n+1$ coefficients $\{c_0, c_1, \dots, c_n\}$ of an n^{th} order polynomial of the form

$$y = c_0 + c_1x + c_2x^2 + \dots + c_nx^n \quad (3.13)$$

to best represent a set of distinct output points $\{y_1, y_2, \dots, y_q\}$ by using a set of distinct input points $\{x_1, x_2, \dots, x_q\}$.² This is better visualized by formulating the equation in a matrix form:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_q \end{bmatrix} = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^n \\ 1 & x_2 & x_2^2 & \dots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_q & x_q^2 & \dots & x_q^n \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \quad (3.14)$$

For simplification purposes the following representation of this matrix form is used:

$$\mathbf{Y} = \mathbf{X}\mathbf{c} \quad (3.15)$$

Therefore $\mathbf{c}$ must be found as the least-squares solution that minimizes the residual vector $\mathbf{X}\mathbf{c} - \mathbf{Y}$:

$$\|\mathbf{X}\mathbf{c} - \mathbf{Y}\|^2 = \mathbf{c}^T \mathbf{X}^T \mathbf{X} \mathbf{c} - 2\mathbf{Y}^T \mathbf{X} \mathbf{c} + \mathbf{Y}^T \mathbf{Y} \quad (3.16)$$

This can be done by taking the derivative of (3.16) with respect to $\mathbf{c}$ and equating it to zero, giving

$$2\mathbf{c}^T \mathbf{X}^T \mathbf{X} - 2\mathbf{Y}^T \mathbf{X} = 0 \quad (3.17)$$

The LS coefficients of the polynomial can therefore be found from

$$\mathbf{c} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \quad (3.18)$$

The greater the number of sections or the higher the order of the polynomials, the closer the m polynomials will model the ideal characteristics of

² A total of 1000 points was used to represent the entire characteristics of the HPAs.

the predistorter. Further study into the effects of the number of sections and the order of the polynomials is reviewed in section 4.1.

Therefore, for implementation purposes, the predistorter model will consist of m different n^{th} order polynomials that model the amplitude predistorter, and m different n^{th} order polynomials that model the phase predistorter. This creates a total of $2m$ different polynomials that need to be implemented. These polynomials will form the basis of the predistorter model, which will be of the form:

$$A(\rho) = \begin{cases} a_{10} + a_{11}\rho + a_{12}\rho^2 + \dots + a_{1n}\rho^n, & 0 \leq \rho < \rho_1 \\ a_{20} + a_{21}\rho + a_{22}\rho^2 + \dots + a_{2n}\rho^n, & \rho_1 \leq \rho < \rho_2 \\ \vdots & \vdots \\ a_{m0} + a_{m1}\rho + a_{m2}\rho^2 + \dots + a_{mn}\rho^n, & \rho_{m-1} \leq \rho < \rho_m \\ \alpha\rho_{\max} & \rho \geq \rho_m \end{cases} \quad (3.19)$$

$$\Psi(\rho) = \begin{cases} b_{10} + b_{11}\rho + b_{12}\rho^2 + \dots + b_{1n}\rho^n, & 0 \leq \rho < \rho_1 \\ b_{20} + b_{21}\rho + b_{22}\rho^2 + \dots + b_{2n}\rho^n, & \rho_1 \leq \rho < \rho_2 \\ \vdots & \vdots \\ b_{m0} + b_{m1}\rho + b_{m2}\rho^2 + \dots + b_{mn}\rho^n, & \rho_{m-1} \leq \rho < \rho_m \\ -\Phi(\alpha\rho_{\max}) & \rho \geq \rho_m \end{cases} \quad (3.20)$$

Here, a_{mn} and b_{mn} are the various coefficients of the amplitude and phase predistorter polynomials, and ρ_i is the maximum possible input amplitude of the i -th section, where i is an element from $\{1, 2, \dots, m\}$. One will notice that section $m+1$ is modeled by constants for both the amplitude and phase predistorters.

It is possible to combine the two polynomials that describe one section of the predistorter's characteristics into one complex equation, therefore creating a predistorter model of the form:

$$y(\rho) = \begin{cases} (a_{10} + a_{11}\rho + \dots + a_{1n}\rho^n) \exp[j(b_{10} + b_{11}\rho + \dots + b_{1n}\rho^n + \theta)] , & 0 \leq \rho < \rho_1 \\ (a_{20} + a_{21}\rho + \dots + a_{2n}\rho^n) \exp[j(b_{20} + b_{21}\rho + \dots + b_{2n}\rho^n + \theta)] , & \rho_1 \leq \rho < \rho_2 \\ \vdots & \vdots \\ (a_{m0} + a_{m1}\rho + \dots + a_{mn}\rho^n) \exp[j(b_{m0} + b_{m1}\rho + \dots + b_{mn}\rho^n + \theta)] , & \rho_{m-1} \leq \rho < \rho_m \\ \alpha \rho_{\max} \exp[j(-\Phi(\alpha \rho_{\max}) + \theta)] & \rho \geq \rho_m \end{cases} \quad (3.21)$$

where θ is the instantaneous phase of the input signal.

The implementation of this predistorter model is represented in figure 3.9. Here it can be seen that even though the model is composed of $2m$ polynomials, each of n^{th} order, only two of these will be used for computational purposes at any one instant. For implementation purposes the model depicted in figure 3.10 can be used. This model outlines the components necessary to create such a predistorter. This predistorter uses a device which stores all the possible coefficients of the different sections of the predistorter, and updates the amplitude and phase predistortion branches with them accordingly.

In theory, as m approaches infinity, the polynomials will become of 0^{th} order and the difference between the ideal predistorter and the modeled one will become nonexistent. However, in practice this cannot be achieved. Therefore in the interest of achieving faster computations, optimal values of n and m must be found.

Figure 3.9: Implementation Model of the Predistorter

Figure 3.10: Functional Block Diagram of the Predistorter

This predistortion technique will be referred to from here forth as the multipolynomial predistorter (MP-PD).

3.4 Performance Analysis

As a HPA is driven closer to its saturation point, the amount of distortion in the amplified signal increases. At the same time, if the HPA is operated at a high backoff, its operation becomes ineffective. It therefore becomes necessary to measure the HPA's degradation at different levels of backoff in order to find a point of high efficiency while minimizing distortion. To do this, the output backoff (OBO) of a HPA must be used as a reference against which degradation measurements can be made. The OBO is defined as the dB difference between the output power of the HPA at its saturation point and the average power of the amplified signal $P_{out}(x)$:

$$OBO = P_{out}(P_{sat}) - P_{out}(x) \quad (3.22)$$

Let us consider the following two transmission systems illustrated in figure 3.11:

System 1: The output of a M-QAM modulator is applied to a linear HPA. The amplified signal is corrupted by AWGN with power spectral density N_o .

System 2: The output of the same M-QAM modulator is applied to a nonlinear HPA. The amplified (and distorted) signal is corrupted by AWGN with power spectral density N_o .

When the HPA is perfectly linear, such as the case for system 1, the performance is determined by the signal to noise ratio:

$$SNR|_{linear} = \frac{S|_{linear \text{ amplifier}}}{N|_{thermal \text{ noise}}} \quad (3.23)$$

When the nonlinearity of the HPA is taken into account, the received signal will be corrupted by thermal noise as well as a noise-like error signal caused by the amplifier nonlinearity. In such a case, the system's performance is dictated by the signal to noise plus distortion ratio:

Figure 3.11: Diagrams of the Assumed Transmission Systems

$$SNDR = \frac{S|_{\text{nonlinear amplifier}}}{N|_{\text{thermal noise}} + N|_{\text{nonlinear distortion}}} \quad (3.24)$$

The degradation in performance of system 2 relative to system 1, at a given BER, is defined as:

$$\Delta = 10 \log_{10} \left(\frac{S|_{\text{nonlinear HPA}}}{S|_{\text{linear HPA}}} \right) \quad (3.25)$$

where $S|_{\text{nonlinear HPA}}$ and $S|_{\text{linear HPA}}$ are the power levels that need to be created by the nonlinear and linear HPAs to produce the required BER. The SNR required for system 1, for a given BER, can be determined from well known expressions for M-QAM signals.

At this point we shall make the approximating assumption that the nonlinear distortion component of system 2 is Gaussian and uncorrelated with the thermal noise. This assumption allows us to add the nonlinear distortion to the thermal noise in the channel, in order to numerical approximate the amount of degradation in the system. In this case the degradation Δ can be expressed as:

$$\Delta = 10 \log_{10} \left(\frac{N|_{\text{thermal}} + N|_{\text{nonlinear distortion}}}{N|_{\text{thermal}}} \right) \quad (3.26)$$

$N|_{\text{thermal}}$ can be calculated from:

$$N|_{\text{thermal}} = \frac{S|_{\text{linear HPA}}}{SNR} \quad (3.27)$$

where $S|_{linear\ HPA}$ is the measured average symbol energy, and the SNR is calculated from the SNR-BER relationship for a QAM signal represented by the following equation found in [21]:

$$BER = \frac{1 - \left[1 - 2 \left(1 - \frac{1}{\sqrt{M_{QAM}}} \right) Q \left(\frac{\sqrt{3SNR}}{\sqrt{M_{QAM} - 1}} \right) \right]^2}{\log_2(M_{QAM})} \quad (3.28)$$

Here, M_{QAM} is the constellation size of the QAM signal assuming that Gray-coding is used. This relationship can be converted into the following more useful form:

$$SNR = \frac{2(M_{QAM} - 1)}{3} \left[\operatorname{erf}^{-1} \left(1 - \frac{1 - \sqrt{1 - BER \log_2 M_{QAM}}}{1 - \frac{1}{M_{QAM}}} \right) \right]^2 \quad (3.29)$$

From this equation, the SNR can be found at the specified BER of 10^{-4} for any QAM signal constellation size.

$Q(\bullet)$ and $\operatorname{erf}(\bullet)$ are respectively the Q-function and the error function which are mathematically defined as follows:

$$Q(\varepsilon) = \frac{1}{\sqrt{2\pi}} \int_{\varepsilon}^{\infty} e^{-\lambda^2/2} d\lambda \quad (3.30)$$

$$\operatorname{erf}(\varepsilon) = 1 - 2Q(\sqrt{2}\varepsilon) \quad (3.31)$$

The value of $N|_{nonlinear\ distortion}$ can be found by calculating the mean-square-error (MSE) between the transmitted symbols and the received ones. i.e.:

$$N|_{\text{nonlinear distortion}} = \frac{1}{k} \sum_{i=1}^k |c_i - \hat{c}_i|^2 \quad (3.32)$$

where k is the number of transmitted symbols.

It is customary to calculate the amount of degradation through measurements of the total degradation (TD). Such measurements reflect the amount of distortion in the signal at any one OBO. The minimum TD at a specific OBO occurs if the amplified signal is distortion-free and its value will therefore be the same as the OBO itself. Hence the amount of TD (in dB) is defined as:

$$TD = OBO + \Delta \quad (3.33)$$

By using (3.22) through (3.33) a relationship between the OBO and the TD can be established. The operation point of optimal performance for the HPA is then found to occur at the OBO at which the minimum amount of TD is measured. This point is termed $TD_{\min}$.

All the results were measured using simulation runs where 50000 symbols were transmitted. Figure 3.12 shows the variance of the MSE measured as a function of the number of transmitted symbols. This variance was calculated using averages over 20 simulation runs, where the characteristics of a TWTA and an ideal predistorter were assumed. Measurements of the MSE were made at an OBO of 4 dB using 64-QAM signals. The rolloff factors of the transmitter and receiver shaping filters were set to 0.25. Results are displayed as a percentage variance.

Figure 3.12: MSE Variance at Various Amounts of Transmitted Symbols

This percentage variance was calculated as:

$$\% \text{ Variance} = \frac{\text{Variance of the Measured MSE} \big|_{\text{over 20 simulation runs}}}{\text{Average MSE} \big|_{\text{over 20 simulation runs}}} \times 100 \quad (3.34)$$

By transmitting 50000 symbols in every simulation run, one will allow the system to minimise errors³ while avoiding time-consuming simulation runs.

³ The variations in the transmitted signal will cause the measured MSE to vary. Such variations cause a deviation from the average MSE.

3.5 Model Verification

A test was carried out on the system in order to verify that the simulation was running correctly. For this test, the TWTA model was used to create the characteristics of a HPA. The predistorter was created to be a 5-section MP-PD modeled with 3rd order polynomials. By setting the sections to cut off at 0.35V, 0.7V, 0.9V and 0.99V, a predistorter whose characteristics were almost identical to the ideal predistorter was created. Therefore it can be assumed that an ideal predistorter was generated from the MP-PD. However, implementation complexity issues had to be neglected in order to achieve this.

By knowing the exact behaviour of the predistorter and the HPA, for a given signal travelling through the system, the waveform at the output of the predistorter and the HPA can be theoretically calculated. Thereafter, the simulation can be run using the same signal as an input, and the simulation results can be compared to the theoretical ones. If the simulation is running correctly the simulation results will match the theory.

The input signal chosen was a double sideband suppressed carrier (DSB-SC) wave of the form:

$$x(t) = X \sin(2\pi f_1 t) \sin(2\pi f_2 t) \quad (3.35)$$

where the parameters of the signal were chosen to be: $X=1.5V$, $f_1=1$ and $f_2=20$.

From (3.35) and by knowing the characteristics of the ideal predistorter, shown in figure 3.7, the output of the predistorter, $y(t)$, can be calculated. A model of the exact form of $y(t)$ is not necessary. It is simply enough to know that the general shape of the $y(t)$ will be of a DSB-SC waveform slightly distorted, where all signals above a maximum amplitude have been cut-off. This maximum amplitude can be found from:

$$|Y_{\max}| = \rho_{\max} \cos(-\Phi(\rho_{\max})) \quad (3.36)$$

where it is known from figure 3.7 that $\rho_{\max} = 1$ and $-\Phi(\rho_{\max}) = -30^\circ = -\pi/6 \text{ rad}$. Therefore the maximum amplitude at the output of the predistorter can be found to be:

$$|Y_{\max}| = \cos(-\pi/6) = 0.866 \text{ V} \quad (3.37)$$

The waveform at the output of the HPA, $z(t)$, can also be calculated. The PD-HPA cascade is known to have a linear input-output relationship when the input signal has an amplitude of up to 1V. Any signal of higher amplitude is eliminated as described by equation (3.11). Therefore $z(t)$ will theoretically have the form of $x(t)$ with the difference that any amplitude higher than 1V in $x(t)$ would become exactly 1V in $z(t)$. In other words, $z(t)$ will be a distorted replica of $x(t)$ where amplitudes above 1V have been “chopped-off.”

In order to further verify the accuracy of the simulation system, the MSE between $x(t)$ and $z(t)$ can be analytically computed using:

$$MSE_{theory} = \frac{1}{b-a} \int_a^b |x(t) - y(t)|^2 dt \quad (3.38)$$

where the constants are set to $a=0$ and $b=1$ in order to integrate over one period of the signal. This is expanded into the following form:

$$\begin{aligned}
 MSE_{theory} = 4 \left[\int_{0.157}^{0.168} |1.5 \sin(2\pi f_1 t) \sin(2\pi f_2 t) - 1|^2 dt + \int_{0.206}^{0.219} |1.5 \sin(2\pi f_1 t) \sin(2\pi f_2 t) - 1|^2 dt \right. \\
 + \int_{0.256}^{0.269} |1.5 \sin(2\pi f_1 t) \sin(2\pi f_2 t) - 1|^2 dt + \int_{0.306}^{0.318} |1.5 \sin(2\pi f_1 t) \sin(2\pi f_2 t) - 1|^2 dt \\
 \left. + \int_{0.358}^{0.366} |1.5 \sin(2\pi f_1 t) \sin(2\pi f_2 t) - 1|^2 dt \right]
 \end{aligned} \tag{3.39}$$

From this equation, the MSE was found to be:

$$MSE_{theory} = 0.018497 \tag{3.40}$$

Thereafter, the DSB-SC signal described by (3.35) was simulated and passed through the PD-HPA cascade. The resulting waveforms from this process can be seen in figure 3.13.

From these, figure 3.13a shows the DSB-SC signal that was used to test the simulation model. Figure 3.13b illustrates the signal received at the output of the predistorter. As theory dictates, this waveform has the distorted shape of the original signal with an amplitude cut-off at 0.866V as described by (3.37). Therefore the predistorter is functioning as it should.

The measured waveform at the output of the HPA can be seen in figure 3.13c. As expected, this waveform is an exact replica of the input waveform with the difference that any amplitude above 1V has been eliminated. This is further clarified through figure 3.13d which shows the difference between the input

Figure 3.13: Amplification Process of the DSB-SC Signal

waveform and the output waveform to the system. This figure displays amplitudes above 1V in the input waveform.

The MSE between the input and output waveforms was also measured by the simulation. This was done through the following computation:

$$MSE_{measured} = \frac{1}{k} \sum_k |x(k) - z(k)|^2 \quad (3.41)$$

which resulted in the following value:

$$MSE_{measured} = 0.018496 \quad (3.42)$$

This agrees with the result obtained in equation (3.40).

From these results it can be seen that the simulated system performs in the manner it was designed for. It can therefore be said that the simulation's performance is accurate and dependable.

SIMULATION RESULTS

This chapter studies the effect of changing the polynomial order and the number of sections of the multipolynomial predistorter. The results obtained by using this predistortion technique on the SSPA are illustrated while the results from the TWTA are compared with some more commonly used predistortion techniques. Finally, the effects of sampling the input signal are studied.

4.1 Effect of Sectioning and Polynomial Order

There are various factors that determine the performance of the MP-PD. A main factor is the number of sections (m) that are used to create this predistorter. In order to study the effect of m on the performance of the linearizer, various simulations were carried out. The purpose of these simulations was to identify the amount of linearity achieved by the MP-PD as m varied and compare this to the linearity obtained by an ideal predistorter. Linearity comparisons between different predistorter schemes were made using the method described in section 3.4.

These simulations were carried out using the TWTA model described in section 3.2.2 to simulate the HPA characteristics. By using the ideal predistorter characteristics of the TWTA, a set of simulations were run in order to identify the performance of this ideal linearizer. Since the ideal predistorter represents a case where no further linearity improvements can be obtained, the results from these simulations are considered to be the best performance that any predistortion technique could aim to achieve.

4.1.1 Effect of the Number of Sections

The linearization ability of the MP-PD was initially investigated. To do this, various simulations were run, allowing m to be varied from one section to a maximum of five possible sections. For the case where m was larger than one, each section was adjusted to span an equal range of input voltage. i.e.

$$\rho_i = \frac{i}{m} \quad (4.1)$$

where ρ_i is the maximum possible input voltage at the i -th section, i being an element from the set $\{1, 2, \dots, m\}$.

The shaping filters at the transmitter and receiver ends (G_T and G_R), were set to have a rolloff factor of 0.25. All sections were modeled using a 3rd order polynomial (i.e. $n=3$). Thereafter, simulations were performed using both a 64-QAM and a 256-QAM signal.

(a) Using a 64-QAM signal

(b) Using a 256-QAM signal

Figure 4.1: TD versus OBO for the MP-PD with $n=3$ at various values of m

Figure 4.1 displays the results obtained from the simulation of the system when using a 64-QAM signal and a 256-QAM signal. In the figure 4.1a, $TD_{\min}$ is obtained at an OBO of 4 dB for all values of m . At this operation point, the amount of degradation (Δ) is higher by 0.1 dB when the predistorter has only one section than when it has two or more. It is important to note that when the number of sections is greater than one, the performance of the MP-PD becomes almost identical to that of the ideal predistorter.

The same pattern of behaviour is observed when the system is subjected to a 256-QAM signal. In figure 4.1b it can be seen that $TD_{\min}$ for an ideal predistorter occurs at an OBO of 4.5 dB. This is also true for the MP-PD when m is greater than one. When the predistorter is composed of only one section, $TD_{\min}$ occurs at an OBO of 4.6 dB. However, at such an OBO, a degradation Δ of 1.19 dB occurs. It is also observed that as the OBO increases, so does the degradation of the one-section MP-PD increases. This is due to the fact that the accuracy of the fitted polynomial is reduced mainly around the beginning and end of the section it is modeling. With $m=1$, this loss of accuracy is mainly evident around an input voltage of 0 V.

A reduction in the performance of the two-section MP-PD can be seen when a 256-QAM signal is used. Even though the increase in degradation is of only 0.08 dB, it reveals that as the complexity of the amplified signal increases, the predistorter's ability to linearize a HPA might fall. Therefore, the number of sections of the MP-PD might need to be increased when amplifying highly complex signals.

When the number of sections in the MP-PD is more than one, a slight adjustment in the length of each section can allow the linearization ability of the predistorter to increase. To do this, the maximum possible input voltage at the i -th section should be found from the following equation:

$$A(\rho_i) = i \frac{A(\rho_{\max})}{m} \quad (4.2)$$

In this equation, $A(\rho_i)$ represents the predistorter's output voltage at an input voltage of ρ_i and $\rho_{\max}$ symbolizes the maximum possible input voltage.

For example, by using (4.2), a two-section MP-PD for the TWTA can be found. This would result in the first section covering from 0 to 0.8 V and the second section covering from 0.8 to 1 V. The performance of such a predistorter is illustrated in figure 4.2, where this is compared to a predistorter with equal length sections. Both predistorters have been tested using a 256-QAM signal and shaping filters with rolloff factors of 0.25. In both cases 3rd order polynomials were used to model each section.

Figure 4.2 shows how by applying (4.2) to obtain the intervals over which each section should be defined, the linearization ability of the MP-PD increases and becomes equal to that of an ideal predistorter. Because this method of sectioning provides better performance than creating sections of equal lengths, all simulations from this point forth will be done assuming (4.2) was used to define the span of the m sections of the MP-PD. Therefore sections will be defined by the following parameters:

Figure 4.2: TD versus OBO for the MP-PD with $m=2$, $n=3$ using equal and unequal-length sections (results using a 256-QAM signal)

- For $m=2$: $\rho_1 = 0.8$, $\rho_2 = 1$
- For $m=3$: $\rho_1 = 0.6$, $\rho_2 = 0.9231$, $\rho_3 = 1$
- For $m=4$: $\rho_1 = 0.4706$, $\rho_2 = 0.8$, $\rho_3 = 0.96$, $\rho_4 = 1$
- For $m=5$: $\rho_1 = 0.3846$, $\rho_2 = 0.6897$, $\rho_3 = 0.8824$, $\rho_4 = 0.9756$, $\rho_5 = 1$

4.1.2 Effect of Polynomial order

Another issue that affects the performance of the MP-PD is the order of the polynomials used to model each section (n). It is logical to assume that as the order of the modeling polynomials increases, the performance of the predistorter improves. However, it is also known that an increase in the polynomial order results in an increase in the complexity of the predistorter's structure. Hence the matter of maintaining low predistorter complexity while maximizing performance is an important issue.

In order to measure the impact of the polynomial order on the predistorter's linearization ability, various simulations were performed. The first set of simulations aimed at quantifying the linearization capabilities of a MP-PD with one single section. This type of predistorter is a special case of the MP-PD, where there is no sectioning involved and therefore only one set of polynomials are used to model the amplitude and phase predistorters. The linearization ability of such a predistorter was tested by using a 256-QAM signal. The transmitter and receiver filters were set to exhibit a rolloff factor of 0.25. The results from these simulations, where 2nd, 3rd, 4th and 5th order polynomials were used to model the predistorter, can be seen in figure 4.3a.

The second set of simulations aimed to demonstrate the linearization efficiency of the two-section MP-PD when n was varied. The same system parameters as for the one-section MP-PD were set and simulation runs were

(a) Using $m=1$ (b) Using $m=2$

Figure 4.3: TD versus OBO for the MP-PD using various polynomial orders applying a 256-QAM signal

performed also using a 256-QAM signal. The results from these simulations can be seen in figure 4.3b.

The first conclusion that can immediately be drawn by comparing figures 4.3a and 4.3b is that the effect of using more than one section to create the MP-PD improves the performance of the system drastically no matter what the order of the polynomials used. It can be seen that when only one section is used, a 2nd order polynomial creates a degradation of 2.79 dB at $TD_{\min}$, as opposed to a degradation of 0.46 dB when a two-section predistorter is used. The use of a two-section MP-PD provides a performance equal to that of an ideal predistorter when a polynomial of 3rd order or higher is used. On the other hand, even a 5th order MP-PD with one section is not able to achieve the performance of an ideal predistorter. The degradation of such a predistorter is around 0.1 dB worse than the degradation of an ideal predistorter. Even though this difference is very low, the fact remains that a 5th order MP-PD with one section has almost the same performance as a 2nd order MP-PD with two sections. Therefore the method of dividing the predistorter into more than one section has a significant advantage over the method of modeling the whole characteristics of the predistorter with one single polynomial.

It can be seen from figure 4.3b that with a 3rd order MP-PD, an ideal performance can be obtained by using two or more sections. Because a greater number of sections means that there are more polynomials to implement, in order to minimize implementation complexity, the predistorter with two sections will be considered as the preferred 3rd order MP-PD.

Nevertheless, a 3rd order is not the lowest order polynomial that can be used to model the MP-PD. Because of the sectioning ability of this predistortion technique, with enough sections, a high-performance 2nd order MP-PD could be created. Since such a predistorter would be implemented using only 2nd order polynomials, it would utilize less processing power to run. This would give it an edge over predistorters modeled using 3rd order polynomials or higher.

Figure 4.4 shows the performance of a 2nd order MP-PD with different sections. From this figure, it can be seen that such a predistorter will linearize the HPA in the same manner as an ideal predistorter when it is composed of three or

Figure 4.4: TD versus OBO for the MP-PD with $n=2$ at various values of m (results using a 256-QAM signal)

more sections. Again, to minimize implementation complexity, the predistorter with three sections will be considered to be the preferred 2nd order MP-PD.

For simplicity purposes, an n^{th} order MP-PD with m sections will be referred to using the acronym MP-PD- $[n,m]$. Therefore, it can be said that the MP-PD- $[2,3]$ and the MP-PD- $[3,2]$ both create highly linear predistorters for the TWTA. Even though the MP-PD- $[3,5]$ is more complex to implement, it also creates a highly linear system and its study is therefore of interest.

The polynomial coefficients for all these cases were designed to be as follows:

- MP-PD- $[2,3]$:

$$A(\rho) = \begin{cases} 0.1436\rho^2 + 0.4615\rho + 0.00205, & 0 \leq \rho < 0.6 \\ 1.4771\rho^2 - 1.2561\rho + 0.5610, & 0.6 \leq \rho < 0.9231 \\ 44.3602\rho^2 - 82.13125\rho + 38.7086, & 0.9231 < \rho \leq 1 \end{cases} \quad (4.3)$$

$$\Psi(\rho) = \begin{cases} -0.3119\rho^2 + 0.0167\rho - 0.00097, & 0 \leq \rho < 0.6 \\ -1.1348\rho^2 + 1.0782\rho - 0.3466, & 0.6 \leq \rho < 0.9231 \\ -24.4745\rho^2 + 45.0611\rho - 21.0773, & 0.9231 < \rho \leq 1 \end{cases} \quad (4.4)$$

- MP-PD- $[3,2]$:

$$A(\rho) = \begin{cases} 0.3546\rho^3 - 0.1786\rho^2 + 0.5389\rho - 0.00176 & 0 \leq \rho < 0.8 \\ 68.2148\rho^3 - 175.8725\rho^2 + 152.2789\rho - 43.7006 & 0.8 \leq \rho < 1 \end{cases} \quad (4.5)$$

$$\Psi(\rho) = \begin{cases} -0.2120\rho^3 - 0.1192\rho^2 - 0.0295\rho + 0.0013, & 0 \leq \rho < 0.8 \\ -37.4418\rho^3 + 96.10180\rho^2 - 83.0127\rho + 23.8711, & 0.8 \leq \rho < 1 \end{cases} \quad (4.6)$$

- MP-PD-[3,5]:

$$A(\rho) = \begin{cases} 0.1543\rho^3 - 0.0099\rho^2 + 0.5010\rho - 0.00002, & 0 \leq \rho < 0.3846 \\ 0.8188\rho^3 - 0.9548\rho^2 + 0.9555\rho - 0.0733, & 0.3846 \leq \rho < 0.6897 \\ 8.3302\rho^3 - 18.6404\rho^2 + 14.8770\rho - 3.7369, & 0.6897 \leq \rho < 0.8824 \\ 80.9755\rho^3 - 215.8684\rho^2 + 193.3911\rho - 57.6025, & 0.8824 \leq \rho < 0.9756 \\ 14025.51\rho^3 - 41350.48\rho^2 + 40641.03\rho - 13315.09, & 0.9756 \leq \rho < 1 \end{cases} \quad (4.7)$$

$$\Psi(\rho) = \begin{cases} -0.0573\rho^3 - 0.2467\rho^2 - 0.00134\rho + 0.00003, & 0 \leq \rho < 0.3846 \\ -0.5237\rho^3 + 0.4003\rho^2 - 0.3073\rho + 0.0488, & 0.3846 \leq \rho < 0.6897 \\ -4.9238\rho^3 + 10.7420\rho^2 - 8.4342\rho + 2.1841, & 0.6897 \leq \rho < 0.8824 \\ -44.7607\rho^3 + 118.8693\rho^2 - 106.2775\rho + 31.7007, & 0.8824 \leq \rho < 0.9756 \\ -7389.84\rho^3 + 21785.39\rho^2 - 21410.44\rho + 7014.38, & 0.9756 \leq \rho < 1 \end{cases} \quad (4.8)$$

These three MP-PDs will be the main focus of further study.

4.1.3 Study of the PSD

The power spectral density (PSD) of the amplified signal can also help evaluate the performance of a predistorter. By comparing the PSD of a signal before and after it is passed through a PD-HPA cascade, one can identify the amount by which the PSD has increased due to the amplification process. An efficient amplification would result in small changes of the PSD.

Figures 4.5 and 4.6 show the resultant PSD obtained using different predistorters. The spectrum resulting from four different MP-PDs was found. These four predistorters were namely the MP-PD-[2,3], the MP-PD-[3,2], the MP-

PD-[3,5] and the MP-PD-[5,1]; and the spectrum obtained from their use was compared to the PSD of the input signal and to the PSD created by an ideal predistorter. Figures 4.5a and 4.5b show the various PSDs found using a 64-QAM at an OBO of 6 dB and 4 dB respectively. Figures 4.6a and 4.6b were the result of using a 256-QAM signal at an OBO of 7 dB and 4.5 dB respectively. For all four cases, the rolloff of the SRRC filters was set to 0.25. The PSD was normalized to display a maximum power of 0 dB and the -3dB point occurs at one eighth of the maximum frequency displayed.

It can be seen from these figures that for all cases, the MP-PD-[3,5] creates a PSD which is almost identical to that of the ideal predistorter. This predistorter can be seen to have the best performance of the four MP-PDs. The worst performance is achieved by the MP-PD-[5,1] which has a PSD of around 14 dB higher than the MP-PD-[3,5] for the case in figure 4.5a, and of around 25 dB higher for the case in figure 4.6a. At high OBOs, such as in the cases depicted in a figure 4.5a and figure 4.6a, it is observed that the MP-PD-[3,2] obtains a lower PSD than the MP-PD-[2,3] by about 5 dB. When the system is set to operate at the OBO at which TD_{min} is achieved (figures 4.5b and 4.6b), all the predistorters create a similar PSD. Therefore, the difference which the predistorters create in the PSD of a signal can be mainly perceived at high OBOs.

It can be concluded that when using a particular polynomial order, the performance of a predistorter can be improved through the method of sectioning. As the number of sections increases, the performance of the MP-PD approaches the performance of an ideal predistorter.

(a) 64-QAM at an
OBO of 6 dB

(b) 64-QAM at an
OBO of 4 dB

Figure 4.5: PSD resulting from various MP-PDs on the TWTA
using a 64-QAM signal with $\beta=0.25$

(a) 256-QAM at an OBO of 7 dB

(b) 256-QAM at an OBO of 4.8 dB

Figure 4.6: PSD resulting from various MP-PDs on the TWTA using a 256-QAM signal with $\beta=0.25$

4.2 Performance of the SSPA

Since the modeled SSPA is more linear than the modeled TWTA, a low complexity predistorter will give better results on the SSPA than on the TWTA. Since the MP-PD-[2,3], the MP-PD-[3,2] and the MP-PD-[5,1] gave good results on the TWTA, these same predistorters will be used to linearize the SSPA. The polynomial coefficients for these predistorters were computed through the LS method to be the following:

- MP-PD-[2,3]:

$$A(\rho) = \begin{cases} 1.6879\rho^2 + 0.6930\rho + 0.00001, & 0 \leq \rho < 0.0156 \\ 14.7638\rho^2 + 0.2434\rho + 0.0039, & 0.0156 \leq \rho < 0.0281 \\ 170.8535\rho^2 - 8.6545\rho + 0.1310, & 0.0281 < \rho \leq 0.0338 \end{cases} \quad (4.9)$$

$$\Psi(\rho) = \begin{cases} -2032.01\rho^2 + 12.2012\rho - 0.0140, & 0 \leq \rho < 0.0156 \\ -3180.10\rho^2 + 57.0206\rho - 0.4495, & 0.0156 \leq \rho < 0.0281 \\ -66543.651\rho^2 + 3692.37\rho - 52.6641, & 0.0281 < \rho \leq 0.0338 \end{cases} \quad (4.10)$$

- MP-PD-[3,1]:

$$A(\rho) = \begin{cases} 808.0960\rho^3 - 27.6687\rho^2 + 0.9792\rho - 0.0006, & 0 \leq \rho < 0.0338 \end{cases} \quad (4.11)$$

$$\Psi(\rho) = \begin{cases} -222230\rho^3 + 6898.47\rho^2 - 81.3160\rho + 0.1812, & 0 \leq \rho < 0.0338 \end{cases} \quad (4.12)$$

- MP-PD-[3,2]:

$$A(\rho) = \begin{cases} 154.6676\rho^3 - 2.0003\rho^2 + 0.7161\rho - 0.00002, & 0 \leq \rho < 0.0225 \\ 8369.51\rho^3 - 623.6440\rho^2 + 16.4396\rho - 0.1327, & 0.0225 \leq \rho < 0.0338 \end{cases} \quad (4.11)$$

$$\Psi(\rho) = \begin{cases} -4718.42\rho^3 - 1908.68\rho^2 + 11.2632\rho - 0.0124, & 0 \leq \rho < 0.0225 \\ -3497113\rho^3 + 266232\rho^2 - 6843.21\rho + 58.3023, & 0.0225 \leq \rho < 0.0338 \end{cases} \quad (4.12)$$

- MP-PD-[5,1]:

$$A(\rho) = \begin{cases} 2747912\rho^5 - 185048\rho^4 + 4591.89\rho^3 \\ -46.645\rho^2 + 0.8845\rho - 0.0002, & 0 \leq \rho < 0.0338 \end{cases} \quad (4.13)$$

$$\Psi(\rho) = \begin{cases} -1254291302\rho^5 + 88080572\rho^4 - 2193718\rho^3 \\ + 20864.6\rho^2 - 77.80\rho + 0.0664, & 0 \leq \rho < 0.0338 \end{cases} \quad (4.14)$$

Performance measurements of the TD were taken using these four multipolynomial predistorters as well as the ideal predistorter. The performance of the SSPA without a predistorter was also calculated. To do this, various simulations were run using 64-QAM and 256-QAM signals, at various rolloff factors. The results from these simulations can be seen in figures 4.7, 4.8 and 4.9, where these figures represent the performance of the predistorters when the SSRC filter rolloffs are set to 0.5, 0.25 and 0.125 respectively using both 64-QAM and 256-QAM signals.

It can be seen from these results that the MP-PD-[2,3], the MP-PD-[3,2] and the MP-PD-[5,1] create the same effect on the SSPA. These three predistorters linearize the HPA in the same manner as the ideal predistorter does, both when amplifying 64-QAM and 256-QAM signals. On the other hand, the MP-PD-[3,1] creates a degradation of around 0.1 dB from ideal behaviour for 64-QAM signals and of around 0.3 dB for 256-QAM signals. It must be noted however, that when using 256-QAM signals, this predistorter exhibits an increase in the degradation as this is operated at higher OBOs.

(a) Using a 64-QAM signal

(b) Using a 256-QAM signal

Figure 4.7: Performance of various MP-PDs on the SSPA using 64 and 256-QAM signals with $\beta=0.5$

(a) Using a 64-QAM signal

(b) Using a 256-QAM signal

Figure 4.8: Performance of various MP-PDs on the SSPA using 64 and 256-QAM signals with $\beta=0.25$

(a) Using a 64-QAM signal

(b) Using a 256-QAM signal

Figure 4.9: Performance of various MP-PDs on the SSPA using 64 and 256-QAM signals with $\beta=0.125$

All the predistorters create a large improvement over the non-linearized SSPA. This improvement is of around 2.6 dB for 64-QAM signals and 4.2 dB of 256-QAM signals. $TD_{\min}$ is the same for rolloff factors of 0.5 and 0.25, the difference being that the smaller rolloff creates a larger degradation by 0.05 dB. However, a rolloff of 0.125 produces a higher $TD_{\min}$ by around 0.3 dB.

In order to view the effect of the predistorters on the spectrum of the system, plots of the PSD using 64-QAM and 256-QAM signals at a rolloff factor of 0.25 can be seen in figure 4.10. These measurements were taken for the various predistorters at an OBO of 6 dB for 64-QAM signals and 6.5 dB for 256-QAM signals. One can see from these results that the PSD is generally lowered when the predistorter is composed of more than one section. This effect is more noticeable when using 256-QAM signals. With 64-QAM signals however, it can be said that the MP-PD-[2,3], the MP-PD-[3,2] and the MP-PD-[5,1] all create almost the same PSD as the ideal predistorter. It is noticed that the PSDs obtained from the SSPA are generally higher than those obtained by the TWTA in figures 4.5 and 4.6. This is because the SSPA is not linearized at the saturation point, but far from it. OBO measurements on the other hand, are taken relative to the saturation point.

Various conclusions can therefore be drawn for the SSPA:

- As the complexity of the signal increases from 64-QAM to 256-QAM or higher, a low order/section predistorter will become less effective.

(a) 64-QAM at an
OBO of 6 dB

(b) 256-QAM at an
OBO of 6.5 dB

Figure 4.10: PSD resulting from various MP-PDs on the SSPA

- SSPAs can be adequately linearized with predistorters as low in complexity as a third-order with only one section, but provide better linearity if the number of sections or polynomial order is increased.
- A slight increase in the number of sections of the predistorter will generally provide better results than an increase in the polynomial order.

4.3 Comparison Based on the TWTA Model

In order to evaluate the performance of the MP-PD on the TWTA, a comparison with other analog predistortion schemes can be performed. Three common predistortion schemes have been chosen for this purpose. The first of these is a 5th order predistorter developed by Georges Karam and Hikmet Sari [14]. This technique is developed by initially modeling the AM/AM and AM/PM characteristics of the TWTA using a fifth order polynomial with complex coefficients. Using this model, a fifth order predistorter is then created to linearize the HPA. This fifth order predistorter will be referred to as the Karam and Sari predistorter (K&S-PD) and has the following form:

$$y = x[1 + (0.896 - j0.611)\rho^2 + (1.695 - j1.816)\rho^4] \quad (4.15)$$

Aldo D'Andrea, Vincenzo Lottici and Ruggero Reggiannini developed what they referred to as an amplitude and phase predistorter (A&P-PD) in [8] and [9]. A third-order and a fifth-order model of this predistorter will also be

used for the comparison. The A&P-PD, like the MP-PD, aims to linearize the amplitude characteristics of the TWTA for an input voltage between 0 V and 1 V by modeling the ideal characteristics of the predistorter. At the same time, it aims to eliminate all HPA phase insertions. The predistorter takes the form:

$$y = xA(\rho^2)\exp[-j\Psi(\rho^2)] \quad (4.16)$$

where $A(\rho^2)$ and $\Psi(\rho^2)$ are respectively the amplitude and phase characteristics of the predistorter and are modeled by polynomials. Polynomial coefficients are computed from a LS technique which uses an averaged probability density function over a symbol interval as part of the equation. This leads to the creation of different polynomials whose coefficients depend on the complexity of the signal constellation.

For a third-order A&P-PD, the following were the polynomials used on a 64-QAM constellation and a 256-QAM constellation respectively:

$$\begin{cases} A(\rho^2) = 0.2285\rho^2 + 0.4847 \\ \Psi(\rho^2) = 0.3422\rho^2 - 9.3 \times 10^{-3} \end{cases} \quad (4.17)$$

$$\begin{cases} A(\rho^2) = 0.2072\rho^2 + 0.4895 \\ \Psi(\rho^2) = 0.3284\rho^2 - 6.8 \times 10^{-3} \end{cases} \quad (4.18)$$

One should note that these are only second order polynomials. Nevertheless, they are multiplied by the input signal as expressed in (4.16), which causes the system to perform as a third order polynomial. The same logic applies to the 5th order polynomials for this predistorter.

For the fifth-order A&P-PD, the following were the models used on 64-QAM signals and 256-QAM signals:

$$\begin{cases} A(\rho^2) = 0.2813\rho^4 + 0.0207\rho^2 + 0.5078 \\ \Psi(\rho^2) = 0.1995\rho^4 + 0.2065\rho^2 + 3.2 \times 10^{-3} \end{cases} \quad (4.19)$$

$$\begin{cases} A(\rho^2) = 0.2087\rho^4 + 0.0653\rho^2 + 0.5039 \\ \Psi(\rho^2) = 0.1599\rho^4 + 0.2282\rho^2 + 1.7 \times 10^{-3} \end{cases} \quad (4.20)$$

4.3.1 Performance Assessment Based on MSE Computations

Three different MP-PDs were used for the comparison: The MP-PD-[2,3], the MP-PD-[3,2] and the MP-PD-[3,5]. A comparison of the performance of these three MP-PDs with the three more commonly used predistorters and an ideal predistorter can be seen in figures 4.11, 4.12 and 4.13. Each of these figures contains the TD curves of these 7 systems both using 64-QAM signals and 256-QAM signals at various SRRC filter rolloff factors.

The general results from these figures show that all three MP-PDs achieve the same performance as an ideal predistorter both for 64-QAM and 256-QAM systems. On the other hand, the 5th order A&P-PD obtains $TD_{\min}$ at an OBO of around 0.1 dB higher than that obtained with the ideal performance for both 64-QAM and 256-QAM systems. The 3rd order A&P-PD has the same $TD_{\min}$ as the 5th order A&P-PD for 64-QAM signals but creates a higher degradation whilst performing 0.2 dB worse for 256-QAM signals. As for the 5th order K&S-PD, its performance is inferior to the ideal predistorter's performance by 0.8 dB when

(a) Using a 64-QAM
signal

(b) Using a 256-QAM
signal

Figure 4.11: Performance of various MP-PDs on the TWTA
using 64 and 256-QAM signals with $\beta=0.5$

(a) Using a 64-QAM signal

(b) Using a 256-QAM signal

Figure 4.12: Performance of various MP-PDs on the TWTA using 64 and 256-QAM signals with $\beta=0.25$

(a) Using a 64-QAM signal

(b) Using a 256-QAM signal

Figure 4.13: Performance of various MP-PDs on the TWTA using 64 and 256-QAM signals with $\beta=0.125$

using 64-QAM signals and by 1.4 dB when using 256-QAM signals. On the whole, the ideal predistorter improves the performance of the non-linearized TWTA by around 6 dB when using 64-QAM signals and by around 8.1 dB using 256-QAM signals. The performance curves the non-linearized TWTA cannot be seen in the aforementioned figures due to graph resolution issues. Please consult appendix A for these results.

It can also be seen from these figures that very little difference in the results is seen when using a filter rolloff factor of 0.5 or 0.25. It is generally the case that when the rolloff factor is changed from 0.5 to 0.25, the amount of degradation increases slightly, but leaves $TD_{\min}$ unchanged. However, when the rolloff factor is changed to 0.125, one can generally see a degradation in most systems of around 0.1 dB when using 64-QAM signals and of 0.2 dB for 256-QAM signals.

4.3.2 Performance Assessment Based on BER Computations

Total degradation, however, is not generally measured by finding the MSE between the transmitted and received symbols. In many papers including [6], [7], [8], [9], [14], [16] and [17] the amount of TD for a HPA is measured by using the semi-analytic method described in [13]. This approach assumes that AWGN is introduced into the communication channel after the amplification process, as shown in figure 4.14. From this assumption a simulation can be run to analytically evaluate the probability of symbol error in the received signal; after

**Figure 4.14: Block Diagram of the Transmission System
for the Semi-Analytic Approach**

sampling the received signal, a decision device is used to find the most probable transmitted symbol. The probability of this device making an error is calculated as the conditional symbol error probability $\Pr\{\hat{c}_i \neq c_i | v_k\}$, where c_i and $\hat{c}_i$ are the transmitted and received symbols respectively and v_i is the array of samples at the input of the decision device. Assuming that Gray coding is used, the bit error rate for any symbol can be found as:

$$BER(v_k) = \frac{\Pr\{\hat{c}_i \neq c_i | v_k\}}{2 \log_2 N_Q} \quad (4.21)$$

By applying (4.21) to every received symbol and averaging the result, the BER of the system can be found at any SNR. Therefore, by finding the BER at various values of the SNR, a plot of the BER versus E_b / N_o can be drawn on a

point-by-point basis. An example of this is shown in figure 4.15 where the curve of a linear, distortion-free system is drawn, together with the curves resulting from the use of a MP-PD-[3,2] at an OBO of 4.6 dB and 4.2 dB. From such curves, the degradation (Δ) at any OBO can be found by subtracting the E_b/N_o necessary to create a BER of 10^{-4} at that OBO from the E_b/N_o necessary to create the same BER for a linear system:

$$\Delta_{@OBO} = (E_b/N_o)_{@OBO} \Big|_{BER=10^{-4}} - (E_b/N_o)_{Linear} \Big|_{BER=10^{-4}} \quad (4.22)$$

Figure 4.15: BER versus E_b/N_o for a 64-QAM Signal using the MP-PD-[3,2] at various OBOs

From knowledge of Δ , the total degradation can therefore be calculated.

To better understand the link between MSE measurements and the semi-analytic approach, one can look at the error signal, e , between the input and the output symbols. This can be found on a per symbol basis:

$$e_i = c_i - \hat{c}_i \quad (4.23)$$

To better visualize the link between the two methods, the error signal can be separated into its in-phase and quadrature components. For clarification purposes, the histograms of the in-phase and quadrature components of the error signal are shown in figure 4.16. These histograms were obtained from a simulation of the TWTA using the MP-PD-[3,2] at an OBO of 4.2 dB, where 64-QAM signals were employed with a filter rolloff of 0.25. These measurements have been taken by normalizing the distance between any two adjacent symbols to two units. From such histograms, one can measure both the MSE of the signal and the BER.

The MSE can be easily calculated from the information displayed in these graphs:

$$MSE = \sum F_i I^2 + \sum F_q Q^2 \quad (4.24)$$

The result obtained from this equation will be equal to the result obtained by using (3.32).

Figure 4.16: Histograms of the Error Signal

For any size QAM signal, the BER can be calculated at any SNR using (3.28). This relationship can be generally expressed as a function of the energy per bit over noise density. i.e.

$$BER = f\left(\frac{E_b}{N_o}\right) \quad (4.25)$$

It logically follows that the BER can also be calculated at any SNR using the semi-analytic approach from the information available in figure 4.16. This is done by applying the following relationship:

$$BER = \frac{F_i \cdot f\left(\frac{E_b}{N_o}(I+1)^2\right) + F_i \cdot f\left(\frac{E_b}{N_o}(1-I)^2\right) + F_q \cdot f\left(\frac{E_b}{N_o}(Q+1)^2\right) + F_q \cdot f\left(\frac{E_b}{N_o}(1-Q)^2\right)}{4k} \quad (4.25)$$

where k is the number of symbols transmitted. When the information displayed in the histograms is applied to (4.25), curve 3 in figure 4.15 will be recreated.

The two methods require different calculations; therefore the results are not expected to be the same. However, there is a clear link between the two, which means that the performance results obtained using the two methods should also display a link.

In order to validate the results obtained by using MSE measurements, some of the simulations were repeated using the semi-analytic approach. Figure 4.17 shows the curves obtained from using this method on the TWTA. This figure contains the same information as figure 4.12, since the same predistorters were simulated using a filter rolloff of 0.25.

(a) Using a 64-QAM
signal

(b) Using a 256-QAM
signal

Figure 4.17: Semi-Analytic Analysis of various MP-PDs on the TWTA
using 64 and 256-QAM signals with $\beta=0.25$

Using this method, $TD_{\min}$ is achieved by the ideal predistorter at an OBO of 4.8 dB and 5.5 dB for 64-QAM and 256-QAM signals respectively. This is consistent with the results obtained by in [7]. $TD_{\min}$ for the 3rd order A&P-PD occurs at 5.1 dB and 6.1 dB for 64-QAM and 256-QAM signals. Using the 5th order A&P-PD, $TD_{\min}$ is found at 5 dB and 5.6 dB. These results are in line with those presented in [9]. As for the K&S-PD, $TD_{\min}$ is found at an OBO of 6.1 dB and 7 dB, which are also consistent with the results presented in [14].

By comparing figures 4.12 and 4.17 on a qualitative basis, it can be seen that by using MSE measurements, one can find the amount of total degradation of a HPA. However, results of $TD_{\min}$ might not necessarily be the same as when using the semi-analytic approach. The semi-analytic method will give better estimates of the actual amount of degradation since it assumes an AWGN channel in its computations, but will require longer simulation runs as well as a more complex simulation code.

4.3.3 Performance Assessment Based on the PSD

To better understand the linearization effectiveness of the various predistorters, the PSD of a 64-QAM signal at an OBO of 6.5 dB and a 256-QAM signal at an OBO of 7 dB, using these predistorters, can be seen in figure 4.18. These spectrum representations were taken using a filter rolloff factor of 0.25. From these figures it can be seen that the lowest PSDs are created when using the

(a) 64-QAM at an OBO of 6.5 dB

(b) 256-QAM at an OBO of 7 dB

Figure 4.18: PSD resulting from various MP-PDs on the TWTA with $\beta=0.25$

MP-PDs, the best of which is the MP-PD-[3,5] which creates a spectrum close to that of the ideal predistorter. By looking at the figures around a normalized frequency of 0.2, one can see that when using a 256-QAM signal, the 5th order A&P-PD create the same PSD as the MP-PD-[2,3], but a higher PSD by about 6 dB when using a 64-QAM signal. Also, a 5th order A&P-PD creates a lower PSD than the 3rd order A&P-PD by around 4 dB when using 64-QAM signals and about 8 dB when using 256-QAM signals. The K&S-PD creates a PSD very similar to that of the 3rd order A&P-PD which is lower than the PSD of a non-linearized HPA by about 15 dB.

From all the above results, various conclusions can be drawn about the various predistorters used on the TWTA:

- Predistorters like the A&P-PD or the MP-PD, which try to model the ideal predistorter characteristics, are more effective than predistorters like the K&S-PD that are created based on models of the HPA's characteristics.
- The MP-PD can be adjusted to perform better than single polynomial predistorters such as the 5th order and 3rd order A&P-PD or the 5th order K&S-PD, without having to use high order polynomials.
- By increasing the number of sections of a MP-PD, the PSD can be lowered. With enough sections, the MP-PD can obtain the same performance as an ideal predistorter with even a 2nd order polynomial.
- As the complexity of the AM/AM and AM/PM characteristics of a HPA increase, more sections can be added to the MP-PD to allow for a more linear PD-HPA cascade. Because this does not increase the polynomial order, the complexity of the predistorter is not greatly affected.

- Even though SSPAs are generally more linear than TWTAs, they might not always have a lower $TD_{\min}$ due to the fact that SSPAs are normally not linearized at their saturation point.

4.4 Effect of the Sampling Ratio

The multipolynomial predistorter is designed to be implemented as an analog predistortion device. For this reason, signal sampling and digital signal processing concerns are generally not an issue. However, all tests and analysis of this predistorter were done using computer aided simulation runs. This meant that analog signals had to be sampled and modeled by discrete components. For all simulations, a standard sampling rate of 8 samples per QAM symbol was used. This allowed the analog waveform to be simulated with enough accuracy.

If the multipolynomial predistorter was to be implemented using DSP technology, such a sampling rate would require the use of extensive signal processing power. Consequently, a study of the effect of smaller sampling rates on the predistorter's performance is in place. To carry out such a study, the adjustments shown in figure 4.19 were done to the predistorter. To simulate the use of a lower sampling rate system, the input signal to the predistorter was downsampled from a rate of 8 samples per symbol to the desired rate. After the predistortion process, the signal was upsampled back to the initial sampling rate in order to emulate the signal's change from discrete to analog.

Figure 4.19: Block Diagram of the Sampling Process

The results from this process using 64-QAM and 256-QAM signals can be seen in figures 4.20 (a) and (b) respectively. TD measurements were done using sampling rates (n/R) of 2, 2.5, 3 and 4 samples per symbol. From there results one can see that using 2.5, 3 or 4 samples per symbol gives almost the same outcome. For more accurate results, a plot of $TD_{\min}$ for each sampling rate can be seen in figure 4.21. This figure clearly shows that very small changes occur when the sampling rate of the input signal to the predistorter is reduced to as little as 2.5 samples per symbol. The greatest loss of accuracy occurs when the sampling rate is set to 2 samples per symbol. Such a change creates an increase in the degradation of about 0.25 dB when using a 64-QAM signal and of 0.9 dB when using a 256-QAM signal. It can therefore be concluded from these results that a sampling rate of 2.5 samples per symbol or higher yields satisfactorily accurate linearization effect.

To validate and verify these results, the same simulations were run using the semi-analytic method to measure the total degradation of the system. The results from these simulations can be seen in figure 4.22 where (a) displays the results for a 64-QAM signal and (b) shows the results for a 256-QAM signal.

(a) Using a 64-QAM signal

(b) Using a 256-QAM signal

Figure 4.20: TD Results for various Sampling Rates

Figure 4.21: Minimum TD for various Sampling Rates

Qualitatively, these results display the same information as the curves in figure 4.20. The amount of degradation is seen to change due to the methodology used to calculate the total degradation. This causes $TD_{\min}$ to change. However, the general information obtained from both figures 4.20 and 4.22 is the same.

Figure 4.23 gives a better measurement of $TD_{\min}$ for the various sampling rates. These measurements are very similar to those in figure 4.21. The obtained results agree with the results by D'Andrea et al. in [7], and confirm that a sampling rate of 2.5 or higher can be used on the transmitted signal with minimal increase in the distortion.

(a) Using a 64-QAM signal

(b) Using a 256-QAM signal

Figure 4.22: TD Results for various Sampling Rates
using the Semi-Analytic Approach

Figure 4.23: Minimum TD for various Sampling Rates using the Semi-Analytic Approach

Therefore the results obtained through the semi-analytic method are again confirmed to have a direct link with results from the MSE method. Small increases in the degradation of the signal are better detected using the semi-analytic approach. This is reflected in the slope of the TD curves, where curves obtained through semi-analytic calculations have a higher slope than those obtained through MSE measurement. Due to this fact, that no differences are seen when the rolloff factor of the SRRC filters changes from 0.5 to 0.25. The amount of signal degradation due to such changes become noticeable when using the semi-analytic method as was shown in [14].

CONCLUSION AND RECOMMENDATIONS

This thesis analyzes the performance of a new predistortion technique that uses multiple polynomials to model the characteristics of a HPA. The conclusions drawn from this work are given here. Also, recommendations for future areas of research are stated.

5.1 Thesis Conclusions

HPA nonlinearities need to be combated. RF envelope predistortion is an effective means of accomplishing this task and for this reason, it is commonly used. It is through the use of polynomials that most predistorters are modeled. In the past, these polynomials would be used to model the characteristics of a HPA, and through such a model, a predistorted would be derived. In more recent times however, the ideal characteristics of a predistorter would first be extracted, and then polynomials would be used to model these. This method proved to be more effective.

In this thesis, a new technique to model the ideal characteristics of a predistorter is presented. This technique divides the characteristics of the ideal predistorter of a HPA into multiple sections, and models each section independently. The outcome is a predistorter composed of multiple polynomials. This is named the multipolynomial predistorter (MP-PD).

This predistortion technique is extensively explained and examined in chapter 3. Two types of HPAs are used to carry out these tests: a SSPA and a TWTA. The characteristics of these HPAs, as well as the characteristics of their ideal predistorters are described. A simulation model is created to test the performance of this predistorter, as well as others in literature. An analysis of the simulation process is given and a test is carried out to ensure that it is functioning correctly.

Various simulations are carried out to thoroughly evaluate the performance of the multipolynomial predistorter. These simulations aim at giving a better understanding of this new technique. The following issues are studied:

- How a change in the number of sections affects the performance.
- How changing the polynomial order affects the performance.
- The effect of different OBOs on the PSD of a linearized signal.
- How the MP-PD performs on a SSPA.
- The performance of the MP-PD compared to other predistorters in literature.

- The relationship between evaluating performance using MSE measurements and BER measurements.
- The performance of the predistorter when the input signal was sampled at various rates.

The results from these simulations are detailed in chapter 4. From these results, various conclusions can be drawn:

- The number of sections in the multipolynomial predistorter can be adjusted to best suit the HPA's characteristics without greatly affecting the implementation complexity of the predistorter. HPAs with complex characteristics can be effectively linearized with a low order polynomial provided enough sections are used.
- As the polynomial order used to create the predistorter increases, the linearizer's abilities also increase, along with its implementation complexity.
- The multipolynomial predistorter provides a powerful predistortion technique that enables users to specify and adjust the number of sections together with the polynomial order to suit the specific requirements needed to linearize any HPA. By adjusting these two parameters one can obtain a highly effective linearizer which can be implemented with relatively low complexity.
- As the OBO decreases, the out-of-band spectral emissions increase. At very low OBOs, all predistorters produce the same amount of out-of-band emissions.

-
- It was found that the multipolynomial predistorter would obtain better linearization results on a SSPA than on a TWTA when using low order polynomials or fewer sections. This is consistent with theory, which states that SSPAs are generally more linear HPAs than TWTAs. However, because SSPAs are not linearized at their saturation point, degradation and PSD measurements indicated the TWTA performs better than the SSPA. This is due to the fact that the OBO measurements are taken relative to the saturation point.
 - It was observed that various combinations of polynomial order and number of sections could be found to emulate the performance of an ideal predistorter. These attain such a performance using 2nd and 3rd order polynomials are proved to be more effective than other 5th order predistorters found in literature.
 - Because the MP-PD can be implemented using lower order polynomials than other predistorters found in literature, implementation complexity is reduced.
 - A predistorter's performance can be calculated using MSE measurements. Results from such measurements are found to be consistent with results obtained from BER measurements, which are more commonly used in literature.
 - A study of the effect of sampling the input signal to the predistorter at various rates revealed that using 2.5 samples per symbol or higher would ensure that the amplified signal had minimum distortion. Such a study is useful if the predistorter was to be implemented using DSP technology.

Because high order polynomials can become ineffective models and are more complex to implement, the use of low order polynomials is desirable. The multipolynomial technique allows users to create a predistorter using polynomials as simple as a 2nd order. Various polynomials are used to model the predistorter creating multiple low order polynomials. However, because only one polynomial is used at any one instant, computations become less tedious than having to use a high order polynomial. The advantages of using multiple low order polynomials over one higher order polynomial are numerous. Not only implementation complexity is reduced, but it is also simpler to curve-fit characteristics using lower order polynomials. However, because an increase in the number of sections means that there are more coefficients that need to be found and updated, the technique required to create an adaptive multipolynomial predistorter will have to be more complex.

5.2 Recommendations for Future Work

In this work, common LS curve-fitting technique is used to find the coefficients of the various polynomials that create the multipolynomial predistorter. However, other methods of finding these coefficients such as the LS technique described by D'Andrea et al. in [8] and [9] or the minimum distortion power polynomial model described by Lai and Bar Ness in [17] can be a more effective ways of finding these coefficients. A further look into applying these

techniques to find the coefficients of the multipolynomial predistorter could lead to the creation of a more effective linearizer.

Furthermore, an in-depth study of the implementation complexity required to construct the various forms of the multipolynomial predistorter could be performed. This could possibly lead to finding a relationship for the trade off between implementation complexity and the linearizer's effectiveness.

Finally, because all HPAs are sensitive to drifting and temperature variations, an adaptive multipolynomial predistorter would further improve the linearity of a HPA. Therefore, a study on the creation of an adaptive system for this predistorter is strongly suggested.

Appendix A

NONLINEAR TWTA

Due to graph resolution issues, the performance (using MSE measurements) of a non-linearized TWTA was not displayed in the figures in chapter 4. These results can be seen in figure A.1, where figure (a) shows the results for the HPA when using a 64-QAM signal and (b) shows the results when using a 256-QAM signal. Three curves can be observed for each case. These curves result from the SRRC filter rolloff being set to 0.5, 0.25 and 0.125.

It can be seen from these results that changes of the filter rolloff factor between 0.125 and 0.5 has very little effect on the performance of the nonlinear TWTA. This is mainly due to the fact that a large amount of distortion already exists due to the nonlinearity of the amplifier. Any further distortion due to filter rolloff would not produce large changes in the MSE of the signal. It can be seen that a 64-QAM signal creates a $TD_{\min}$ of around 10 dB while $TD_{\min}$ for a 256-QAM signal occurs around 13 dB.

(a) Using a 64-QAM signal

(b) Using a 256-QAM signal

**Figure A.1: Performance of the Non-Linear TWTA
using 64-QAM and 256-QAM Signals**

References

- [1] Agilent, Application Note AN 1287-1, "Understanding the Fundamental Principles of Vector Network Analysis."
- [2] Agilent, Application Note AN 1287-6, "Using a Network Analyzer to Characterize High-Power Components."
- [3] CAMPBELL, C. F. and BROWN, S. A., "Application of the Unequal Two-Tone Methods for AM-AM and AM-PM characterization of MMIC Power Amplifiers," *2001 IEEE Emerging Technologies Symposium on BroadBand Communications for the Internet Era Symposium Digest*, 2001, pp. 103 –106.
- [4] CAVERS, J. K., "Amplifier Linearization Using a Digital Predistorter with Fast Adaptation and Low memory Requirements," *IEEE Transactions on Vehicular Technology*, vol. 39, no.4, November 1990, pp. 374-382.
- [5] CRIPPS, S. C., *RF Power Amplifiers for Wireless Communications*, Massachusetts: Artech House, 1999.

-
- [6] D'ANDREA, A. N., LOTTICI, V., and REGGIANNINI, R., "A Digital Approach to Efficient RF Power Amplifier Linearization," in *IEEE Global Telecommunications Conference*, vol. 1, November 1997, pp. 77-81.
 - [7] D'ANDREA, A. N., LOTTICI, V., and REGGIANNINI, R., "Efficient Digital Predistortion in Radio Relay Links with Nonlinear Power Amplifiers," *IEEE Proceedings-Communications*, vol. 147, no. 3, June 2000, pp. 175-179.
 - [8] D'ANDREA, A. N., LOTTICI, V., and REGGIANNINI, R., "Parallel Amplitude and Phase Predistortion for RF Amplifier Linearization," in *IEEE Global Telecommunications Conference*, vol. 1, November 1995, pp. 198-202.
 - [9] D'ANDREA, A. N., LOTTICI, V., and REGGIANNINI, R., "RF Power Amplifier Linearization through Amplitude and Phase Predistortion," *IEEE Transactions on Communications*, vol. 44, no. 11, November 1996, pp. 1477-1484.
 - [10] GHANNOUCHI, F. M., WAKANA, H., and TANAKA, M., "A New Unequal Tree-Tone Signal Method for AM-AM and AM-PM Distortion Measurements Suitable for Characterization of Satellite Communication Transmitters/Transponders," *IEEE Transactions on Microwave Theory and Techniques*, vol. 88, August 2000, pp. 1404-1407.
 - [11] GHORBANI, A., and SHEIKHAN, M., "The Effect of Solid State Power Amplifiers (SSPAs) Nonlinearities on MPSK and M-QAM Signal

- Transmission," in *Sixth International Conference on Digital Processing of Signals in Communications*, 2-6 September 1991, pp. 193-197.
- [12] HEUTMAKER, M. S., WELCH, J. R., and WU, E., "Using Digital Modulation to Measure and Model RF Amplifier Distortion," *Applied Microwave and Wireless*, March/April 1997, pp. 34-39.
- [13] JERUCHIM, M. C., "Techniques for Estimating the Bit Error Rate in the Simulation of Digital Communication Systems," *IEEE Journal on Selected Areas in Communications*, vol. SAC-2, no. 1, January 1984, pp. 153-170.
- [14] KARAM, G., and SARI, H., "Analysis of Predistortion, Equalization, and ISI Cancellation Techniques in Digital Radio Systems with Nonlinear Transmit Amplifier," *IEEE Transactions on Communications*, vol. 37, no. 12, December 1989, pp. 1245-1253.
- [15] KATZ, A., "Linearization: Reducing Distortion in Power Amplifiers," *IEEE Microwave Magazine*, vol. 2, no. 4, December 2001, pp. 37-49.
- [16] LAI, H., and BAR-NESS, Y., "A New Predistorter Design for Nonlinear Power Amplifiers Using the Minimum Distortion Power Polynomial Model (MDP-PM)" in *54th Vehicular Technology Conference*, vol. 4, 2001, pp. 2216-2220.

-
- [17] LAI, H., and BAR-NESS, Y., "Minimum Distortion Power Polynomial Model (MDP-PM) of Nonlinear Power Amplifiers and its Application on Analog Predistorters," in *50th Vehicular Technology Conference*, vol. 3, 1999, pp. 1501-1505.
- [18] LAICO, J. P., MCDOWELL, H. L., and MOSTER, C. R., "A Medium Power Travelling-Wave Tube for 6,000-Mc Radio Relay," *The Bell Systems Technical Journal*, November 1956, pp. 1285-1346.
- [19] MANN, S., BEACH, M., WARR, P., and MCGEEHAN, J., "Increasing the Talk-Time of Mobile Radios with Efficient Linear Transmitter Architectures," *Electronics and Communications Engineering Journal*, April 2001, pp. 65-76.
- [20] PERLOW, S. M., and SARNOFF, D., "Basic Facts about Distortion and Gain Saturation," *Applied Microwaves*, May 1989, pp. 107-117.
- [21] PROAKIS, J. G., *Digital Communications*, New York: Mc-Graw-Hill, 2001.
- [22] PUPOLIN, S., and GREENSTEIN, L. J., "Performance Analysis of Digital Radio Links with Nonlinear Transmit Amplifiers," *IEEE Journal on Selected Areas in Communications*, vol. SAC-5, no. 3, April 1987, pp. 534-546.
- [23] RAAB, F. H., ASBECK, P., CRIPPS, S., KENINGTON, P. B., POPOVIĆ, Z. B., POTHECARY, N., SEVIC, J. F., and SOKAL, N. O., "Power Amplifiers and

-
- Transmitters for RF and Microwave," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 3, March 2002, pp. 814-826.
- [24] SALEH, A. A. M., "Frequency-Independent and Frequency-Dependent Nonlinear Models of TWT Amplifiers," *IEEE Transactions on Communications*, vol. COM-29, no. 11, November 1981, pp. 1715-1720.
- [25] STAPLETON, S. P., "Amplifier Linearization Using Adaptive RF Predistortion," *Applied Microwave and Wireless*, March 2001, pp. 40-46.
- [26] YI, S. J., NAM, S., OH, S. H., and HAN, J. H., "Prediction of a CDMA Output Spectrum Based on Intermodulation Products of Two-Tone Test," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 5, May 2001, pp. 938-946.