DIGITAL PREDISTORTION FOR NONLINEAR RF POWER AMPLIFIERS IN WIDEBAND COMMUNICATION SYSTEMS

by

Ibrahim Fatungase

Submitted in partial fulfillment of the requirements
for the degree of Master of Applied Science

at

Dalhousie University
Halifax, Nova Scotia
August, 2023

To my family, whose unwavering support has been my constant strength and inspiration throughout this journey.

Table of Contents

List of Tables

List of Figures

Abstract

This thesis provides a comprehensive study of linearity enhancement in Radio Frequency (RF) Power Amplifiers (PA) while transmitting wideband signals with a high Peak Average Power Ratio (PAPR). With the advancement of digital modulation techniques, signals with a highly variable envelope create challenges for the linearity of a transmitter's exciter electronics. The main focus of this work is on modeling existing Power Amplifier designs and improving their linearity to accommodate these advanced modulation techniques without violating spectrum emission requirements. A key aspect of this research involves the development and implementation of a memory polynomial pre-distortion method for the mitigation of memoryless nonlinearity, which is demonstrated to be effective through computational simulations and hardware implementation. The thesis also acknowledges existing challenges in addressing memory distortion, as it identifies specific areas where current methodologies fall short, underscoring the potential for continued research and innovation in this domain. Overall, this research holds significant implications for improving power efficiency and spectral efficiency in RF communication systems, which ultimately contributes to the advancement of wireless transmission technologies.

List of Abbreviations Used

ACLR Adjacent Channel Leakage Ratio

AM Amplitude Modulation

AM/AM Amplitude-to-Amplitude

AM/PM Amplitude-to-Phase

API Application Programming Interface

dB decibel

DC Direct Current

DPD Digital Pre-Distortion

FCC Federal Communications Commission

FIR Finite Impulse Response

FM Frequency Modulation

FPGA Field Programmable Gate Arrays

IBOC In-Band on-Channel

IMD Intermodulation Distortion

IMD3 Third Order Intermodulation Distortion

IP3 Third Order Intercept Point

ISED Innovation, Science and Economic Development

LMS Least Mean Squares

LO Local Oscillator

LS Least Square

LTI Linear Time-Invariant

LUT Look-Up Table

MSE Mean Squared Error

OFDM Orthogonal Frequency-Division Multiplexing

PA Power Amplifier

PAPR Peak to Average Power Ratio

RF Radio Frequency

RFPA Radio Frequency Power Amplifier

SDR Software Defined Radio

SGD Stochastic Gradient Descent

UHD USRP Hardware Driver

USRP Universal Software Radio Peripheral

VHF Very High Frequency

Acknowledgements

I would like to extend my heartfelt gratitude to my supervisor, Jean-Francois Bousquet, for his invaluable guidance, support, and patience throughout this project. His expertise and dedication have been instrumental in shaping my research. Equally, my profound thanks goes to Philipp Schmid, whose insightful feedback and deep subject knowledge greatly refined my work. Their remarkable involvement, guiding me every step of the way, is a testament to their commitment to my growth and learning. Their influence on this work cannot be overstated, and I am truly grateful for their guidance and support.

Chapter 1

Introduction

In Radio Frequency (RF) communication systems, data and information may be transmitted via modulated electromagnetic waves through the air/space in wireless applications or through a physical medium such as an electrical conductor to transmit signals. With the digitization of information, advanced modulation techniques have been defined to increase the throughput available over a transmission medium. However, these complex techniques create signals with a highly variable envelope, which burdens the linearity of the transmitter's front-end power electronics. The broadcaster must have high power efficiency to transmit over large distances at a low cost. For this purpose, state-of-the-art Power Amplifier (PA)s must now include linearization features that correct for the signal distortion before it is transmitted, ensuring the transmitter operates in favorable conditions. This work investigated and improved the linearity of current RF amplifier designs. The findings from this work generally apply to communication systems that transmit wide-band signals.

1.1 Background

Historically, the Frequency Modulation (FM) band spans 87.5 MHz to 108 MHz for sound broadcasts. However, with technological advances to produce high-quality digital information, there is an interest in offering a high throughput broadcast link on the spectrum over the existing standard. For this purpose, spectrally efficient modulation techniques have been proposed [30, 33, 6]. Orthogonal Frequency-Division Multiplexing (OFDM) which organizes the spectrum in sub-carriers, has also been proposed in this frequency band, for example, in [16]. However, these modern digital modulation techniques produce a signal with a high Peak to Average Power Ratio (PAPR) compared to the simple FM technique based on a single signal carrier in everyday use today.

In RF communication systems, the broadcast links are often required to cover several 100 km. A transmitter's signal exciter must be designed carefully to support high signal power peaks while optimizing power-added efficiency. Using FM modulation, nonlinear circuit topologies can be considered since the signal comprises a

single carrier with a constant power envelope. Any non-linearity is manifested as harmonic content that can be removed using a harmonic filter in hardware. However, for modern digital modulation techniques, such as OFDM, the PAPR is high due to many added carriers that create in-band intermodulation products that violate spectrum emission requirements in a non-linear amplifier. This places a burden on the amplifier. More sophisticated PA designs must rely on advanced system architectures and use Digital Pre-Distortion (DPD) algorithms and frequency domain equalization. Using DPD, the signal produced at the output of the PA has improved inter-modulation distortion, and Adjacent Channel Leakage Ratio (ACLR). Since the required signal spectrum occupies a large bandwidth, the distortion includes memory effects, and digital filters that take this into account must be designed [1, 25, 2].

While traditional PA circuits can provide suitable power efficiency when applied to single tones as used by FM, broadband signals with high PAPR's, require additional power overhead above the average power level. To minimize the operation of the PA in backoff, it is possible to apply PAPR reduction using pre-distortion algorithms to shape the signal and compensate for the effect of nonlinearity. Various pre-distortion architectures are available [14, 20, 7, 3, 34, 24], and can be optimized depending on the type of nonlinearity experienced by the PA. In this work, the nonlinearity is expected to include memory effects, which must be modeled using advanced mathematical models like the Volterra series. The proposed technique will dynamically adjust the DPD by sensing the PAs output for the signal of interest and adjusting the DPD coefficients accordingly. This will thus improve linearity of the signal and will effectively allow it to transmit at higher power while respecting the standard spectral mask in the FM band. This can be useful for new applications, such as vehicular communications and additional digital radio channels broadcast.

1.2 Objectives & Contributions

The study emphasizes the development of a linearization technique aimed at augmenting transmission power while curbing in-band nonlinearities and reducing power transferred to adjacent bands [30, 29].

Objectives set forth involve modeling Radio Frequency Power Amplifier (RFPA)'s when processing wideband signals encompassing both the conventional FM voice signal and an in-band OFDM symbol integrated into the sidebands of the analog FM signal. The aspiration is to devise a linearization technique that enables the transmission of signals with highly variable envelopes. Linearization strategies commonly

applied in the industry [29], as well as advanced techniques considering nonlinear correction [28, 24, 34, 3, 7], will be evaluated. The hardware constraints, mainly the processor's maximum sampling frequency, will be considered [20, 14].

The project concludes with testing the digital pre-distortion algorithm on existing PAs to improve transmitter linearity and facilitate increased transmission power. Behavioral modeling of the system will be applied to account for distortion with memory. Using models from previous projects and existing models with memory, such as the Volterra series, the behavior of PAs and the influence of varying settings on the model can be assessed.

The project will leverage models like the Wiener and Hammerstein models [22], which provide a unique way to capture the frequency dependent behavior of amplifiers. Combined with measurements, these models will help define a realistic procedure for extracting a model of the PA[10].

Subsequently, the study involves an analysis of pre-correction algorithms and the development of a system-level amplifier model, including this pre-correction algorithm. The model is expected to provide superior performance while considering hardware complexity by incorporating adaptive filters to attain the optimum predistortion weights.

The final part of the study will see the full PA realized, which includes interfacing with the amplifier core, sensing the output voltage, and implementing the pre-correction technique to observe the real-world effects. This procedure evaluates linearity enhancement and transmits power at the PA's output.

1.3 Thesis Organization

The rest of this thesis is organized as follows: Chapter Two discusses the challenges in RF communication related to PAs, encompassing signal transmission, system complexities, and the implications of memoryless nonlinearity and memory distortion. In Chapter Three, PA modeling is presented, featuring a detailed analysis of nonlinearity as a cascade of distortion properties and the usage of Reduced Volterra Models. Chapter Four investigates PA distortion reduction methods, focusing on equalizing a linear filter and amplifier distortion compensation. Chapter Five elaborates on the software-defined radio system implementation for transmitting a wide-band signal in the Very High Frequency (VHF) Band, supplemented by experimental results. The final chapter, Chapter Six, summarizes the findings and provides insights into potential future work.

Chapter 2

Power Amplifier Challenges for RF Communication

This chapter reviews state-of-the-art waveforms for high throughput applications, specifically in Section 2.1, where the transmit signal characteristics are understood. Then in Section 2.2, the transmit front end/exciter is described. Next, in Section 2.3, the amplifier impairments are defined for a narrow-band signal. From Section 2.4 onward, the frequency dependence of an amplifier is introduced, and modes of quantification of distortion are discussed.

2.1 Transmission of RF Communication Signals

First, in Section 2.1.1, how the spectrum is typically allocated and shared between multiple users is described; then, in Section 2.1.2, a wideband signal is defined; finally, in Section 2.1.3, the high PAPR characteristics of specific modulation techniques are reviewed.

2.1.1 Spectrum Allocation

Reliable communication over an RF spectrum requires a complex set of processes and techniques. This includes the digital modulation of the signal data onto a carrier frequency. Theoretically, a signal can be modulated onto any frequency on the electromagnetic spectrum. However, the choice of frequency depends on several factors, including frequency band availability, the nature of the signal, and the characteristics of the communication medium.

Firstly, frequency bands possess varying properties and are regulated and reserved by different governing bodies depending on the region. Certain frequency bands are booked for specific applications, such as radio and television broadcasts, cellular and satellite communication, etc. In most cases, the use of channels within these bands requires a license.

The choice of frequency also depends on the nature of the transmitted signal. Different signals have varying requirements based on their information content and transmission needs. For instance, a video signal, which inherently carries more data

due to its combination of audio and visual elements, requires more bandwidth than an audio signal with a good codec [5].

Moreover, the frequency of modulation in wireless communication is significantly influenced by the environmental condition in which the communication occurs. Environmental factors, such as indoor, urban, rural, satellite, and marine settings, shape the unique channel characteristics. For instance, these conditions can determine the noise, interference, or attenuation experienced during transmission. Different environmental settings might emphasize specific characteristics over others due to the unique physical properties and challenges they present [17].

To maximize the capacity of communication systems, reserved frequency bands are typically portioned into several channels. While transmitting on an isolated channel, the signal must remain within its channel boundary. This minimizes the risk of interference with adjacent channels that may also be in use. The portioning of frequency bands also allows for the frequency reuse of a specific channel. The predetermined bandwidth used for all the channels in a reserved frequency band is also determined by the application of the communication system. Therefore, the carrier signal capacity is limited to its corresponding channel bandwidth.

2.1.2 Wideband Signal Transmission

This section describes the impact of transmitting a signal over a wide bandwidth. Narrowband and wideband systems are fundamental concepts in wireless communications. In a narrowband system, the signal occupies a small fraction of the channel bandwidth. This means that the signal's frequency range is relatively narrow compared to its carrier frequency, and it does not significantly change over the course of transmission. In contrast, a wideband signal is one where the signal's bandwidth takes up a substantial fraction of its carrier channel bandwidth. With a broader frequency range, the wideband signal can carry more information, but it may also face more interference and require more complex transmission techniques [26].

This thesis researches the effects of a channel's frequency response on a wideband signal, primarily due to the significant amount of data a wideband signal can carry. However, as the bandwidth of these signals increases, so do the challenges associated with maintaining signal integrity amidst potential interference and the complexity of transmission techniques. Increasingly intricate modulation techniques are required to accommodate rising throughput demands in response to the evolution of technology and communication standards. One such technique is OFDM,

extensively used in digital communication. OFDM operates by transmitting data signals on closely spaced sub-carriers [4]. This strategic design allows more data to be represented using significantly less bandwidth, thereby improving efficiency. A visualization of the OFDM spectrum is provided in Figure 2.1. This technique finds applications in various technologies, including TV, WIFI, digital radio, and cellular data transmission.

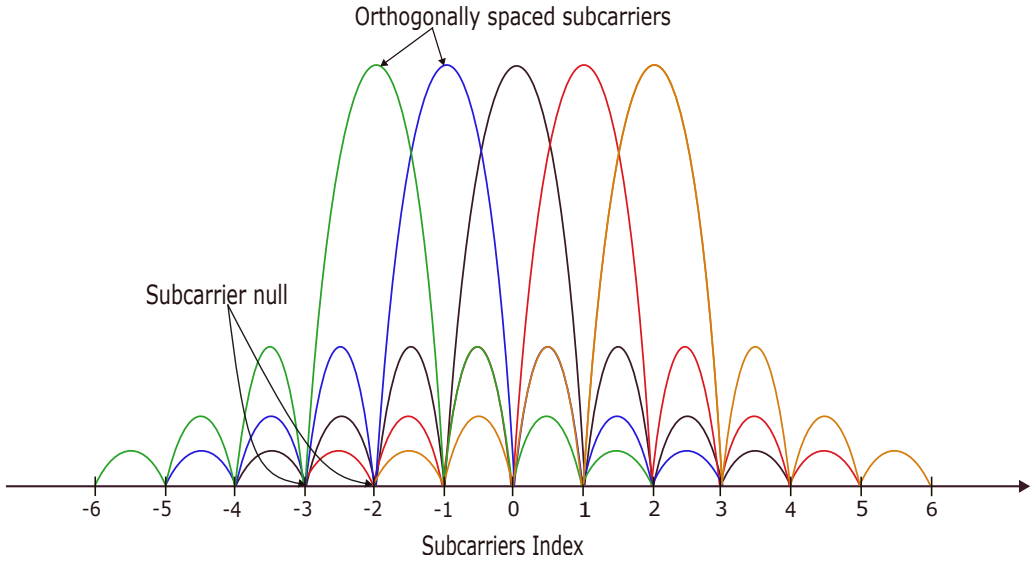


Figure 2.1: OFDM signal spectrum.

OFDM extend in applicability to even more complex systems. As analog radio broadcast slowly phases out, including digital signals for radio broadcasting becomes increasingly significant. Digital signals are more easily stored and compressed while transmitting than analog signals. The In-Band on-Channel (IBOC) system transmits both signals simultaneously as a single unit without the requirement for new spectrum allocations [30]. The analog radio signal may be modulated using analog modulation techniques, such as Amplitude Modulation (AM) or FM. In contrast, the digital portion of the radio signal may be modulated using a more complex technique such as OFDM. This hybrid method transmits a signal's digital and analog portions using the same carrier frequency. The OFDM subcarriers use additional sidebands outside the analog signal's band.

As illustrated in Figure 2.2, the spectrum of the IBOC signal displays the allocation of digital and analog signals on the same carrier frequency. The analog signal occupies the central portion of the spectrum, while the digital signal, modulated using OFDM, are distributed on the sidebands. This visual representation

demonstrates how IBOC leverages OFDM to accommodate both digital and analog broadcasts without the need for new spectrum allocations, underlining the versatility and efficiency of OFDM in diverse communication contexts.

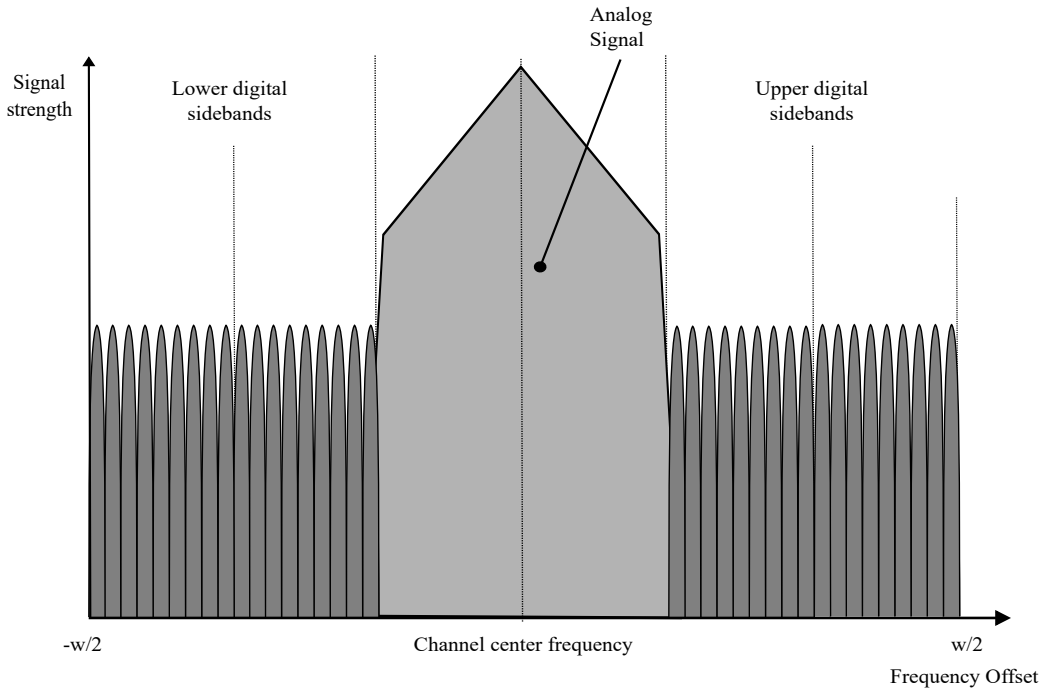


Figure 2.2: IBOC signal spectrum.

2.1.3 The Impact of High PAPR Signals

One of the key characteristics of signals in wireless communication systems, particularly those modulated using OFDM, is high PAPR. This occurs because the subcarriers of the OFDM signal which are in phase, constructively interfere with one another, producing peaks that are much higher than the average power level of the signal. The PAPR represents the ratio between the peak power and the average power of a signal. It depicts how much the power level of the signal can spike in contrast to its average level.

For power amplification, high PAPR poses a critical challenge. PAs are designed to operate close to their saturation point to ensure maximum power efficiency. However, when dealing with OFDM signals with high PAPR, the PA must operate at a point far from its saturation. This is to prevent distortion of the signal's high peak levels, a requirement that results in lower average transmitted power.

The PAPR is mathematically represented as follows:

$$\text{PAPR (dB)} = 10 \cdot \log_{10}\left(\frac{P_{peak}}{P_{avg}}\right). \tag{2.1}$$

Where P_{peak} represents the peak power level that the signal reaches, and P_{avg} is the average power level of the signal.

The criticality of managing high PAPR lies in its interaction with the non-linearity of the PA. When a signal's PAPR is high, the signal becomes harder to amplify and transmit without distortion. Any distortion can result in spectral leakage, where the signal data spills over its designated channel, potentially corrupting adjacent channels. This is an undesirable scenario in communication systems, threatening the integrity of the transmitted data.

Mitigation techniques, such as clipping, filtering, and peak reduction, can be applied to decrease a signal's PAPR and enhance its performance. Likewise, the amplifier's linearity can also be improved for performance enhancement. High PAPR puts a significant strain on PAs, demanding them to be more linear and therefore highlighting the importance of PA linearity improvement techniques in achieving optimal communication system performance.

2.2 RF Power Amplifier Systems

This section discusses PAs for the purpose of RF communication. For a system to be considered an RF communication system, the transmission frequency must fall between 1 MHz and 300 GHz [31]. It is important to note that most RF communication systems are wireless, this is not a requirement. Any data and information transmission within the specified frequency band could be considered RF communication regardless of the medium of transmission.

Figure 2.3 depicts a simplified RF transmitter's system architecture, which constitutes a crucial component of such communication systems. This high-level schematic shows the path of a signal as it navigates through various stages in the system. These stages represent the fundamental operation required for transmitting an RF signal.

An RFPA is essential for transmitting communication signals. Its job is to increase the magnitude of the modulated signal so that it can be received at distances it otherwise would not reach.

The behavior of an amplifier heavily depends on the biasing point, which is the Direct Current (DC) operating voltage or current that is applied to set the amplifier's

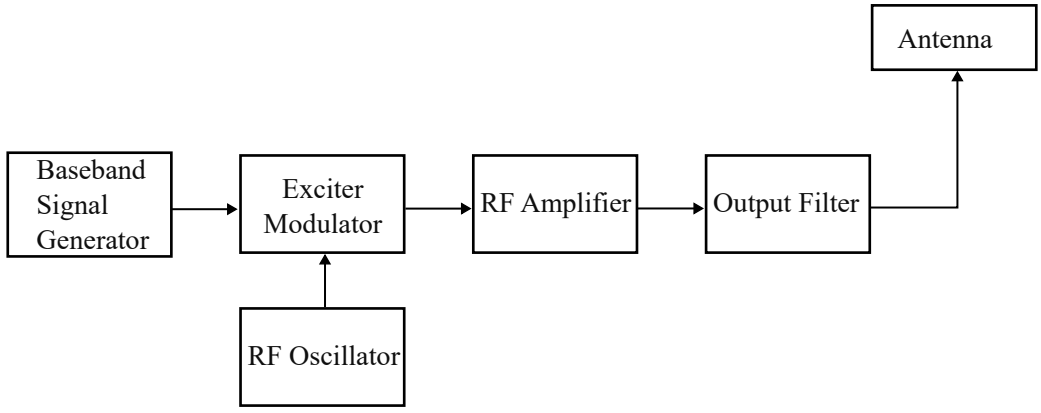


Figure 2.3: System architecture of a simple RF transmitter.

operating state [35]. The biasing point sets the operating conditions that dictate when the amplifier circuitry will be active in conducting electrical current. This active phase and other characteristics play a crucial role in categorizing amplifiers for specific design and application purposes.

Figure 2.4 offers a graphical illustration of amplifier biasing. The graph shows the biasing point for a Class A amplifier. Unlike other amplifier classes, a Class A amplifier is designed to conduct throughout the entire input cycle at the linear portion of the transfer characteristic curve. This allows for the most linear amplification.

Table 2.1 provides a comparative overview of different amplifier classes along with several key operating properties. In the table, "Peak Voltage" refers to the maximum instantaneous voltage that the amplifier can accommodate, and it essentially measures the maximum signal size that the amplifier can amplify without any distortion. "Max Efficiency" measures the amplifier's capability to convert the power supplied to it into power used to produce the output. Higher efficiency means more power is transformed into a useful output. The "Conduction Angle" is the fraction of the input signal cycle during which the amplifier conducts. This angle varies from 2π radians for Class A, to anywhere less than π radians for Class D - F amplifiers.

Table 2.1: Characterization of PA Classes [35].

Class	A	AB	B	C	D - F
Peak Voltage	$2V_{dd}$	$2V_{dd}$	$2V_{dd}$	$2V_{dd}$	$>= 2V_{dd}$
Max Efficiency	50%	50 - 75%	75%	$\approx 90\%$	$\approx 90\%$
Conduction Angle	$\theta = 2\pi$	$\theta = \pi$	$\pi < \theta < 2\pi$	$\theta < \pi$	$\theta < \pi$

Ideally, RFPA used to boost a signal should not contribute any additional distortion. This would interpret to the PA operating at high efficiencies while also

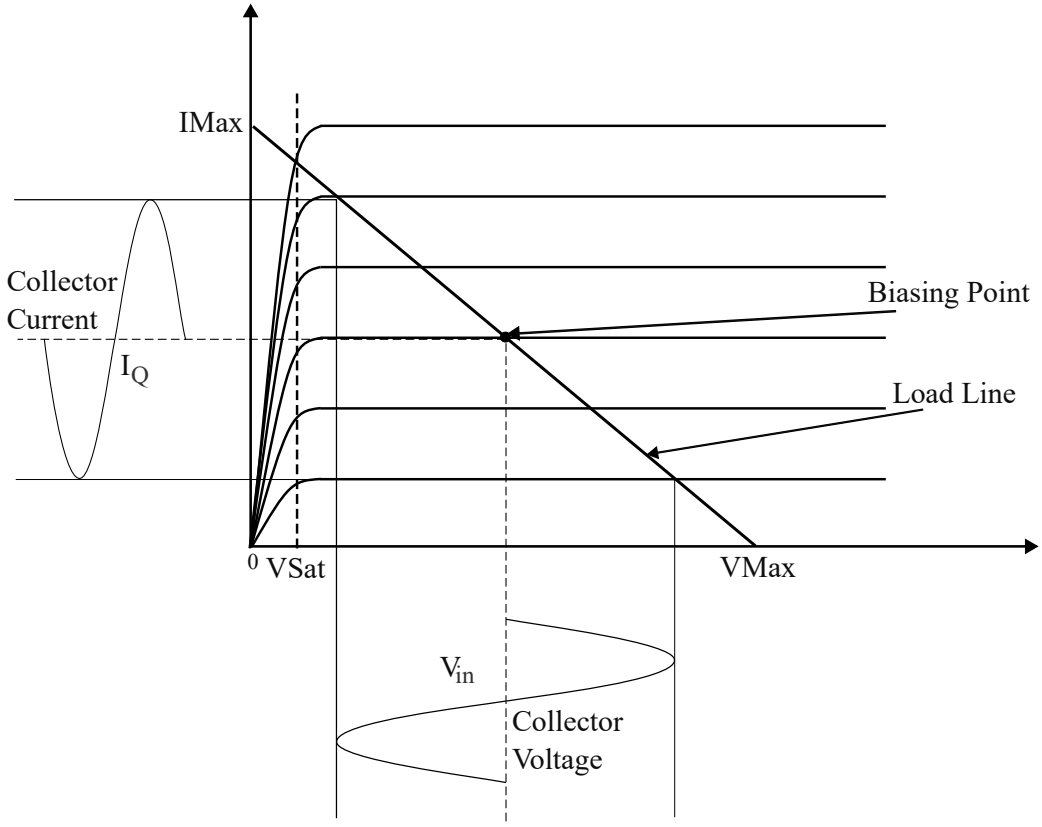


Figure 2.4: Biasing point and input cycle in a Class A Amplifier

maintaining linearity at various signal amplitude levels. These are conflicting properties. An amplifier working at high efficiency is close to saturation, and therefore, it will produce a higher level of nonlinearity.

The RFPA is one of the most power-consuming components in a wireless transmitter. The current allocated to the PA must be used as efficiently as possible to reduce energy waste [35], wherein the linearity must then be enhanced in ways that do not compromise the efficiency of the RFPA.

The DPD technique is widely employed for amplifier linearization, A simplified representation of an amplifier system utilizing DPD linearization is shown in Figure 2.5. The system depicted includes a transmitter chain that utilizes a digital pre-distortion block that feeds a distorted version of the baseband signal to the PA, represented as a Wiener model in this context. The analysis and simulations of model and DPD for this work are completed in baseband using Matlab.

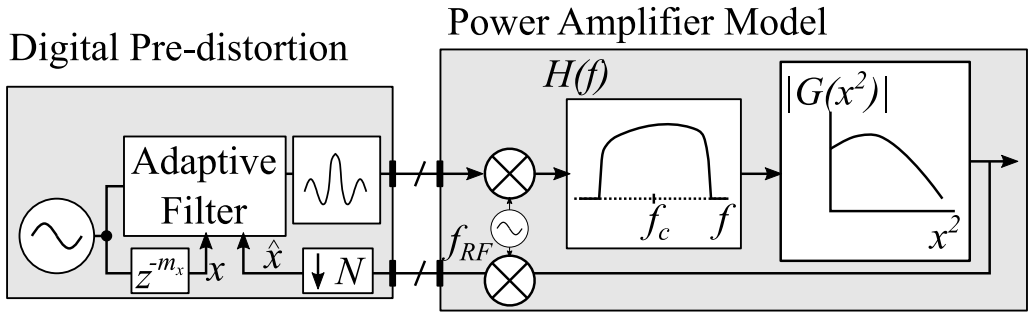


Figure 2.5: System Architecture of a Non-Linear Amplifier with Pre-distortion.

2.3 PA Memoryless Nonlinearities

The superposition principle generally describes a linear system, comprising of additivity and scaling (homogeneity) properties. A system of equations can generally classify various ideal and real applications. Linear systems describe an ideal PA. The output of this system is mathematically represented as

$$Y_t = \alpha X_t, \tag{2.2}$$

where X_t is the input signal and α is the gain. This system has a constant gain value for all input signal magnitudes.

This would be an inaccurate assumption in the real world as PAs are never entirely linear. Amplifier components, application-specific biasing, and environmental

conditions are some of the factors that may introduce both nonlinear and frequent dependent properties to an amplifier system.

Memoryless nonlinearity in a system can be described as the constant distortion property which does not vary with the amplitude of the input signal. The output of a system with memoryless nonlinearity is mathematically expressed in its simplest form as

$$Y_t = \alpha(X_t)X_t, \tag{2.3}$$

where X_t also represents the input signal and $\alpha(X_t)$ is the gain value of the signal at a value of X_t. Increasing memoryless nonlinearity may be observed in the system as the complexity increases. To show a simple form of PA nonlinearity, the system output can be expressed as a Taylor series where varying gain values are multiplied by several orders of the input value, written as

$$Y_t = \alpha_1 X_t + \alpha_2 X_t^2 + \alpha_3 X_t^3 + \alpha_4 X_t^4 + \cdots . \tag{2.4}$$

Equation (2.4) can be described as a memoryless polynomial representation of memoryless nonlinearity. The nonlinearity in this system is expressed using the additional terms in the series. With increasing order the nonlinear relationship between the input signal and the output increases in complexity.

2.3.1 PA Memoryless Nonlinearity Representation

If a PA can be represented with a set of equations that describes its behavior, it is helpful to understand how they translate into distortions. This is necessary because to model and analyze a PA properly; the non-linearity must be quantified to a degree. Such analysis could reveal where an amplifier stops providing a linear output.

A characteristic nonlinearity effect of PA's pushed to saturation can be described as gain compression. This would generally result in the amplifier providing less power on the output signal than an increasing input signal. A commonly used metric to observe this effect is the 1 dB compression point, which indicates the onset of noticeable distortion in the amplifier's output.

The 1dB compression point concept will be further explored and visualized in Section 2.3.1.1, presenting a figure illustrating this phenomenon. This figure will provide a more intuitive understanding of how gain compression manifests in the context of PAs operation and will serve as a practical guide for understanding the

nonlinear behavior of amplifiers.

2.3.1.1 Single-Tone Analysis

This section analyzes the impact of a nonlinear amplifier, as represented by (2.2),on a single-tone input.

A single-tone sinusoidal input signal is introduced, as represented by the following equation:

$$X_t = A\cos(\omega t). \tag{2.5}$$

Considering only the first three orders of the model, the output signal can be written as [27]

$$Y_t = \alpha_1 A\cos(\omega t) + \alpha_2 A^2\cos^2(\omega t) + \alpha_3 A^3\cos^3(\omega t). \tag{2.6}$$

The sinusoidal input in a nonlinear system produces an expression at its fundamental frequency ω and distortion at integer multiples ω described as harmonic distortion. By expanding (2.6), the resulting equation is represented by

$$Y_t = \frac{\alpha_2 A^2}{2} + \left(\alpha_1 A\frac{3\alpha_3 A^3}{4}\right)\cos(\omega t) + \frac{\alpha_2 A^2}{2}\cos(2\omega t) + \frac{\alpha_3 A^3}{4}\cos(3\omega t). \tag{2.7}$$

This modeled system also reveals an added DC value to the output. This change in operating point is described as the self-biasing nature of an amplifier [35]. From (2.7), a gain compression is also observed. By isolating the gain value of the fundamental frequency, as Razavi [27] suggests, the small signal gain value can be estimated with the assumption that the even harmonics are negligible. Then, the gain is expressed as

$$G = \frac{Y_t}{X_t} = \frac{(\alpha_1 A + \frac{3\alpha_3 A^3}{4})\cos(\omega t)}{A\cos(\omega t)}. \tag{2.8}$$

(2.8) can be reduced further and the gain G as a function of the input amplitude A is simply

$$G = \alpha_1 + \frac{3\alpha_3 A^2}{4}. \tag{2.9}$$

The gain expression is where the amplifier has crossed its linear region into its saturation region.

Figure 2.6 visually captures the 1-dB Compression Point concepts and the Third Order Intercept Point (IP3) that will be presented in the next section. The 1 decibel (dB) compression point, symbolizing the beginning of significant gain compression, marks the transition from linear to non-linear amplifier operation.

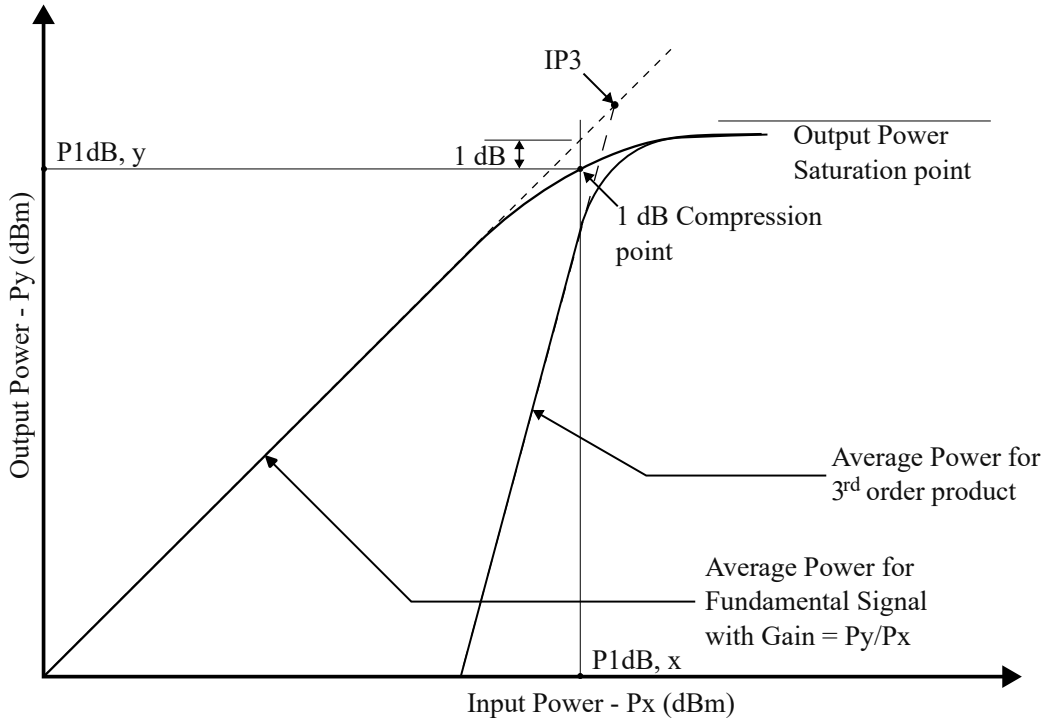


Figure 2.6: 1dB Compression point and input third-order intercept point.

2.3.1.2 Two Tone analysis

The above analysis provides a model of the nonlinear amplifier in response to a single tone. However, when excited by a signal that extends in frequency, a simple analysis can be obtained using a signal with two dominant fundamental frequencies. A two-toned signal with a peak amplitude A_1 can be mathematically represented as

$$X_t = A_1(\cos(\omega_1 t) + \cos(\omega_2 t)). \tag{2.10}$$

By applying this signal to the same memoryless polynomial model, and considering only the first three terms, the output signal Y_t is represented by [32]

$$Y_t = \alpha_1(A_1(\cos(\omega_1 t) + \cos(\omega_2 t))) + ... + \alpha_3(A_1(\cos(\omega_1 t) + \cos(\omega_2 t)))^3. \tag{2.11}$$

The system output Y_t expansion results in a complex solution that reveals non-negligible energy at new frequencies in the spectrum that are not harmonics of the fundamental frequencies. This phenomenon is known as the Intermodulation Distortion (IMD), which occurs due to applying the two-toned signal to a non-linear system [27]. Specifically, the distortion due to the 3rd order coefficient appears at the Third Order Intermodulation Distortion (IMD3) frequencies, at $\omega_{12} = 2\omega_1 - \omega_2$ and $\omega_{21} = 2\omega_2 - \omega_1$. The IMD3 is the signal amplitude that appears at ω_{12} and it can be expressed as [32]:

$$\text{IMD}_3(2\omega_1 - \omega_2) = a_3 A_1^3 \frac{3}{4} \cos(2\omega_1 - \omega_2)t. \tag{2.12}$$

As observed in Figure 2.7, The IMD3 products can be generated very close to ω_1 or ω_2. This proximity to the fundamental frequencies increases the risk of causing interference, as they fall may fall within a receiver's desired frequency band.

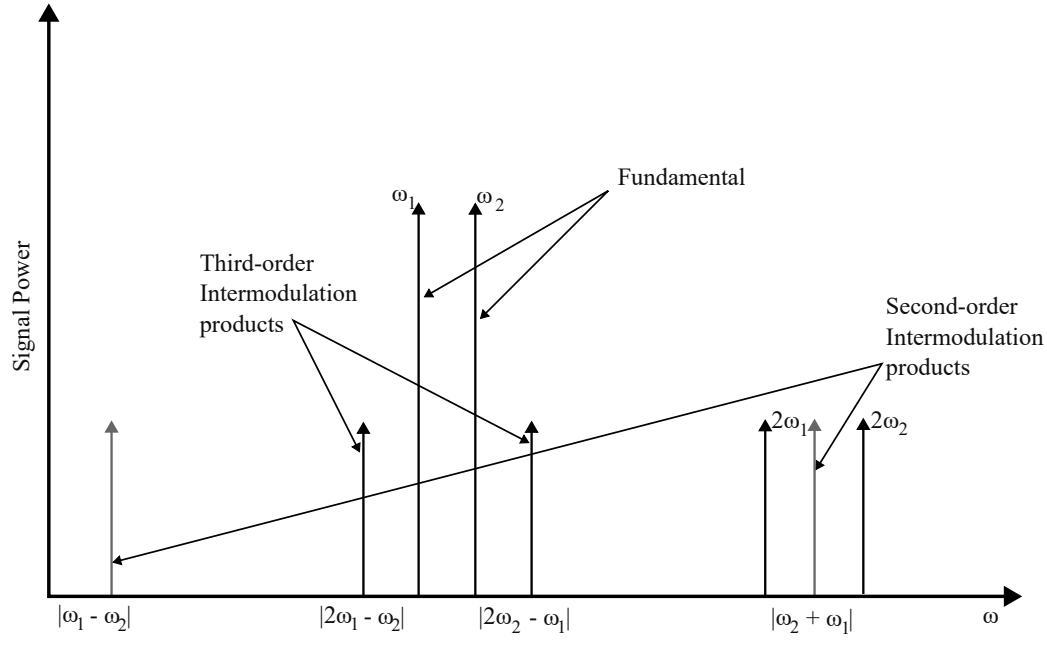


Figure 2.7: Signal spectrum with higher order products.

A performance metric, called the IP3 [27], has been defined to characterize the third-order intermodulation behavior. The IP3 occurs when the magnitude of the signal at the fundamental frequency equals the magnitude of its third-order intermodulation product. This is possible because the intermodulation product grows three times faster than the main component.

The IP3 shown in Figure 2.6 occurs when the extrapolations of the fundamental

signal power and third-order intermodulation distortion product intersect. This intersection is a theoretical point where the intermodulation distortion's power equals the input signals' power.

2.4 Phase Distortion due to Nonlinearity

The narrowband model of the PA presented in Section 2.3.1.1 assumes a sinusoidal signal is applied to the amplifier. This section describes a baseband model of the PA in response to a low-frequency signal modulating the carrier. As will be shown, the distortion effects of the nonlinear system can be represented using Amplitude-to-Amplitude (AM/AM) and Amplitude-to-Phase (AM/PM) distortion. AM/AM and AM/PM characteristics can be explored with different models, one of which is the memoryless Saleh model. The transfer function for an AM/AM response describes how the amplitude of the input signal factors into the distortion of the output amplitude, and that of the AM/PM response shows how it affects the phase of the output of the system. If provided a complex input b_x, which is represented in polar form as

$$b_x = u_x e^{j\alpha_x}, \tag{2.13}$$

where u_x represents the magnitude of the input signal and α_x represents its phase. Drotar et al [8] describes the base-band transfer functions for AM/AM and AM/PM characteristics represented by G_{u_x} and ϕ_{u_x} respectively as

$$G_{u_x} = \frac{k_g u_x}{1 + x_g u_x^2}, \tag{2.14}$$

and

$$\phi_{u_x} = \frac{k_\phi u_x^2}{1 + x_\phi u_x^2}. \tag{2.15}$$

In (2.14) and (2.15), k_g and k_ϕ are coefficients that characterize the relationship between the input and output amplitudes and phases respectively, while x_g and x_ϕ are coefficients that represent the nonlinearity of the amplifier. These values are typically extracted empirically using a Least Square (LS) approximation to minimize the error between the model and the behavior of the PA being replicated [23]. Introducing the input to this characterized system, the output b_y is equal to

$$b_y = G_{u_x} e^{j(\alpha_x + \phi_{u_x})}. \tag{2.16}$$

From (2.16) it is observed that the phase of the output is distorted by the AM/PM characteristic ϕ_{u_x}.

2.5 PA Memory Distortion

This section delves into the intricacies of memory effects in PA's and their impact on the amplified signal. The term "memory" in PA's refers to the influence of past data samples on the current output of the system. This influence is manifested as distortions in the output signal, as previous samples essentially intermingle with the current sample. As Wood [35] discusses, memory effects in RFPA's are primarily a consequence of energy storage within the system. Various parts within a PA, which form the basic building components, are capable of storing energy. In capacitors, for example, charges accumulate, forming an electric field that can affect current flow in subsequent amplifier operations.

Thermal effects which have a significant influence on memory distortion stem from the thermal energy generated by the operation of the PA. The thermal energy produced can cause fluctuations in the amplifier's performance over time, as it can influence the physical properties of components in the amplifier circuit, leading to changes in their operational behavior.

The most substantial contribution to memory distortion in PAs, however, comes from electrical memory effects. These effects can arise from a variety of sources within the amplifier, such as stored charges in transistor components, or the influence of the biasing line in the circuit.

2.6 Spectral Regrowth Characterization

This section examines the concept of spectral regrowth characterization, a critical phenomenon in wireless communication systems. The focus is on its impact on signal quality, how it is quantified through ACLR, and the importance of complying with regulatory standards.

Spectral regrowth, primarily due to the IP3 non-linearity of the power amplifier, refers to the unintended spread of signal power beyond the intended bandwidth or channel [36]. When a signal passes through a nonlinear amplifier, it can suffer distortions that manifest as additional frequency components outside of the original signal's bandwidth. This leads to the unintended spread of the signal's power, known as spectral regrowth. They pose a significant issue in wireless communication due to

the potential interference they can cause in adjacent channels.

Memory effects in the power amplifier, caused by the inter-dependencies between current and past input values, can also contribute to spectral regrowth. This is because the distortions induced by memory effects could increase the bandwidth of the amplified signal, contributing to the power leakage into adjacent frequency bands.

The ACLR is a crucial parameter that quantifies this effect. It measures the power ratio between the carrier signal in the designated channel and the power leaked into the adjacent channel. Higher ACLR values are preferred as they indicate less interference and better fidelity of the transmitted signal.

The ACLR can be mathematically represented as:

$$ACLR = 10 \log_{10} \left(\frac{P_{Carrier}}{P_{Adjacent}} \right) \quad \text{dB}. \tag{2.17}$$

$P_{Carrier}$ represents the power in the desired channel or carrier frequency, and $P_{Adjacent}$ represents the power leaked into the adjacent channel. The resulting ACLR is expressed in decibels (dB). Understanding and controlling spectral emissions and the resulting ACLR are critical to ensuring high-quality, interference-free wireless communication. Regulatory bodies like the Innovation, Science and Economic Development (ISED) and Federal Communications Commission (FCC) impose stringent requirements on these parameters to uphold the reliability and integrity of communication networks.

Chapter 3

Power Amplifier Modelling

This chapter reviews the methods of estimating an amplifier's behaviors. The requirement for proper modeling of amplifier distortion is explained in section 3.1 This section will also describe how the memoryless and memory distortion properties are represented for this thesis. In section 3.2, the model described in the previous section is mathematically represented using a simplified pruned Volterra Taylor series model for further understanding.

3.1 Nonlinearity Expressed as a Cascade of Distortion Properties

Previously, the independent effects of memoryless nonlinearity and memory distortion in an amplifier were discussed. This section presents a block-oriented model that describes a reference amplifier system as a cascade of these distortion properties. Figure 3.1 illustrates a Wiener model system architecture as an example of such a representation. In this model, a linear frequency selective system precedes a memoryless nonlinearity. This is one way of organizing the components of the system, though other models arrange these elements differently. The Hammerstein model, as opposed to the Wiener model, places the nonlinear component before the frequency-selective filter [22].

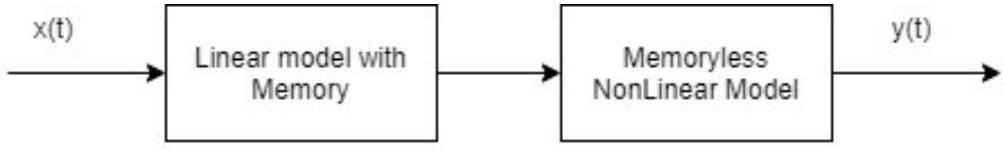


Figure 3.1: Wiener model system architecture.

Both the Wiener and the Hammerstein models are effective at identifying nonlinear systems with memory effects. However, subtle differences between them lead to different outputs. A Wiener model is more suitable for a system where the frequency-selective circuit occurs before the nonlinear amplification. Conversely, a Hammerstein model provides a better characterization for a system where the nonlinearity occurs before the memory effects. The Wiener-Hammerstein model, which consists of a

linear dynamic block before and after the nonlinear block, can characterize complex systems where memory effects occur before and after the nonlinearity.

In real-world scenarios, disentangling the memory distortion from the memoryless nonlinearity might not always be straightforward, as discrete active electronic components can exhibit both frequency selectivity and nonlinearity. Regardless of this challenge, understanding and accurately representing these phenomena are crucial steps toward effective system design and analysis.

3.1.1 Memoryless Distortion

In modeling a PA, it is necessary to consider the nonlinearity and computational complexity involved. The severity of the nonlinearity can be captured by the finite order K. The output signal $y[n]$ can be expressed as a polynomial function of the input signal $x[n]$ in such models. This is simplified as

$$y[n] = \sum_{k=0}^{K-1} \alpha_k x[n]|x[n]|^k, \tag{3.1}$$

where α_k are the polynomial coefficients and $|x[n]|^k$ are the powers of the input signal.

Here, a gain curve is used to model a real amplifier which is shown in Figure 3.2, with amplitude-dependent gain and phase distortion. It is evident that the amplifier does not remain constant prior to compression point as is observed in the region below 15 dBm. The gain curve is indexed using the instantaneous input power $x[n]^2$ to obtain the complex gain. The response from this process is expressed as

$$y[n] = x[n]G|x^2[n]|, \tag{3.2}$$

where $G(|x^2[n]|)$ denotes the PA gain as a function of the instantaneous input power.

While this method has been suggested in a previous study by Fatungase et al. [9], it's essential to acknowledge that both the polynomial and the Look-Up Table (LUT) approaches can be helpful in their own right for modeling RFPAs. The choice between them can often depend on the specific requirements and constraints of a given application. For instance, the LUT method can be advantageous when direct mapping from input to output is preferred.

On the other hand, polynomial models offer flexibility in adjusting the level of complexity and, thus, the model's accuracy. By increasing the order K of the polynomial, we can achieve a higher degree of accuracy in modeling the amplifier's behavior.

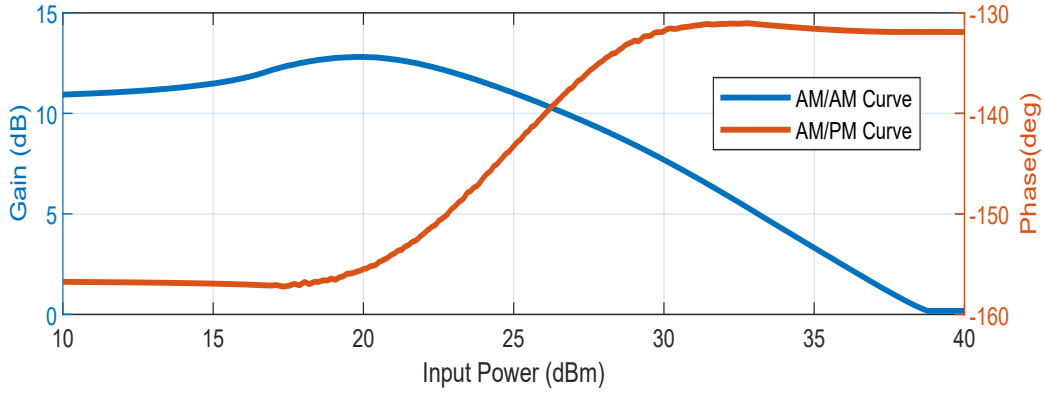


Figure 3.2: Model amplifier gain curve.

However, this added precision comes at the cost of increased computational complexity. Consequently, it is crucial to identify an appropriate value of K that balances the need for a sufficiently accurate model with the practical constraints of computational efficiency.

A normalized gain value may be derived for progressing input signal value to characterize a gain curve. This method of sweeping the input amplitude and providing a corresponding output value demonstrates a simple yet fundamental way of analyzing the behavior of an amplifier. It is possible to obtain an ideal region for input signal operation from a gain curve.

3.1.2 Front-End Model with Memory using Digital Filters

A linear system with memory distortion can be characterized as a Linear Time-Invariant (LTI) system. The Dirac delta function, which produces the filter's impulse response, can be used as an ideal representation of the memory effect in the system. The convolution of the input and the impulse produces the resulting response of the system. This response is expressed as

$$y(t) = x(t) * h(t), \tag{3.3}$$

where $x(t)$ is the input signal data and $h(t)$ is the filters response. When expressed in discrete format with time index n, the response of the system $y[n]$ can be mathematically represented as

$$y[n] = (x * h)[n] = \sum_{-\infty}^{\infty} x[m] \cdot h[n-m]. \tag{3.4}$$

As a result of the cumulative property of a convolution operation, the order of the input x and the filter response h can be inverted with no consequences to the system's response. Considering a Finite Impulse Response (FIR) with M coefficients, the response $y[n]$ is now written as

$$y[n] = (h * x)[n] = \sum_{m=0}^{M-1} h[m] \cdot x[n-m]. \tag{3.5}$$

From (3.5), the representation of linear memory distortion as an LTI system is characterized more intuitively. The value of M determines how many discretized historical inputs affect the current response. The severity of the system thus depends on the number of coefficients that ideally represent the filter's response and the weight of each coefficient.

Fatungase et al [9] describes the implementation of a linear memory distortion using a series of digitally programmable filters that are cascaded to meet the complexity required for an adequate representation of a system with memory. The first filter is a unity gain low-pass FIR filter, denoted as $F_1(\omega)$, which defines the system bandwidth as a function of discrete frequency ω. Following this filter is $F_2(\omega)$, a filter that exhibits a gain with a slope as a function of frequency. The slope is measured in dB/MHz.

An all-pass filter $F_3(\omega)$, featuring a group delay based on the delay taps present in the cascaded filter, is employed to model the delay m_x [9]. This configuration ensures a match between the transmit baseband signal at the input of the front-end electronics and the sampled output of the PA. The increasing steepness of the filter response slope dictates the severity of the memory effect. This approach enables the definition of a slope within a specified bandwidth for the system.

The discrete-time FIR filter $F_1(k)$ is designed in Matlab using a standard digital filter with a linear in-band phase to implement the unity gain low-pass filter [9]. A high-order filter with 200 coefficients is chosen to avoid in-band variation. The impulse response is zero-padded to have a radix-2 window size equal to 1024, then converted to the frequency domain to obtain $F_1(\omega)$. In the frequency domain, the filter frequency response is multiplied by $F_2(\omega)$, which varies as a function of frequency. Filter $F_3(\omega)$ is initially defined in the time domain as an impulse with a delay m_x,

and later converted to the frequency domain. The three cascaded filters are multiplied in the discrete Fourier domain, and the result undergoes an inverse Fourier transform to obtain an equivalent discrete-time impulse response for the cascaded filters. A convolution of the input signal with the cascaded filters impulse response as described in (3.3) mathematically yields skewed samples due to the effects of historical data. Figure 3.3 depicts a series of filters that shows the spectrum of LTI filters with varying slopes ranging from 0 dB/MHz to 2 dB/MHz.

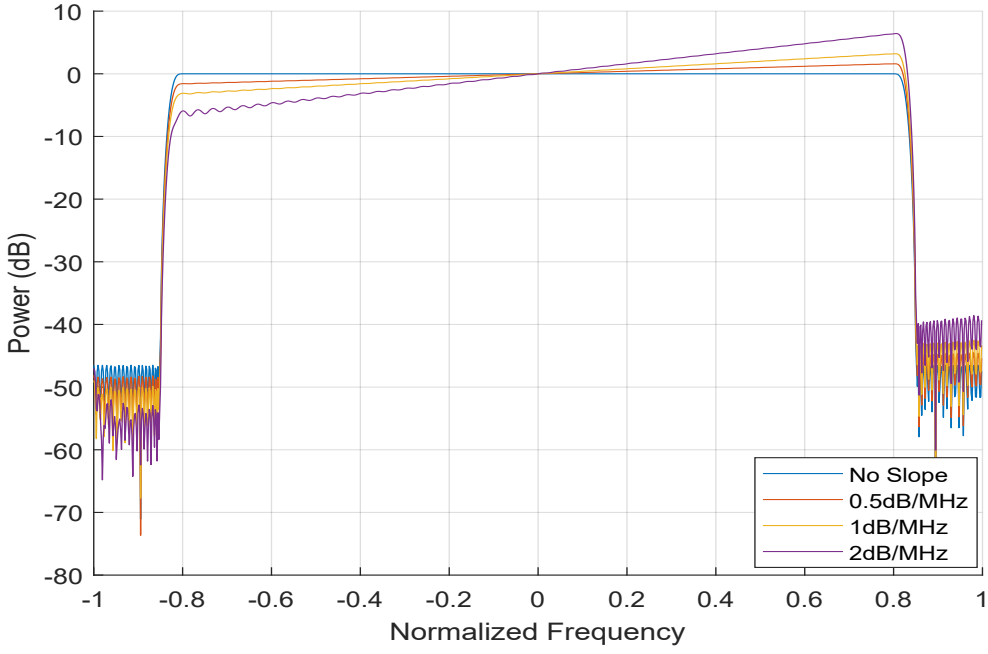


Figure 3.3: Spectrum of LTI filters with varying slope.

To evaluate the functionality of the cascaded filters, they are tested with different characteristics [9]. When a filter response with a zero slope is applied to a signal in the filter band, the signal output exhibits only a group delay, confirming the absence of inter-symbol interference. When the delay between the input and output of the filter response is synchronized, the output of the filter matches its input. Synchronization can be achieved, for example, using cross-correlation between the input to the transmit front-end and the sampled PA output.

3.2 The Reduced Volterra Model

The characterization of an amplifier described in Section 3.1 is widely studied for general application theory and offers a simplified perspective on amplifier behavior. Though simple, its use in precisely correcting individual distortions in RFPAs is limited [37]. As a Wiener model characterization does not fully represent the amplifier's complex behavior. This section discusses the translation of the represented system in 3.1 to a memory polynomial model and determines how its coefficients are ideally estimated.

3.2.1 Memory Polynomial Model

The Volterra series is a mathematical method used to model nonlinear systems, particularly those with memory effects. The Volterra series expands the output of a nonlinear system in terms of its inputs. It is expressed as the sum of multidimensional convolutions of terms, with each term in the series representing a different order of nonlinearity. The first term of the series corresponds to linear behavior, while higher-order terms represent increasingly complex nonlinear effects. Each term in the series is characterized by a set of coefficients, which quantify the system's response to different combinations of input signals. The response of a Volterra series can be mathematically represented in discrete-time as

$$y[n] = \sum_{k=0}^{K} \sum_{m_1=0}^{n-1} \cdots \sum_{m_k=0}^{m_{k-1}-1} c_k[m_1, \cdots, m_k] \prod_{i=1}^{k} x[m_i], \tag{3.6}$$

where $y[n]$ is the output of the system at a time n, $x[m]$ are the input signals and $c_k[m_1, \cdots, m_k]$ are the Volterra Kernel coefficients [22].

While the full Volterra model is capable of capturing all orders of nonlinearity, it requires a large number of coefficients and memory elements to represent the behavior of the system accurately. This makes the model computationally expensive and challenging to implement in practice. A reduced Volterra model defines a memory structure, which is a more straightforward case of the general formula described in(3.6). Reduced Volterra models are often used instead of the complete Volterra series because they provide a less resource-intensive way of modeling nonlinear systems with memory effects.

This thesis further explores the characterization of amplifier systems using the memory polynomial model. The memory polynomial model is a reduced Volterra

model that extends the memoryless polynomial to include the memory effects. The memoryless nonlinear effects and linear memory effects expressed in (3.1) and (3.5) respectively may be represented using a single Taylor series. The response of which is written as

$$y[n] = \sum_{k=0}^{K-1} \sum_{m=0}^{M-1} c_{km} x[n-m]|x[n-m]|^k, \tag{3.7}$$

In the equation above, c_{km} represents the weight applied to the input signal at the m^{th} memory depth and at the k^{th} nonlinearity order [22].

3.2.2 Direct Matrix Inversion

In this section, a new approach for coefficient estimation using forward and reverse signals in the direct matrix inversion process is proposed for a memory polynomial model.

The response of the PA system can be expressed as a matrix multiplication using the memory polynomial model. The output $y[n]$ can be obtained by a summation of all entries of a vector that is the result of the matrix multiplication of the input transformation matrix $\mathbf{X}[\mathbf{n}]$ with the memory polynomial model coefficient vector c_{PA}, This vector is expressed as (3.8)

$$y[n] = \mathbf{X}[\mathbf{n}] c_{PA}, \tag{3.8}$$

Assuming a model with M and k number of memory and memoryless nonlinear distortion coefficients, respectively. The output signal vector is mathematically written as

$$y[n] = \begin{bmatrix} x[n] & x[n]|x[n]| & \dots & x[n]|x[n]|^{K-1} & \dots & x[n-M]|x[n-M]|^{K-1} \end{bmatrix} \begin{bmatrix} c_{11} \\ c_{12} \\ \vdots \\ c_{1K} \\ \vdots \\ c_{MK}. \end{bmatrix} \tag{3.9}$$

In Equation (3.9), the output signal vector $\mathbf{y}[n]$ represents a single discretized

output signal at a given time n. Each element of this vector is calculated by multiplying a row of the input transformation matrix $\mathbf{X}[n]$ with the memory polynomial model coefficient vector $\mathbf{c_{PA}}$ which has a size of M^*K [21].

This general equation can simplify to different specific cases, thereby showcasing its adaptability to diverse scenarios. For example, when $M = 1$, there is no memory effect, and the model collapses to an instantaneous non-linearity polynomial, making it suitable for scenarios where only instantaneous effects are of concern.

Moreover, when $K = 1$, the model simplifies to a linear model, and the formula becomes equivalent to an FIR filter. This can be especially useful in applications where a simple linear response is desired and nonlinear effects are negligible.

Understanding these cases provides a more intuitive grasp of the functionality and flexibility of the memory polynomial model for different use cases in power amplifier modeling.

Figure 3.4 provides a two-dimensional visualization of the architecture of a memory polynomial model and the application of the model's coefficient for the estimation of models output to further understanding of the simulation conditions

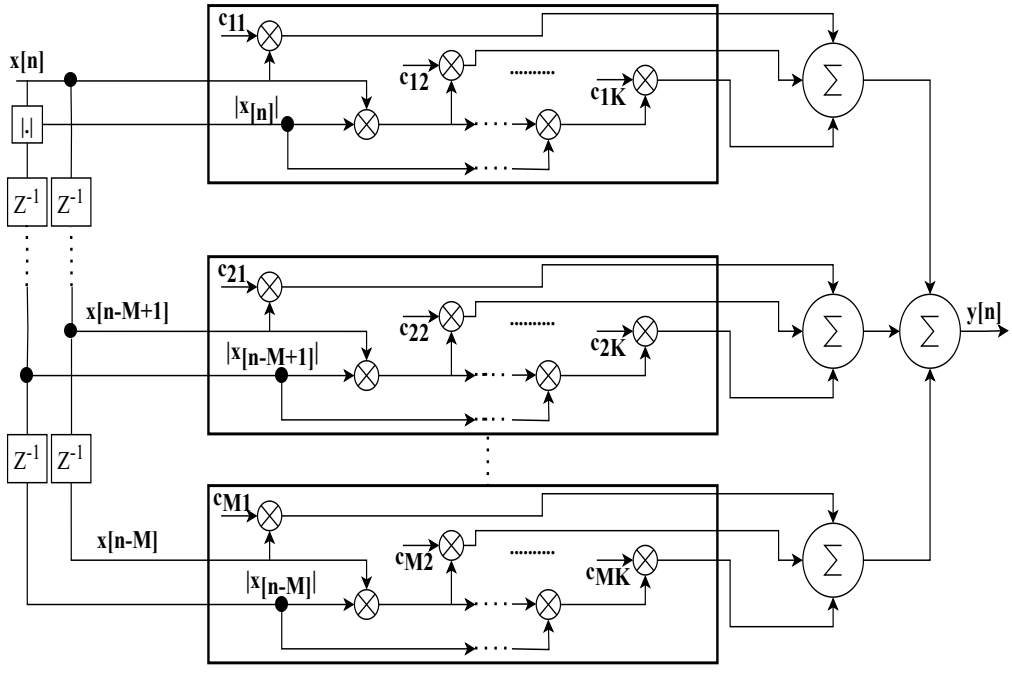


Figure 3.4: A realistic 2-D coefficient application for a memory polynomial model.

There is a proposed method that involves using a series of live signals to establish

the input-output correlation. This process involves applying the signal to the PA model to obtain both forward and reverse signals. The direct matrix inversion process is then applied to estimate the coefficients [22].

In addition to using a memory polynomial model to represent the behavior of a PA, it is also essential to determine the ideal number of coefficients to use in the model. The number of coefficients is related to the order of the model, which corresponds to the memory depth as well as the nonlinearity order of the system.

One approach for determining the ideal number of coefficients is to use the Mean Squared Error (MSE) metric. The MSE is a commonly used metric to quantify the average squared difference between predicted values and actual values. In the context of coefficient estimation for a memory polynomial model, the MSE equation can be defined as follows:

$$\text{MSE} = \frac{1}{N} \sum_{n-1}^{N} \left(y[n] - \hat{y}[n] \right)^2 , \tag{3.10}$$

where N is the total number of samples, $y[n]$, represents the desired output of the PA system, and $\hat{y}[n]$ represents the predicted output obtained using the estimated coefficients from memory polynomial model.

The MSE is a key metric used to assess the accuracy of coefficient estimation for a memory polynomial model. By evaluating the MSE for different numbers of coefficients, we can determine the ideal model complexity.

When too few coefficients are used, the model may lack the necessary complexity to capture the nonlinear behavior and memory effects of the PA. As a result, the predicted output will deviate significantly from the actual output, leading to a high MSE.

Conversely, when an excessive number of coefficients is used, the model may overfit the training data, capturing noise and irrelevant features. This results in an overly complex model that fails to generalize well to new input signals. In such cases, the predicted output may closely match the training data but perform poorly with unseen data, leading to a high MSE.

The goal is to find the ideal set of coefficients where the MSE is minimized, indicating the optimal number of coefficients for an accurate and efficient representation of the PA's behavior. This optimal number strikes a balance between capturing the essential nonlinearities and memory effects of the system while avoiding excessive complexity and overfitting.

To illustrate this process, a Matlab simulation is executed using a signal, which is divided into two portions: a testing data stream and a validation data stream. An amplifier model represented by the cascade of the gain curve expressed in Figure 3.2 and a 1dB/MHz slopped filter is used to generate the output for both streams. The testing data stream served as the basis for estimating the coefficients. Following this, the output of the validation stream is calculated using the coefficients estimated from the testing data. The MSE is then calculated as expressed in (3.10) by comparing this calculated output with the actual output of the validation data stream. Using a nested loop, the simulation explored various combinations of memory and nonlinearity orders, ultimately identifying the combination that yielded the lowest MSE.

Table 3.1 summarizes the results of the simulation. The table signifies the MSE obtained for a specific combination of memory and nonlinearity orders. The columns correspond to the number of memory terms, and the rows represent the number of nonlinearity terms.

Table 3.1: MSE coefficient estimation for a model with 1dB/MHz slopped filter.

Memory Terms	1	2	3	4
Nonlinearity Terms				
1	6.30 e-4	1.01 e-4	2.69 e-4	2.62 e-4
2	8.22 e-4	1.27 e-4	1.01 e-4	1.02 e-4
3	9.46 e-4	1.06 e-4	1.07 e-4	1.07 e-4
4	9.67 e-4	1.08 e-4	8.12 e-5	8.70 e-5
5	9.93 e-4	1.05 e-4	1.05 e-4	1.06 e-4

Analyzing the table, it becomes evident that using three memory terms and four nonlinearity terms yielded the smallest MSE, indicating an optimal representation of the model's behavior.

Chapter 4

Power Amplifier Predistorion Algorithms

This chapter expands upon the various methods with which an amplifier's distortion properties could be diminished. An iterative approach is used to equalize a frequency selective low pass filter in section 4.1. Then in section 4.2, approaches to compensate for memoryless nonlinearity and memory distortion properties are explored. The direct matrix inversion process is used to correctly estimate a set of correction coefficients, where its effects on an input signal result in an inverse amplifier behavior.

4.1 Equalization of a Linear Filter

Equalization of the linear filter provides insight into the compensation of memory distortion in PA's. The primary objective of linear filter equalization is to mitigate distortion introduced to a signal as it passes through a communication channel. These distortions are typically caused by a range of factors such as noise, interference, spectral leakage from adjacent channels as well as inherent characteristics of the carrier channel. A linear filter equalizer operates by inversely modeling the effects of the channel. Essentially, it applies the inverse of the channel's impulse response to the received signal, intending to regenerate the original transmitted signal. An equalizer may be used to determine the exact inverse of a channel's impulse response which may be helpful in some applications. In other applications, the equalizer considers not only the channel's impulse response but also the noise introduced by the channel. The performance of a linear filter equalizer for a specific application is heavily dependent on the precise estimation of the channel's characteristics.

In this section, the application of an Least Mean Squares (LMS) equalizer using the Stochastic Gradient Descent (SGD) technique for the equalization of a linear filter is discussed. The SGD is particularly suitable for applications where the objective function has a complex landscape with multiple local minima. An adaptive equalizing system is more suitable because it provides flexibility in adjusting its parameters dynamically in real time to counteract the changing characteristics of the communication channel. Adaptive equalizing systems measure the discrepancy between the output signal and the transmitted signal, which constitutes the error and

is subsequently fed back into the adaptive filter. This process facilitates retraining the system's tunable complex multipliers, also known as weights [26]. This feedback mechanism underscores the central role of error in controlling the system.

Figure 4.1 illustrates the architecture of an adaptive equalizer. The chosen adaptive algorithm determines the weights $\mathbf{w}$. The selection of this algorithm and its parameters critically influences the rate of convergence and overall stability of the system. The chosen algorithm for this system operates by optimizing the equalizer coefficients to minimize the MSE between the desired and the actual equalizer output signal [26].

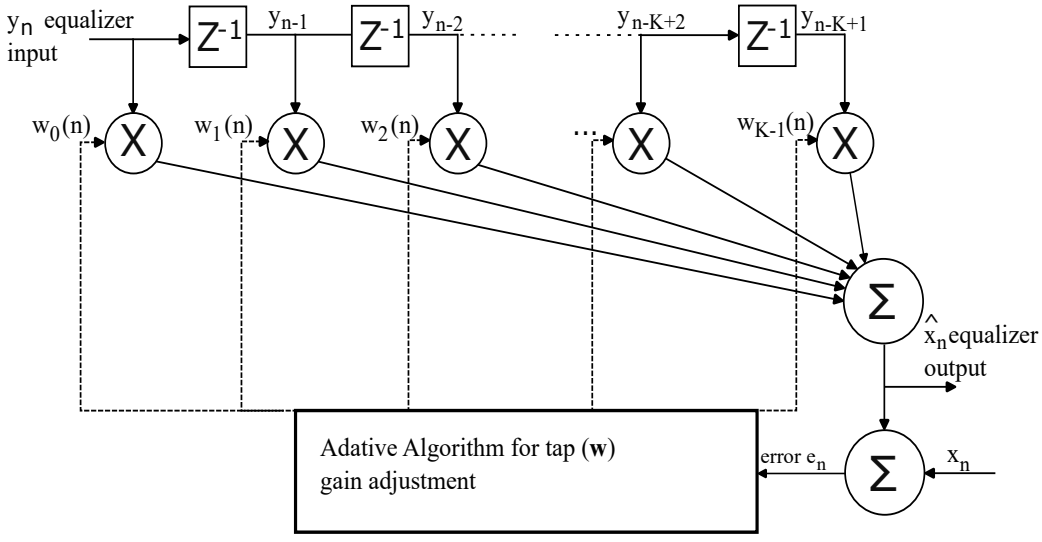


Figure 4.1: Architecture for an adaptive equalizer.

If the equalized output, denoted as $\hat{x}_n$, is described as [26]:

$$\hat{x}_n = \mathbf{w}_n^T \mathbf{y}_n, \tag{4.1}$$

and the error, denoted as e_n, is calculated as [26]:

$$e_n = x_n - \hat{x}_n, \tag{4.2}$$

Then, the weights for the subsequent iteration can be expressed as [26]:

$$\mathbf{w}_{n+1} = \mathbf{w}_n - \alpha e_n^* \mathbf{y}_n, \tag{4.3}$$

where n signifies the current sequence of iteration.

A crucial aspect that is important for the proper function of an adaptive algorithm is the setup of its parameters. The performance of the algorithm heavily relies on the proper configuration of these parameters which include the step size or learning rate (α) as well as the filter tap length.

The step size determines the rate at which the algorithm adapts the filter coefficients based on the error signal. A more significant step size may result in faster convergence but may result in an increased possibility for divergence, while a smaller step size can lead to slower convergence but enhanced stability. Finding the optimal step size involves balancing convergence speed and stability.

The filter tap length, on the other hand, determines the number of filter taps used in the LMS algorithm. A longer tap length allows for more complex and accurate modeling of the filter's characteristics but can increase computational complexity and introduce overfitting. Conversely, a shorter tap length simplified the model but may result in insufficient compensation for the filter distortion.

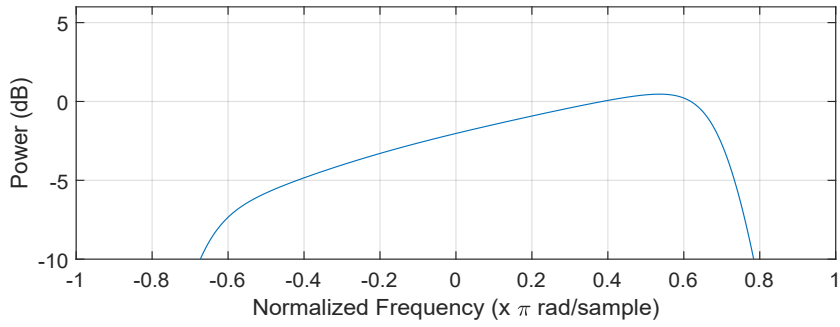
The efficacy of the LMS adaptive equalizer can be demonstrated using MATLAB simulations, which mirror the mathematical modeling of the system. The objective of the filter is to rectify a signal distorted by transmission through a frequency response inducing a memory effect. Over multiple iterations, the algorithm adjusts the weights from initial values to optimal ones that correct the distortion. In this case, the distortion takes the form of a ramp, and the system generates values to counter this ramp. The convolution of the weights and memory filter should yield a flat response, as corroborated by the following simulation below.

The simulation setup involves generating an IBOC signal which represents a wideband broadcast signal at baseband. This signal passed through the linear filter with the specified ramp. The LMS algorithm is subsequently employed for equalizing the distorted signal, aiming to correct the effects caused by the filter.

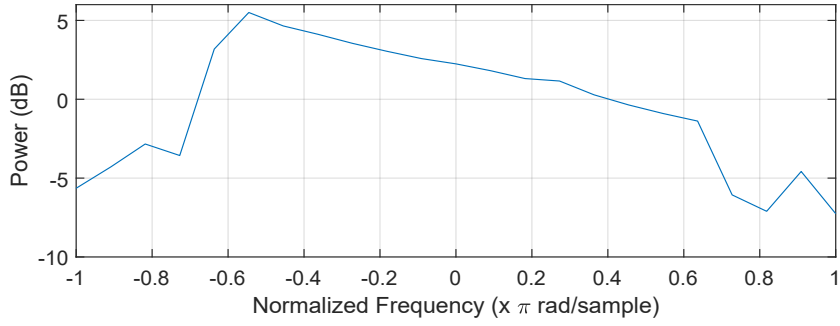
During the simulation, particular attention is given to observing the behavior of the LMS response in the spectral domain. Notably, it is observed that the LMS response exhibited a ramp inversely proportional to that of the filter. This inverse ramp indicates the effective functioning of the LMS equalizer, as shown in Figure 4.2b, counteracting the original filter's response depicted in Figure 4.2a.

However, it should be noted that the LMS equalizer primarily corrects over the input signal bandwidth. This limitation is highlighted by the fact that the correction ends at 0.55 rad. Beyond this range, which signifies the extent of the input signal bandwidth, the signal tends to be attenuated. Therefore, while the LMS response appears to flatten the frequency response as shown in Figure 4.2c, this compensation

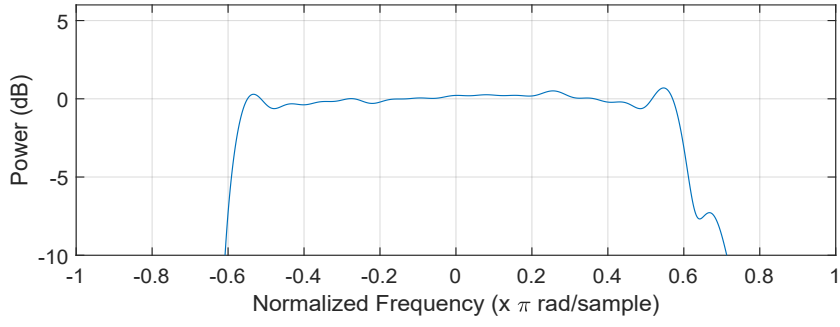
is particularly effective only within the confines of the input signal bandwidth.



(a) Normalized linear filter spectrum.



(b) The LMS algorithm response as a function of frequency.

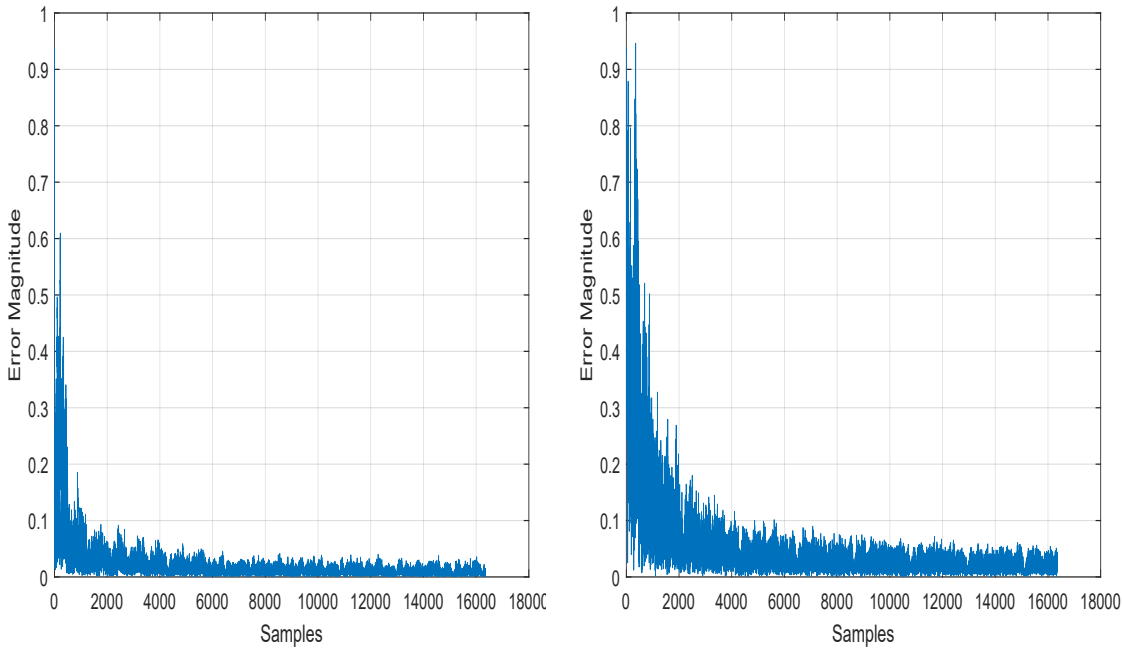


(c) Convolution of the adaptive equalizer response and the linear filter

Figure 4.2: Simulations demonstrating the impact of the equalizer.

Figure 4.3 presents the convergence of the error magnitude during the equalization process for α values of 0.05 and 0.005. A more rapid decrease in error magnitude across successive iterations is observed with an alpha value of 0.05, indicating faster convergence due to accelerated adaptation of the filter coefficients. On the other hand, an α value of 0.005 results in a slower rate of convergence, emphasizing the

impact of the alpha parameter on the efficiency of the equalization process.



(a) Convergence for $\alpha = 0.05$.

(b) Convergence for $\alpha = 0.005$.

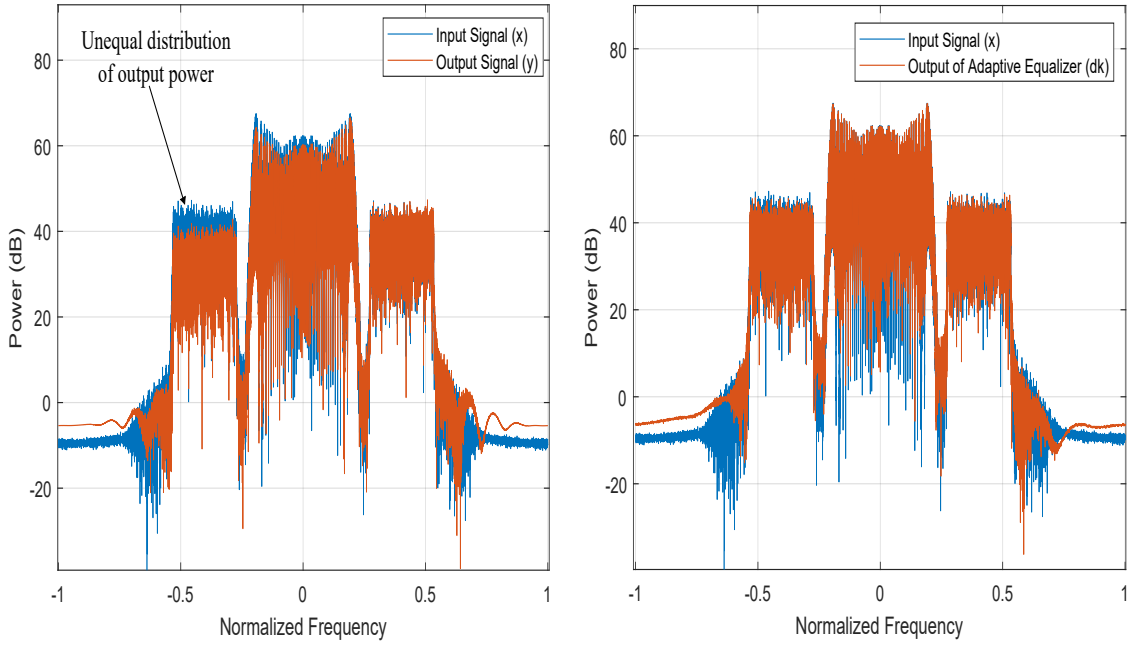
Figure 4.3: Error magnitude as a function of discrete time.

The comparison between the two α values illustrates the balance between convergence speed and system stability. While a larger α value can facilitate faster convergence and enhance the adaptability of the LMS algorithm, it may also risk system instability if increased excessively. Thus, carefully adjusting this parameter is crucial for maintaining an equilibrium between convergence speed and system stability.

Therefore, the convergence of the error magnitude not only serves as a quantitative measure of the algorithm's performance but also provides insight into the importance of proper parameter selection. This contributes to the effective identification and correction of distortions introduced by the linear filter, leading to a more accurate equalized output.

The equalization effect can be observed on the signal spectrum as shown in Figure 4.4. In Figure 4.4a which depicts the non-equalized output, the IBOC signal appears distorted and exhibits noticeable artifact, most notable being the sloped effect on the signal spectrum as a result of the frequency-dependent characteristics of the linear filter. In contrast, Figure 4.4b illustrating the equalized output showcases the significant improvement achieved by applying the LMS equalization algorithm.

The equalized IBOC signal exhibits a more accurate representation of the original transmitted signal.



(a) Spectrum with uncorrected output. (b) Spectrum with Corrected Output

Figure 4.4: Signal spectrum of input and output signals.

The findings of this simulation offer insights into the performance of the LMS linear filter equalization method for mitigating the effects of the linear filter on the IBOC signal. The results contribute to the understanding of equalization techniques and their effectiveness in compensating for distortions that may be induced by memory.

4.2 Compensation for Memoryless Nonlinearity and Memory Distortion

In the preceding section, a comprehensive exploration of the compensation for memory distortion was undertaken, employing a linear equalization strategy. The pursuit for enhanced signal clarity led to a series of advancements that successfully reduced the distortion introduced by system memory. However, efforts to ensure perfect signal transmission does not stop there. Proceeding, the focus is shifted toward compensating for a signal that exhibits both distortion properties.

This section aims to provide an extended examination into more complex strategies used for distortion mitigation, mainly using the memory polynomial model as

described in 3.2. However, before venturing into the compensation of distortion with more than one dimension, it is vital to recognize a fundamental approach utilized to counteract memoryless nonlinearity briefly.

Memoryless nonlinearity, often a culprit behind degraded signal quality in RF communication, can be counteracted using a simple yet effective tool known as a LUT. The response of a system with memoryless nonlinearity that has been corrected using a simple LUT is mathematically expressed as

$$\hat{x}[n] = x[n]G|x^2[n]|W|y^2[n]|, \tag{4.4}$$

where $\hat{x}[n]$ represents the reverse signal which is expected to be analogous to the forward input signal. The power of the amplifier's output signal as defined in (3.2) represents the indices of a compensation LUT W. A gain value of the LUT W as a specified index is represented as

$$W|y^2[n]| = \frac{\hat{x}[n]}{y[n]} = \frac{\hat{x}[n]}{x[n]G|x^2[n]|}, \tag{4.5}$$

which is the ratio of the expected reverse signal versus the actual output signal obtained from the system.

Our primary emphasis, though, will be on compensating for 2-dimensional distortion. This process provides a means of compensation for a non-linear time-varying system. The section shall look into the intricacies of the memory polynomial model, its implementation, and its performance in the context of RF communication. This review will bring forth a thorough understanding of both memoryless nonlinearity and memory distortion and the strategies used to manage it efficiently.

Fatungase et al [9], present a system architecture for DPD, as shown in Figure 2.5. The RFPA is modeled at baseband, with an equivalent local oscillator frequency of $f_{RF} = 0$. To evaluate the DPD requirements, a filter designed using the process described in 3.1.2 is cascaded with a LUT that implements the nonlinear distortion in the PA model. The system is fed with a realistic input signal, and the DPD is trained to minimize the error between the observed signal $\hat{x}$ and the expected signal x. To reduce the memory depth of the adaptive filter, a delay m_x is introduced to the input signal.

This section discusses the generation of the coefficients of the memory polynomial model as well as the correction of distortions using the inverse coefficients of the memory polynomial, specifically focusing on the use of direct matrix inversion to

obtain the coefficients [9]. Furthermore, simulations are run to estimate the resources required to implement the DPD.

4.2.1 Estimating Coefficients for Memory Polynomial Linearization

The direct matrix inversion process is efficient for estimating memory polynomial linearization coefficients. (3.8) expresses the output of an amplifier as the matrix multiplication between the input transformation matrix and the coefficients vector of the model c_{PA}. In a similar mathematical manner, the input of the system can also be expressed as

$$x[n] = \mathbf{Y}[n]\mathbf{c_{dpd}}, \tag{4.6}$$

where $\mathbf{c_{dpd}}$ is the inverse coefficient used for correction in the model. $\mathbf{Y}[n]$ represents the output transformation matrix at a discrete time n [9]. Then, the set of pre-correction weights for the same system can be expressed as

$$\mathbf{c_{dpd}} = \mathbf{Y}^{-1}[n]x[n]. \tag{4.7}$$

A direct matrix inversion process, as previously described, facilitates the efficient computation of the DPD coefficients. Central to this method is the assumption that the output transformation matrix of the system, which can be characterized as a time-dependent complex-valued matrix, does not have linearly dependent rows. This means each row should be distinct and not derived as a combination of other rows, a condition vital for direct matrix inversion. In practice, this may not always be the case as the nature of the signal and the system condition could influence the characteristics of the transformation matrix. In situations where the transformation matrix can not be directly inverted, using a generalization of the transformation matrix inverse might offer a viable alternative solution as explained in [19].

Some other challenges associated with direct matrix inversion is the computational complexity and potential instability when working with large or ill-conditioned matrices [11]. The method breaks down in such scenarios. The inversion of an ill-conditioned matrix can lead to numerical instability and non-negligible errors in the estimated coefficients. The direct inversion of large matrices may not be feasible for all real-time applications.

Despite these challenges, direct matrix inversion still provides a simple and straightforward solution to coefficient estimation for specific cases. For instance, direct inversion may offer a quick and efficient means to obtain DPD coefficients if the system is carefully designed to ensure a reasonable matrix size and a well-conditioned system matrix.

4.2.2 Block-Based Iterative Estimation for Wideband Signal

In the realm of high-rate wideband signals for RF communication purposes, it becomes imperative to optimize coefficient estimation methods for efficient and computational costs. Direct matrix inversion, although powerful, can become very expensive for large matrices. Signals such as the OFDM or the IBOC require a long train of sampled data to represent correctly. An iterative process that uses a static piecewise function that partitions the signal into smaller, manageable blocks can be used [12, 18]. This approach reduces the computational requirements and provides flexibility for implementation in a real-time process. Wideband signals with a higher bandwidth demand even higher sampling rates for accurate representation and processing, before applying DPD. Up-sampling a signal before applying DPD allows for the capture of more signal details which provides a smoother transition between data points and reduces the risk of further corrupting the intended signal, This increases the size of the matrix being processed. By adopting an iterative approach, the impact of up-sampling on computational resources can be significantly mitigated. The proposed technique uses the SGD method for estimating memory polynomial coefficient [22]. The mathematical equation for coefficient estimation is written as

$$c_{dpd,i} = c_{dpd,(i-1)} + \mu(Y^H Y)^{-1} Y^H e, \tag{4.8}$$

Here, $c_{dpd,i}$ represents the coefficients of the DPD model at the i^{th} iteration. The iteration begins from the initial estimate of the coefficients, which are recalculated in each successive iteration. The correction term which is added after each iteration defines the basis of this approach. Here, μ represents the learning rate of the algorithm which controls the speed of convergence. The learning rate is multiplied by the product of the Pseudo-inverse of the auto-correlation of output transformation matrix Y and the cross-correlation between the output and the error e. Where, e is equal to the difference of the input signal x from the output signal obtained with the previous set of pre-distortion coefficient $c_{dpd,(i-1)}$.

This method offers several benefits including its computational efficiency, an attribute that's particularly important for wideband applications. Its iterative nature allows it to optimize the coefficients to suit the signal characteristics better. However, this method may require careful selection of the step size, μ, to ensure stability and ultimately convergence of the iterations.

The efficacy of the iterative block-based memory polynomial coefficient estimation method is demonstrable via MATLAB simulation. In the simulation, an IBOC signal is employed as the test signal. This choice is influenced by the real-world applicability for high-throughput use cases, especially in FM radio. The signal has been carefully designed to respect the FCC's standards for FM radio, particularly with respect to the signal bandwidth. The IBOC signal is upsampled in the simulation to offer a higher resolution, which is essential for observing the behavior in adjacent channels.

This simulation uses methods discussed in Section 3.2 to obtain coefficients representing the amplifier model. The output transformation matrix Y can be estimated using the input signal and the obtained coefficient. The gain curve used to model the memoryless nonlinearity is maintained for all simulations, but the slope of the LTI filter used to model the memory distortion is varied to simulate a range in memory impairment. Filters demonstrating an in-band slope of 0.5 dB/MHz and 1 dB/MHz serve as the basis of this experiment.

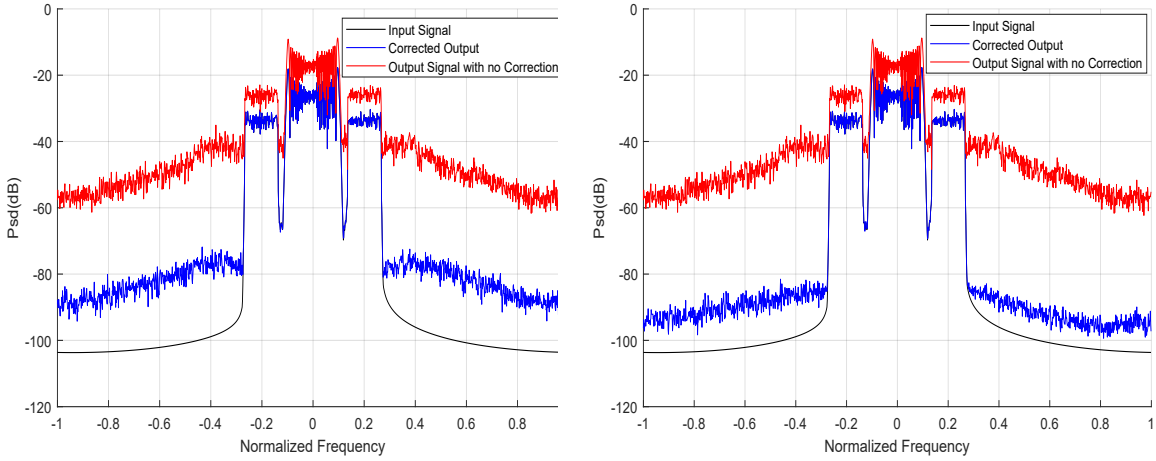
Identifying the optimal orders for the system's coefficients demands an MSE comparison across models of differing orders. This crucial step, involving the MSE evaluation, ensures that the chosen orders accurately reflect the system's behavior. Table 3.1 reveals that with a model that has a slope of 1 dB/MHz, the coefficient orders of $K = 4$ and $M = 3$ generate the smallest MSE, indicating optimum system performance. Consequently, these are chosen as the orders for the system's coefficients in this simulation.

The inclusion of memory effects is pivotal for this study, and the simulation contrasts the outcomes when memory terms are employed against scenarios where they are omitted. The influence of memory effects on signal recovery and overall system performance will be discussed in the forthcoming analysis.

Figures 4.5 and 4.6 provide the results for simulations carried out. They provide critical insight into the efficacy of the iterative block-based memory polynomial coefficient estimation method. Two major observations can be drawn from these results, one on the difference in slopes used to model memory, and the other related to the impact of memory compensation.

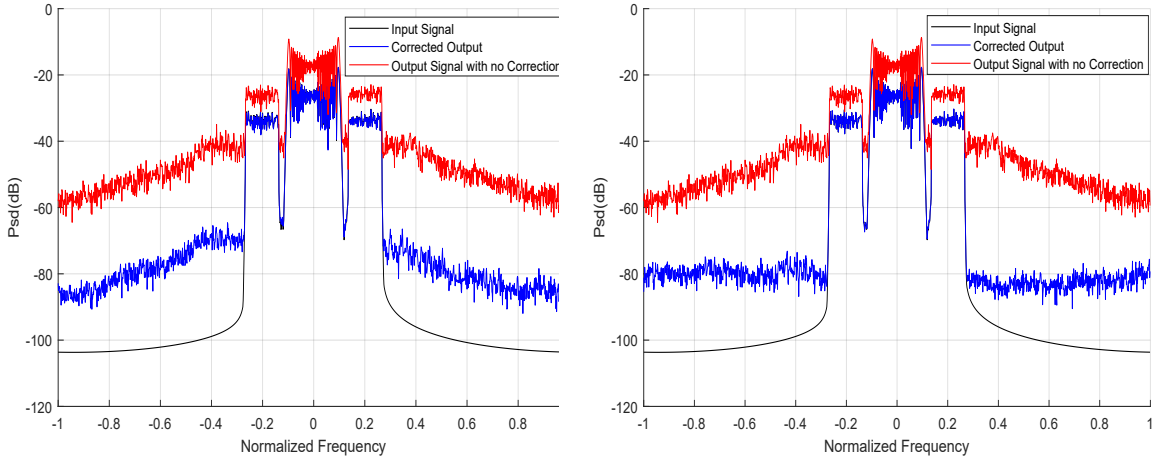
When comparing the results for the different slopes used to model memory, it can

be seen that the system exhibits different responses. A higher slope tends to have a more pronounced memory impairment, affecting the performance of the memory polynomial model and the quality of the corrected signal.



(a) Memoryless corrected signal spectrum. (b) Memory corrected signal spectrum.

Figure 4.5: Iterative Simulation with a Frequency Selectivity of 0.5dB/MHz



(a) Memoryless corrected signal spectrum. (b) Memory corrected signal spectrum.

Figure 4.6: Iterative Simulation with a Frequency Selectivity of 1dB/MHz

Furthermore, the results highlight the significant role of memory terms in the system's performance. By comparing results for simulations with and without memory terms, it becomes clear that omitting memory terms can lead to sub-optimal system performance. When the memory terms are introduced the system can adequately compensate for the memory distortion, resulting in a more accurate signal representation. Conversely, when these terms are neglected, the system fails to adequately

correct for the memory distortion, thereby impacting the quality of the recovered signal.

Table 4.1 presents the ACLR value of the simulation run in this context. ACLR, a measure typically used in certain communication standards, is used here to compare the power in the adjacent frequency band (200-600 kHz) to that in the desired channel (-200 to +200 kHz). The table denotes a decreasing ACLR as the slope of the frequency-dependent filters increase. The table also delineates a significant increase in ACLR when the simulation considers memory.

Table 4.1: ACLR test results in dB

	Memoryless Correction	Memory Correction
0.5 MHz	-54.0	-57.1
1 MHz	-49.9	-51.2

These comparisons between scenarios with and without memory terms underline the importance of including memory effects in the model. The necessity for an accurate and comprehensive model for effective distortion correction is emphasized, especially in high-throughput applications like FM radio transmission. The simulation results demonstrate the potential of the iterative block-based memory polynomial coefficient estimation method in handling wideband signals.

Chapter 5

Software Defined Radio System Implementation

This chapter attempts to demonstrate the efficiency of the DPD algorithm discussed thus far in this thesis in a practical implementation. The system has a software-defined radio setup and a RFPA with a frequency response operating in the VHF band, the system architecture is described and the purpose of individual components are expanded upon. Following this, the experimentation procedures are elaborated, and the results are discussed.

5.1 System Setup in the VHF Band

This section presents a hardware implementation to validate the theoretical constructs and simulations delineated in the preceding chapters. This real-world demonstration provides validation to findings under practical constraints and conditions. Furthermore, it offers the opportunity to identify any potential discrepancies or additional factors that might remain undetected in purely theoretical or simulated environments. Here, an experimentation setup for transmitting an IBOC signal in the VHF band is elaborated on.

The system setup shown in Figure 5.1 primarily consists of three integral components: a workstation computer, an Ettus B200 Software Defined Radio (SDR) board, and an RFPA provided by industry partners.

The workstation, running Ubuntu, hosts GNU Radio, an open-source toolkit that provides signal processing block functions to implement SDR applications. To enable interaction with the Ettus B200 SDR board, the USRP Hardware Driver (UHD) package includes the UHD Application Programming Interface (API), which is a library that applications use to interface with the SDR board. The Ettus B200 is a compact SDR board that operates with a frequency ranging from 70 MHz to 6 GHz. It integrates a direct conversion transceiver providing up to 56 MHz of real-time bandwidth, alongside a re-programmable Spartan6 Field Programmable Gate Arrays (FPGA). This device is adaptable to various applications, making it an ideal tool for versatile radio frequency experimentation. The system under test consists of two cascaded power amplifiers with a combined nominal gain of 42 dB. They derive

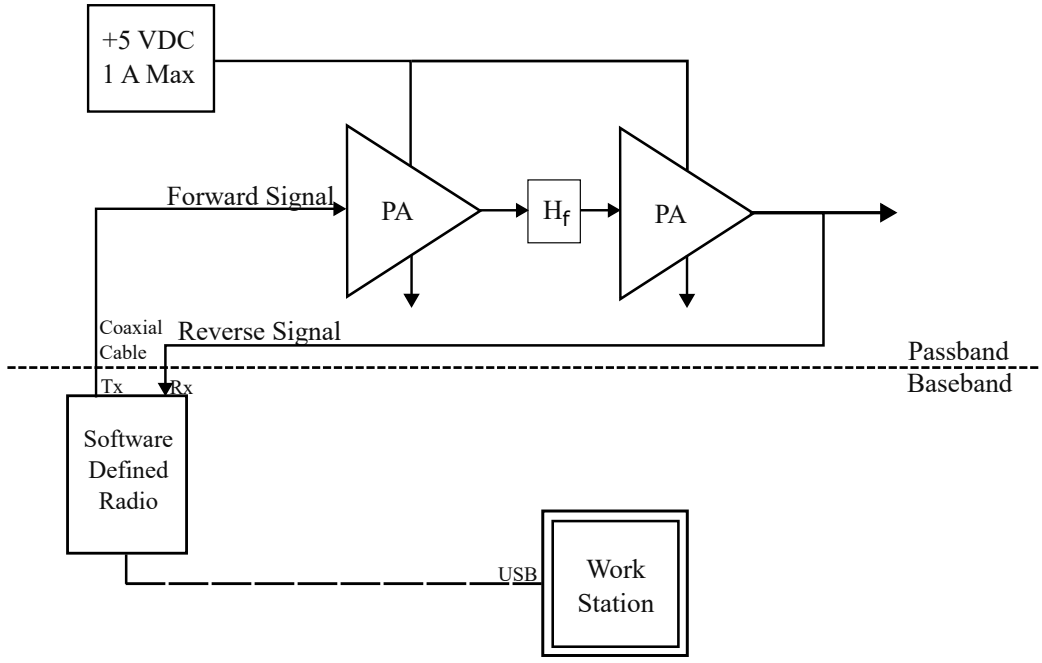


Figure 5.1: Experimentation system architecture.

power from a shared 5 VDC supply, and their combined maximum input direct current is 1 A. It is worth noting that Figure 5.1 denotes a block H_f between the 2 gain stages of the amplifier system, which is where memory distortion is expected. This system differs from the Wiener model developed in 3.1 in that it includes two gain stages surrounding the memory distortion. However, the first PA operates at such a low power level that it is assumed to be operating in a linear region and not contribute significant non-linearity. The model defined in 3.1 is assumed to apply for this topology, as well.

The signal utilized in this experiment is read from a file on the computer and is streamed repeatedly into GNU Radio, which feeds the signal into a Universal Software Radio Peripheral (USRP) sink block.

A picture of the actual setup is shown in Figure 5.2. The transmission (Tx) port of the USRP device is connected to the input of the PA. This PA supplied by the industry partner, amplifies the signal, which is then transmitted. Through a directional coupler, the output of the amplifier is wired to the receiver (Rx) of the USRP device, which captures the amplified signal. In GNU Radio, a USRP Source block is employed to receive this signal from the Rx of the USRP device. The forward and reverse signals are streamed to a Matlab data file of the workstation for further observation and analysis in Matlab. This system setup and configuration are the

foundation for the complex theory and experiments discussed in this thesis.

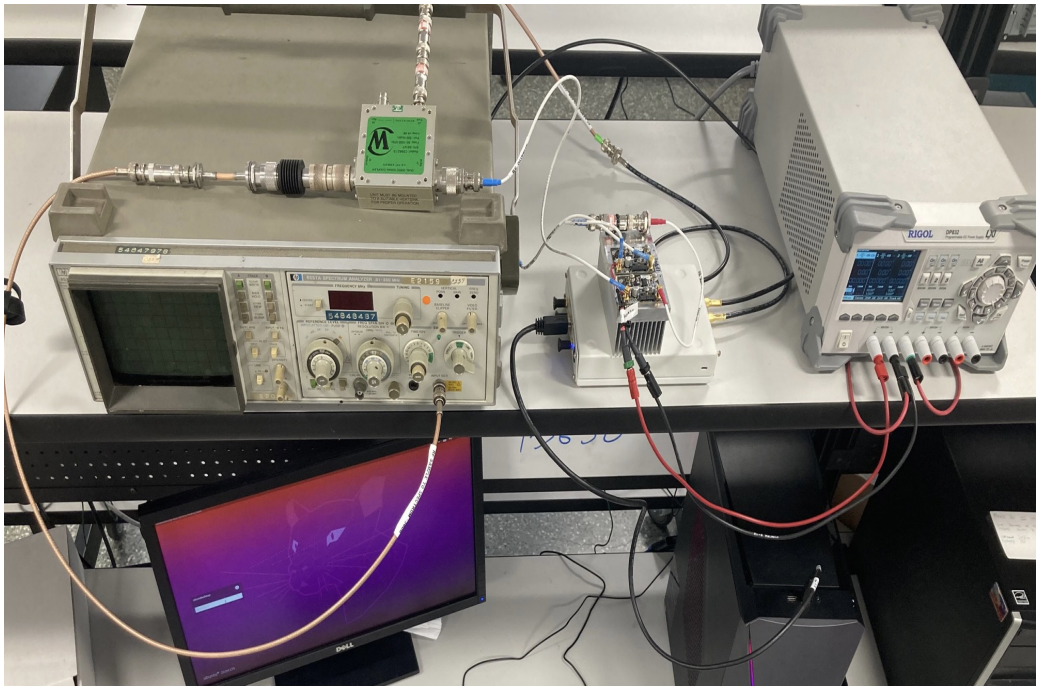


Figure 5.2: Experimentation system setup.

5.1.1 Single Tone Test

A series of specific experiments are carried out to explore the conditions and properties of the amplifier using a single tone. The initial experiment uses a modulated ramp signal as the amplifier input signal. This experiment demonstrates the amplifier's operational capabilities at a carrier frequency of 80 MHz. This information is critical as it outlines the boundary of the amplifier's linear region, a zone where the amplifier functions optimally without introducing distortion.

Figure 5.3 presents the gain curve of a single sawtooth sampled at 3 MHz input signal capture with a 0.1 to 1 amplitude range, applied over discrete samples at a carrier frequency of 80 MHz. The system gain (G) is calculated as the ratio between the output and the input signal expressed in dB. Thus, the observed changes in gain directly reflect the differences between the two signals. The figure illustrates the 1 dB compression point, which is used to determine the amplifier operating range for the experiment.

Following this, a frequency sweep is performed by applying a low amplitude single tone, transmitted at a carrier frequency ranging from 75 MHz to 90 MHz with a

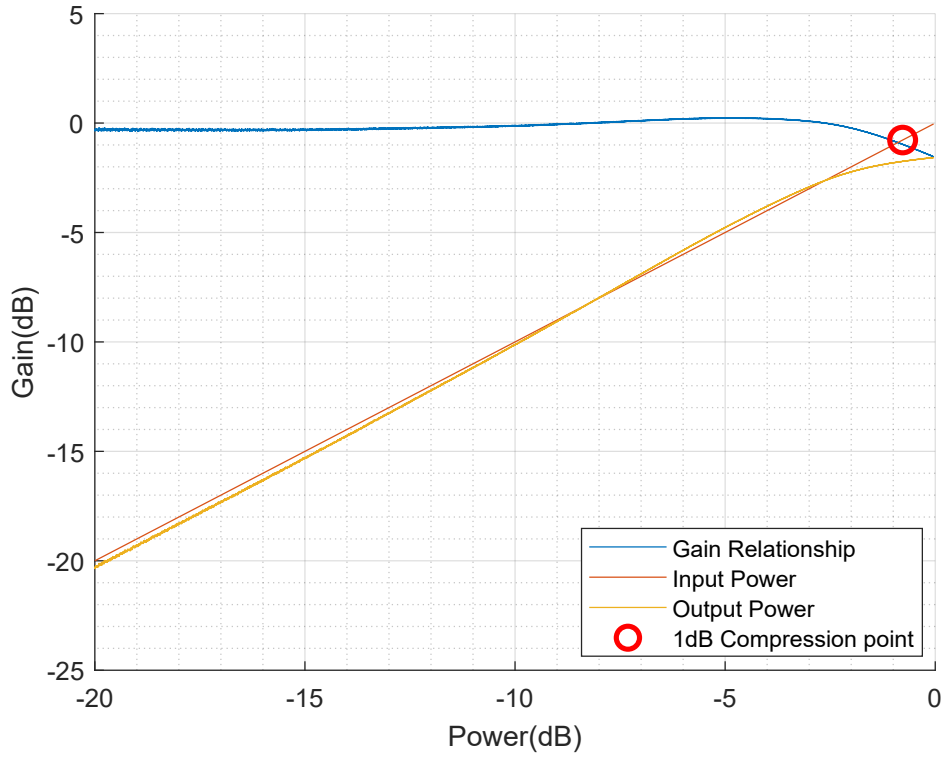


Figure 5.3: Amplitude sweep.

resolution of 100 kHz. Using this signal, no memoryless nonlinear distortion is expected because the signal is a single tone. This test allows to produce the frequency response of the amplifier.

Figure 5.4 presents the system gain as a function of frequency, illustrating the systems' memory characteristics. It is observed that the gain varies across the frequency spectrum, demonstrating the memory effect in the amplifier's behavior. Most notably, the spectrum exhibits a benign slope of approximately 0.5dB across 10 MHz, demonstrating the frequency response to some extent. However, it's interesting to note a relatively flat region between 85 MHz and 87 MHz. This suggests a frequency range where the system's memory distortion characteristics are minimal. It is interesting to note that the frequency response behaves differently as the amplifier approaches operation in compression, as observed by the PA Sweep in Compression plot in 5.4.

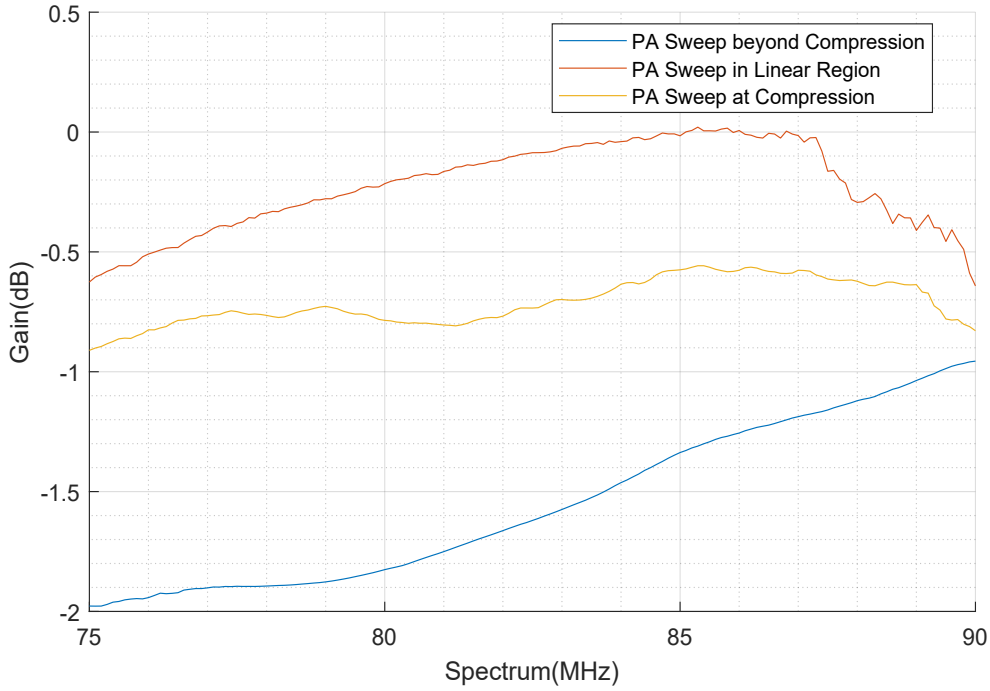


Figure 5.4: Gain as a function of frequency

5.1.2 Full Signal Transmission

In this experiment, the USRP transmit gain is set to a constant value of 55 dB across all trials. This configuration standardizes the experiment conditions, ensuring a fair comparison of results obtained from different tests.

The signal utilized in this setup is an IBOC signal, up-sampled by a factor of four for increased resolution. The sample rate selected for operation must be managed as excessive sample rate may cause inconsistent behavior, such as data-rate mismatch between the device and the workstation. The test signal, sampled at 3 MHz, was transmitted at a carrier frequency of 80 MHz in this experiment. It is worth noting that the signal is upsampled to 3 MHz because a linearization bandwidth that is greater than the Nyquist of the signal is necessary for compensation using DPD [15]. Additionally, the amplitude of the signal is adjusted appropriately. The goal is to ensure the peak amplitude of the signal matches the 1-dB compression point of the amplifier allowing for a small amount of amplifier compression that typically still meet spectral mask requirements. Such a signal used with this system is expected to experience nonlinearity as well as memory distortion.

The signal reception uses a heterodyne receiver setup, a scheme frequently employed in modern radio frequency communication systems. This setup mixes the incoming RF signal with a locally generates signal of a different frequency. Specifically, in this setup, the USRP receiver down-converts from a carrier of 80 MHz using a 79.5 MHz Local Oscillator (LO) to an intermediate frequency of 500 KHz. In the process of this down-conversion, the heterodyne architecture also introduces an unwanted component at 79.5 MHz because the LO leaks directly into the output.

This undesired signal component at 79.5 MHz, results from the frequency mixing process and is a common characteristic of heterodyne receivers. The unwanted signal is subsequently removed by applying a low-pass filter. This ensures that only the desired signal and sufficient bandwidth for adjacent channels are preserved, maintaining the purity of the received signal. This setup allows the system to effectively process incoming signals while managing computational load efficiently [13].

Figure 5.5 offers a visual representation of the GNURadio Setup and interaction with the UHD. In this architecture, a heterodyne receiver setup is implemented. It can be observed that the reverse signal is mixed along with an intermediate frequency of 500 KHz and the forward signal is delayed to compensate for lag in the system. In the USRP device, the UHD FPGA performs a host of tasks most importantly the Interpolation or decimation of the signal for transmission and reception purposes respectively.

The delayed forward signal and the reverse are saved to a single Matlab data file on the workstation using a custom block provided by the industry partner. A major challenge while attempting to save results during the experiments was the occurrence of 'under-runs' on the USRP device, which refers to a situation where data is not supplied fast enough to the USRP device from the workstation for continuous transmission. In other words, the data rate between the workstation and the USRP device is not matching up, creating an unstable system. This effect is usually due to a disparity in processing speeds between the USRP and the workstation. If the workstation cannot keep up with the data demands of the USRP, perhaps due to a high sampling rate or inadequate processing capabilities, the smooth data flow to the USRP is interrupted.

In light of these challenges, an adjustment to the experiment is made. Instead of real-time streaming, an amplifier's forward and reverse signals are recorded in a controlled state, ensuring no under-runs occur. This requires that the workstation is not overloaded, reducing the probability of under-flows.

The optimal configuration of nonlinearity and memory coefficient orders used

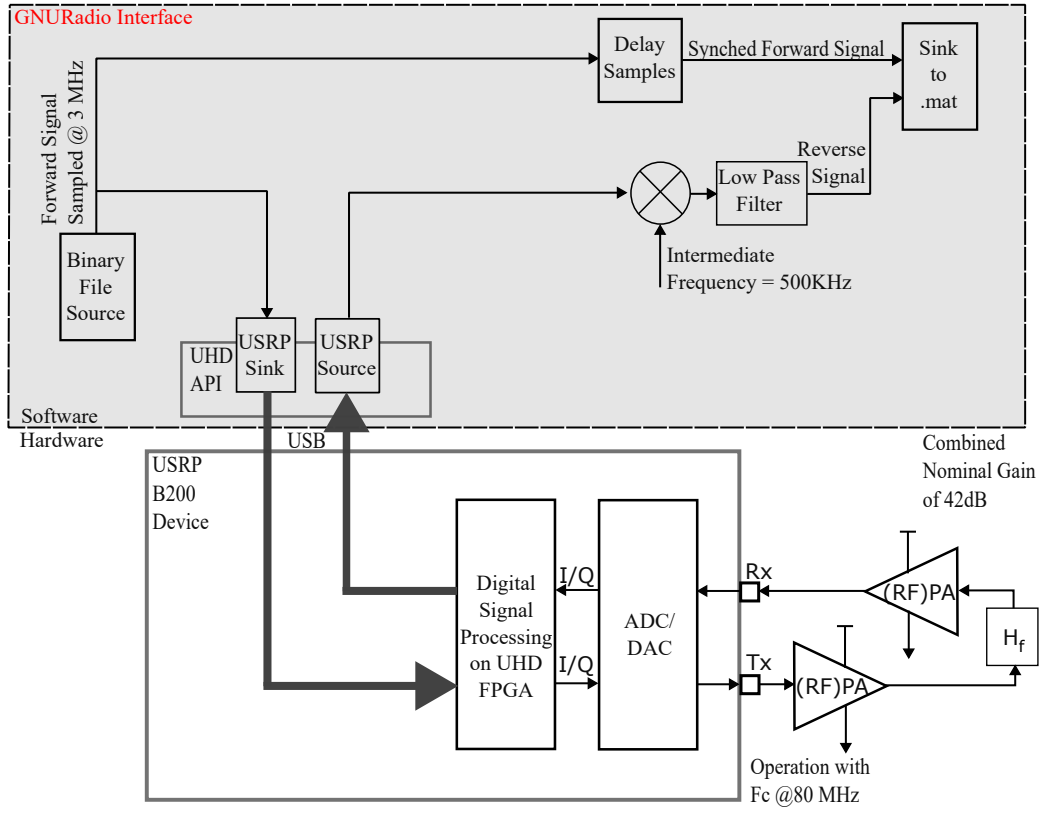


Figure 5.5: Block diagram showing UHD and GNURadio setup.

for the setup as described can be discerned from the MSE coefficient estimation method presented in this thesis. Here, a stream of data is processed through the system. Using the synchronized forward and reverse signal that is saved onto the workstation, the coefficient is estimated offline for varying orders of nonlinearity and memory. Following this, individual MSE's are calculated using the error between a computed output and the reverse signal. From Table 5.1, it is observed that the configuration with three memory coefficients and four nonlinear coefficients minimizes the MSE, suggesting this as the optimal configuration, as higher orders of coefficient than required offers no significant improvement with diminishing returns.

Table 5.1: MSE Coefficient estimation

Memory Terms	1	2	3	4
Nonlinearity Terms				
1	7.83e-05	7.25e-05	5.55e-05	5.45e-05
2	4.72e-05	4.24e-05	4.20e-05	4.13e-05
3	4.57e-05	4.15e-05	3.97e-05	4.02e-05
4	4.51e-05	4.06e-05	3.80e-05	3.89e-05
5	4.56e-05	4.14e-05	3.92e-05	4.01e-05

Employing the iterative block-based method, coefficients are estimated offline, and subsequently integrated with a new stream to produce a pre-distorted signal. This signal is the new input, with its corresponding reverse signal re-recorded. The outcomes from these tests provide insightful data for understanding the performance and implications of the memory polynomial pre-distortion method.

The figures shown present the results of experiments carried out. Figure 5.6 shows the result for simulation carried out without memory coefficient $M = 1$ and Figure 5.7 presents results of a simulation that utilizes three orders of memory $M = 3$. Here, both experiments use the same amplifier model and hardware setup. Each plot shows four spectra, indicating the forward signal, the reverse signal, the simulated results, and the experimental results. Here, the simulation results represent a post processed signal which used data that was saved for estimation of coefficient. The experimental result was the data saved using the predistorted signal as the input to the experimental setup.

A noteworthy observation from the results is that simulations using a memory order of 3 show improved outcomes. However, this improvement in simulation does not strongly reflect in the experimental results. Despite this, there is a demonstrated capability to correct memoryless nonlinearity. Table 5.2 further supports this point,

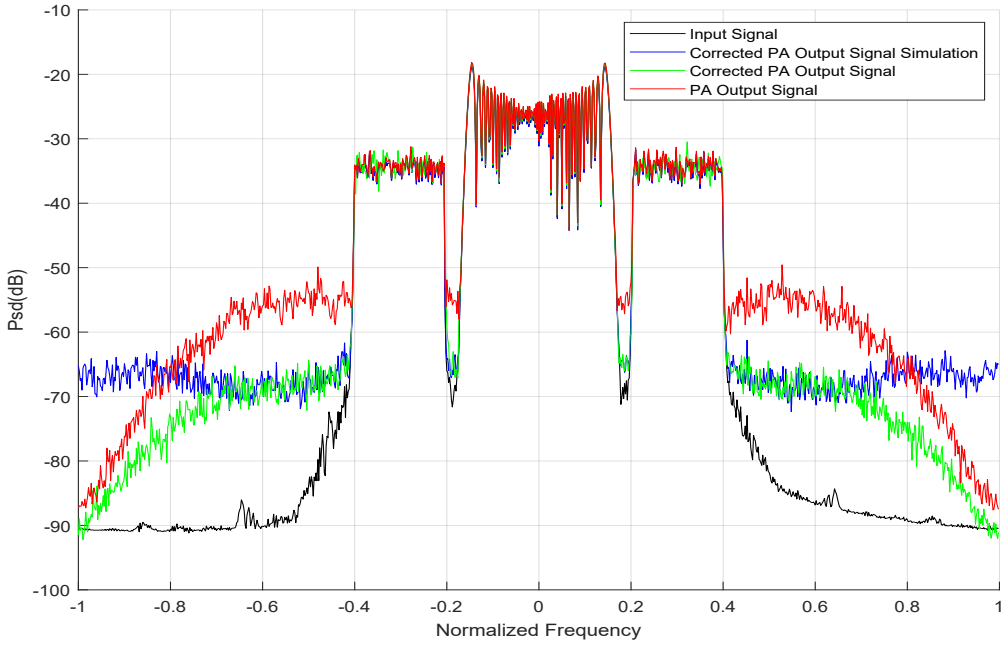


Figure 5.6: Spectrum of memoryless Experimentation results.

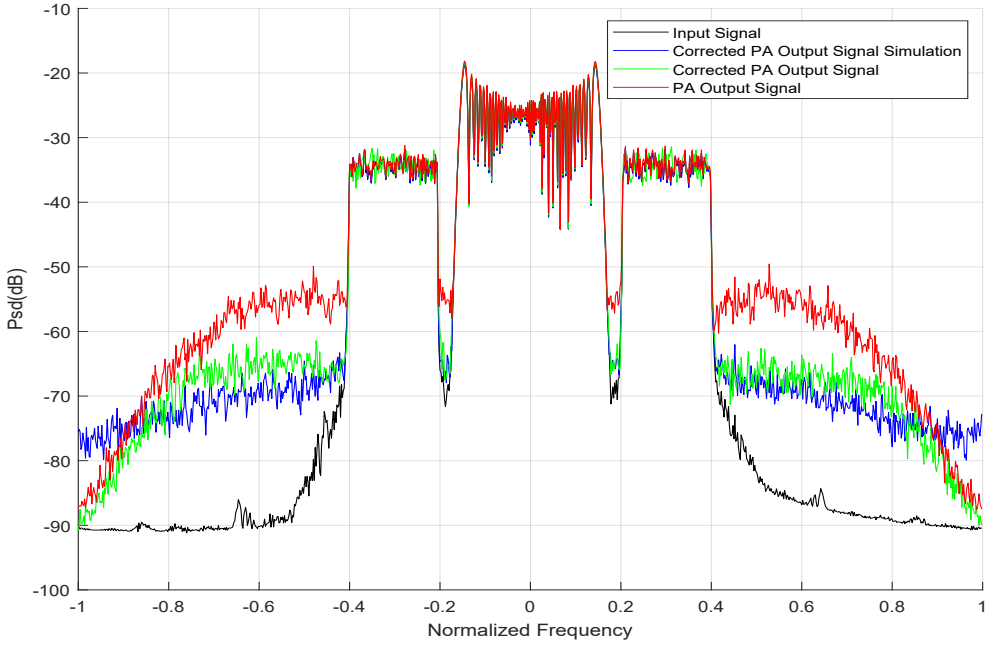


Figure 5.7: Spectrum of Experimentation results with memory correction.

as it presents the ACLR of the experiments achieved. The context defined in 4.2.2 for measuring the ACLR remains consistent. Here, a significant reduction in the ACLR is observed when observing a memory corrected signal from the output of the PA.

Table 5.2: Experimentation ACLR results in dB

	Memoryless Correction	Memory Correction
Simulated	-40.9	-42.5
Experimental	-39.2	-38.9

The measurement results do not allow us to make a clear conclusion about the improvement of the performance in presence of memory. This could be attributed to different limitations in the measurement setup. For example, a linearization bandwidth of 3 MHz was used in the experiment, but in actuality, using a lowpass filter reduced this linearization bandwidth to 800 kHz. It is possible that this filter, while necessary to remove unwanted lower-frequency components, might also be eliminating vital information necessary for adequate memory compensation. It is also worth noting that the frequency response between the two amplification stages may have been too benign to present an issue. The frequency responses modeled in earlier chapters presented a more pronounced slope of 0.5-1 dB/MHz compared to what was observed from the frequency sweep in 5.4. This preliminary test used to inform the experimental setup showed that the frequency response between the amplification stages had a slope of 0.06d dB/MHz. Furthermore, the system demonstrates a limitation regarding its sampling rate. Increasing the sampling rate might enhance the correction performance but also lead to system instability. This constraint suggests that the existing configuration might not be ideal for tacking memory distortion, mainly when the system cannot operate effectively at higher sampling rates.

Chapter 6

Conclusion

This thesis analyzes the challenges power amplifiers face in RF communications, examining the intricacies of wideband signal transmission and the implications of high PAPR. The progression led to an exploration of power amplifier modeling, detailing cascaded distortion properties and introducing frameworks such as the memory polynomial model. The subsequent sections delved into pre-distortion strategies, from basic linear filter equalization to more complex compensation methods, and after that was the hands-on implementation of a software-defined radio system in the VHF band.

The objective of this work was to craft a holistic end-to-end model of a power amplifier, integrating both computational and practical elements. The memory polynomial pre-distortion method stood out as a critical solution, with simulation results reinforcing its effectiveness in enhancing system performance.

While the overall experiment achieved its primary objectives, we observed some considerations around memory distortion. The use of the low pass filter, which was necessary for the heterodyne receiver setup, may have hindered correction to some degree. The results suggest that while these setups are necessary for the system's functioning, they might also introduce complexities and remove vital information necessary for adequate memory compensation.

While it's observed that the current configuration and setup could benefit from significant enhancements, it's also noteworthy to understand the specific constraints of the system in the real-time context. The PA device under test, for instance, has been identified as not having significant memory, which presents a discrepancy with the models that exhibited more severe memory distortion. This implies that while the theoretical aspects of the research underscore the necessity for further refinement in this area, practical implementation might require a more measured approach.

With these observations in mind, future endeavors may aim to address these discrepancies, seeking alternative setups or methodologies that are more accurately aligned with the performance of the actual device under test. Notably, the system's measured slope is about 0.5 dB/10 MHz, significantly smaller than the models' most

moderate tests. Therefore, one potential avenue of exploration to confirm the theory discussed in this thesis could be the use of a signal with a wider bandwidth, as such a signal is more susceptible to frequency selectivity. It would be interesting to investigate the possibilities of achieving this within the framework of GNU Radio. This highlights the potential for improvements on the current experimental setup and paves the way for future real-time optimization and exploration.

Despite these identified limitations and potential refinements, it remains essential to highlight that the research conducted and presented in this thesis significantly contributes to understanding the complexity involved in distortion mitigation for wideband signals in RF applications. It provides a basis for future work, elucidating the path to integrate real-time considerations and practical device constraints into the theoretical framework and simulations, ultimately pushing towards a more practical and applicable system design.

Bibliography

[1] S. Ali, R. Ramzan, and S. Azam. High efficiency 88–108MHz, 25W class-E PA for transmitters in smart cities. In *2017 International Conference on Electrical and Computing Technologies and Applications (ICECTA)*, pages 1–4, 2017.

[2] D. K. Choi and S. I. Long. A Physically Based Analytic Model of FET Class E Power Amplifiers Designing for Maximum PAE. *IEEE Transactions on Microwave Theory and Techniques*, 47:1712–1720, 1999.

[3] S. Choi, E. Jeong, and Y. H. Lee. Adaptive predistortion with direct learning based on piecewise linear approximation of amplifier nonlinearity. *IEEE Journal of Selected Topics in Signal Processing*, 3(3):397–404, 2009.

[4] P. B. Crilly and A. B. Carlson. *Communication Systems*. McGraw-Hill Education, 2009.

[5] G. A. Davidson, M. A. Isnardi, L. D. Fielder, M. S. Goldman, and C. C. Todd. ATSC Video and Audio Coding. *Proceedings of the IEEE*, 94(1):60–76, 2006.

[6] G. de Boer, C. Kupferschmidt, D. Bederov, and H. P. Kuchenbecker. Digital Audio Broadcasting in the FM Band Based on Continuous Phase Modulation. *IEEE Transactions on Broadcasting*, 49(3):293–303, 2003.

[7] L. Ding. A robust digital baseband predistorter constructed using memory polynomials. *IEEE Transactions on Communications*, 52(1):159–165, 2004.

[8] P. Drotar, J. Gazda, D. Kocur, and P. Galajda. MC-CDMA Performance Analysis for Aifferent Spreading Codes at HPA Saleh Model. In *2008 18th International Conference Radioelektronika*, pages 1–4, 2008.

[9] I. B Fatungase and J.-F. Bousquet. Real-time Power Amplifier Impairment Assessment using Nonlinear Polynomial with Memory. In *2022 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 890–894, 2022.

[10] F. M. Ghannouchi and O. Hammi. Behavioral modeling and predistortion. *IEEE Microwave Magazine*, 10(7):52–64, 2009.

[11] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, 2012.

[12] J. Guan, X. A. Nghiem, and R. Negra. Iterative characterisation approach using realistic excitation signals for linearisation of transmitters. In *2014 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 197–200, 2014.

[13] N. Hosseini and D. W. Matolak. Wide band channel characterization for low altitude unmanned aerial system communication using software defined radios. In *2018 Integrated Communications, Navigation, Surveillance Conference (ICNS)*, pages 2C2–1–2C2–9, 2018.

[14] H. Huang, P. Mitran, and S. Boumaiza. Digital Predistortion Function Synthesis using Undersampled Feedback Signal. *Microwave and Wireless Components Letters IEEE*, 26(10):855–857, 2016.

[15] Hai Huang, Jingjing Xia, and Slim Boumaiza. Novel parallel-processing-based hardware implementation of baseband digital predistorters for linearizing wideband 5g transmitters. *IEEE Transactions on Microwave Theory and Techniques*, 68(9):4066–4076, 2020.

[16] S. Kim, K. Park, K. Lee, and H. Choi. Detection method for Digital Radio Mondiale plus in hybrid broadcasting mode. *IEEE Transactions on Consumer Electronics*, 59(1):9–15, 2013.

[17] J. F. Kurose and K. W. Ross. *Computer Networking: A Top-down Approach.* Always learning. Pearson, 2013.

[18] Kwang-Pyo Lee, Soon-Il Hong, and Eui-Rim Jeong. A polynomial digital predistortion technique based on iterative architecture. *International Journal of Electrical and Computer Engineering (IJECE)*, 6:106–112, 02 2016.

[19] Y. Lei, B. Liao, and J. Chen. Limited-time convergent znn for computing time-dependent complex-valued matrix pseudoinverse. In *2019 9th International Conference on Information Science and Technology (ICIST)*, pages 479–484, 2019.

[20] Y. Liu, W. Pan, S. Shao, and Y. Tang. A General Digital Predistortion Architecture Using Constrained Feedback Bandwidth for Wideband Power Amplifiers. *IEEE Transactions on Microwave Theory and Techniques*, 63(5):1544–1555, 2015.

[21] MathWorks. Power amplifier - matlab & simulink, 2017. Accessed: 2022-05-22.

[22] D.R. Morgan, Z. Ma, J. Kim, M.G. Zierdt, and J. Pastalan. A Generalized Memory Polynomial Model for Digital Predistortion of RF Power Amplifiers. *IEEE Transactions on Signal Processing*, 54(10):3852–3860, 2006.

[23] M. O'Droma, S. Meza, and Y. Lei. New modified saleh models for memoryless nonlinear power amplifier behavioural modelling. *IEEE Communications Letters*, 13(6):399–401, 2009.

[24] W. Pan, Y. Liu, and Y. Tang. A predistortion algorithm based on accurately solving the reverse function of memory polynomial model. *IEEE Wireless Communications Letters*, 1(4):384–387, 2012.

[25] F. H. Raab et al. Class-F Power Amplifiers With Maximally Flat Waveforms. *IEEE Transactions on Microwave Theory and Techniques*, 45(11):2007–2012, 1997.

[26] T. S. Rappaport. *Wireless Communications: Principles and Practice*. Prentice Hall communications engineering and emerging technologies series. Prentice Hall PTR, 2002.

[27] B. Razavi. *RF Microelectronics*. Wiley, 2014.

[28] P. Roblin, C. Quindroit, N. Naraharisetti, S. Gheitanchi, and M. Fitton. Concurrent linearization. *IEEE Microwave Magazine*, 14(7):75–91, 2013.

[29] P. Schmid. *An improved method of peak-to-average power ratio reduction for FM + IBOC broadcast transmission*. Halifax, 2009.

[30] P. Schmid. Interleaving IBOC Signals for a Digital HD Radio Multiplex. *IEEE 7th Workshop on Signal*, 2015.

[31] J. S. Seybold. *Introduction to RF Propagation*. Wiley, 2005.

[32] W. Tsai, C. Liou, Z. Peng, and S. Mao. Intermodulation Distortion Analysis for Power Amplifier with Various Collector Voltages. In *2018 IEEE International Symposium on Radio-Frequency Integration Technology (RFIT)*, pages 1–3, 2018.

[33] F. Wang, J. Huang, X. Chen, and A. Men. Simultaneous Broadcasting of Analog FM and Digital Signals by Separating Co-Channel FM Signals. *IEEE Communications Letters*, 20(11):2197–2200, 2016.

[34] Y. Woo, J. Kim, J. Yi, S. Hong, I. Kim, J. Moon, and B. Kim. Adaptive digital feedback predistortion for linearizing power amplifiers. *IEEE Transactions on Microwave Theory and Techniques*, 55(5):932–940, 2007.

[35] J. Wood. *Behavioral Modeling and Linearization of RF Power Amplifiers*. Artech House microwave library. Artech House, 2014.

[36] G. T. Zhou. Analysis of spectral regrowth of weakly nonlinear power amplifiers. *IEEE Communications Letters*, 4(11):357–359, 2000.

[37] A. Zhu, P. J. Draxler, J. J. Yan, T. J. Brazil, D. F. Kimball, and P. M. Asbeck. Open-Loop Digital Predistorter for RF Power Amplifiers Using Dynamic Deviation Reduction-Based Volterra Series. *IEEE Transactions on Microwave Theory and Techniques*, 66(4):2017–2027, 2008.