Mixed-signal predistortion for small-cell 5G wireless nodes

Venkata Narasimha Manyam

▶ **To cite this version:**

Venkata Narasimha Manyam. Mixed-signal predistortion for small-cell 5G wireless nodes. Electronics. Université Paris Saclay (COmUE), 2018. English. NNT : 2018SACLT015 . tel-01997230

NNT : 2018SACLT015

Mixed-Signal Predistortion for Small-Cell 5G Wireless Nodes

Thèse de doctorat de l'Université Paris-Saclay
préparée à Télécom ParisTech

Ecole doctorale n°580 Sciences et technologies de l'information et de la communication (STIC)
Spécialité de doctorat : Réseaux, Information et Communications

Thèse présentée et soutenue à Paris, le 09 Novembre 2018, par

Venkata Narasimha Manyam

Composition du Jury :

Dominique Dallet	
Professeur, Bordeaux INP	Président
Geneviève Baudoin	
Professeur, ESIEE	Rapporteur
Myriam Ariaudo	
Maître de conférences HDR, ENSEA	Rapporteur
Philippe Meunier	
Ingénieur, NXP Semiconductors France	Examinateur
Patricia Desgreys	
Professeur, Télécom ParisTech	Directeur de thèse
Chadi Jabbour	
Maître de conférences, Télécom ParisTech	Co-directeur de thèse

Thèse de doctorat

Abstract

Small-cell base stations (picocells and femtocells) handling high bandwidths ($> 100\,\text{MHz}$) will play a vital role in realizing the 1000X network capacity objective of the future 5G wireless networks. Power Amplifier (PA) consumes the majority of the base station power, whose linearity comes at the cost of efficiency. With the increase in bandwidths, PA also suffers from increased memory effects. Digital predistortion (DPD) and analog RF predistortion (ARFPD) tries to solve the linearity/efficiency trade-off. In the context of 5G small-cell base stations, the use of conventional predistorters becomes prohibitively power-hungry.

Memory polynomial (MP) model is one of the most attractive predistortion models, providing significant performance with very few coefficients. We propose a novel FIR memory polynomial (FIR-MP) model which significantly augments the performance of the conventional memory polynomial predistorter. Simulations with models extracted on ADL5606 which is a $1\,\text{W}$ GaAs HBT PA show improvements in adjacent channel leakage ratio (ACLR) of $7.2\,\text{dB}$ and $15.6\,\text{dB}$, respectively, for $20\,\text{MHz}$ and $80\,\text{MHz}$ signals, in comparison with MP predistorter. Digital implementation of the proposed FIR-MP model has been carried out in $28\,\text{nm}$ FDSOI CMOS technology. With a fraction of the power and die area of that of the MP a huge improvement in ACLR is attained. An overall estimated power consumption of $9.18\,\text{mW}$ and $116.2\,\text{mW}$, respectively, for $20\,\text{MHz}$ and $80\,\text{MHz}$ signals is obtained.

Based on the proposed FIR-MP model a novel low-power mixed-signal approach to linearize RF power amplifiers (PAs) is presented. The digital FIR filter improves the memory correction performance without any bandwidth expansion and the MP predistorter in analog baseband provides superior linearization. MSPD avoids 5X bandwidth requirement for the DAC and reconstruction filters of the transmitter and the power-hungry RF components when compared to DPD and ARFPD, respectively. The impact of various non-idealities is simulated with ADL5606 ($1\,\text{W}$ GaAs HBT PA) MP PA model using $80\,\text{MHz}$ modulated signal to derive the requirements for the integrated circuit implementation. A resolution of 8 bits for the coefficients and a signal path SNR

of 60 dB is required to achieve ACLR1 above 45 dBc, with as little as 9 coefficients in the analog domain. Discussion on the potential circuit architectures of subsystems is provided. It results that an analog implementation is feasible. It will be worth in the future to continue the design of this architecture up to a silicon prototype to evaluate its performance and power consumption.

Acknowledgements

First and foremost, I would like to express my deepest and sincere gratitude to my supervisor Prof. Dr. Patricia Desgreys and co-supervisor Dr. Chadi Jabbour for entrusting me and providing me (an analog IC designer) with an interesting digital system challenge resulting in a *nouvelle* mixed-signal solution. Their unparalleled supervision at each step during the thesis has positively molded me and taught me many things for a lifetime, be it research methodologies, presentation skills or scientific writing. I feel immensely proud to have worked at Télécom ParisTech on the future 5G telecommunications, where the word *telecommunication* was first coined by a Télécom ParisTech alumni Édouard Estaunié in the year 1904.

My sincere thanks to Dr. Dang-Kièn Germain Pham for mentoring me and getting me started with the MATLAB codes. I have thoroughly benefited from his meticulous feedback and technical discussions during my thesis. I used to learn a new thing on each and every visit to his office.

I would also like to thank all the members of the jury, Prof. Geneviève Baudoin, Prof. Dominique Dallet, Dr. Myriam Ariaudo and Dr. Philippe Meunier for their invaluable feedback on my work. It is an honor to be able to have such experts in the jury. I was always inspired by their research work in this field.

I am particularly grateful to Prof. Yves Mathieu and Dr.Tarik Graba at the Safe and Secure Hardware (SSH) group at the COMELEC (Communications & Electronics) department for the help with RTL codes and digital ASIC implementation flow.

I am very thankful to the faculty and colleagues at the COMELEC department, especially Circuits & Communications systems (C2S) group, Prof. Patrick Loumeau, Dr. Hervé Petit, Dr. Hussein Fakhoury, Kelly Tchambake, Dr. Elias Solieman, Dr. Reda Mohellebi, Dr. Minh Tien Nguyen, Dr. Han Le Duc, Dr. Yosra Gargouri, Dr. Raphaël Vansebrouck, Dr. Ta Duc-Tuyen, Dr. Chetan Joshi, Dr. Manuj Mukherjee, Dr. Sumanta Chaudhuri for the interesting discussions and camaraderie.

Contents

List of Figures

List of Tables

Acronyms

3GPP	3rd Generation Partnership Project
AAF	Anti-Aliasing Filter
ACLR	Adjacent Channel Leakage Ratio
ACPR	Adjacent Channel Power Ratio
ADC	Analog-to-Digital Converter
AFE	Analog Front End
AM	Amplitude Modulation
APD	Analog Predistortion
ARFPD	Analog Radio Frequency Predistortion
ASK	Amplitude Shift Keying
BB	Baseband
BBU	Base-Band Unit
BS	Base Station
BTS	Base Transceiver Station
CA	Carrier Aggregation
CC	Component Carrier
CDMA	Code Division Multiple Access
CFR	Crest Factor Reduction
CMOS	Complementary Metal Oxide Semiconductor
CPRI	Common Public Radio Interface
DAC	Digital-to-Analog Converter
DAC	Digital-to-Analog Converter
DFT	Discrete Fourier Transform
DPD	Digital Predistortion
DSP	Digital Signal Processing
EMP	Envelope Memory Polynomial
EVM	Error Vector Magnitude
FBMC	Filter Bank Multi-Carrier

FDSOI	Fully-Depleted Silicon-on-Insulator
FM	Frequency Modulation
FSK	Frequency Shift Keying
GaAs	Gallium Arsenide
GaN	Gallium Nitride
GMP	Generalized Memory Polynomial
HBT	Heterojunction Bipolar Transistor
HD	Harmonic Distortion
HPA	High Power Amplifier
HWN	Heterogeneous Wireless Network
ICT	Information and Communications Technology
IF	Intermediate Frequency
IM	Intermodulation
IM3	Third-order Intermodulation
IMD	Inter-Modulation Distortion
IoT	Internet of Things
IP3	Third-order Intercept point
IQ	Inphase and Quadrature
ISI	Inter-Symbol Interference
LNA	Low noise Amplifier
LTE	Long Term Evolution
LTI	Linear Time Invariant
LUT	Look-Up Table
MIMO	Multiple-Input Multiple-Output
ML-LUT	Memory Less Look-Up Table
MP	Memory Polynomial
MSPD	Mixed-Signal Predistortion
NMSE	Normalized Mean Square Error
OBSAI	Open Base Station Architecture Initiative
OFDM	Orthogonal Frequency-Division Multiplexing
PA	Power Amplifier
PAE	Power Added Efficiency
PAPR	Peak-to-Average Power Ratio
PD	Predistortion
PM	Phase Modulation
PSD	Power Spectral Density

PSK	Phase Shift Keying
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
RBS	Radio Base Station
RF	Radio Frequency
RFPD	Radio Frequency Predistortion
RRH	Remote Radio Head
SSAPI	Small-Signal Assisted Parameter Identification
TNTB	Twin Nonlinear Two-Box
TWT	Traveling Wave Tube
UE	User Equipment
UFMC	Universal Filtered Multi-Carrier
UHD	Ultra High Definition
VGA	Variable Gain Amplifier
WCDMA	Wideband Code Division Multiple Access
NF	Noise Figure
FFT	Fast Fourrier Transform
FPGA	Field Programmable Gate Array
AGC	Automatic Gain Control
AWG	Arbitrary Wave Generator
PSD	Power Spectral Density
LMS	Least Mean Square
SFDR	Spurious Free Dynamic Range
BW	BandWidth
FS	Full Scale
SNR	Signal-to-Noise Ratio
SNDR	Signal-to-Noise and Distortion Ratio
FIR	Finite-Impulse-Response
IIR	Infinite-Impulse-Response
LPF	Low Pass Filter
BPF	Band Pass Filter

Chapter 1

Introduction

The phenomenal increase in the number of mobile devices, coupled with an exponential growth of the data traffic to support emerging applications, such as cloud storage and computing, over the past few years has led to an extensive increase in energy consumption by cellular networks. This scenario is expected to sustain or even exacerbate in foreseeable future. The energy consumption of the Information and Communications Technology (ICT) is expected to grow from $600\,\text{TWh}$ in the year 2009 to $1700\,\text{TWh}$ by the year 2030 [14]. This is around 3-4% of the total world electrical energy consumption [15, 14]. A significant portion (around a third) of it is consumed by the mobile communication networks.

Base stations are at the heart of these mobile communication networks, which account for more than 50% of the total network energy consumption [16, 17, 18], and can even be up to 85% [19]. The component which is still consuming the majority (60%) of the power in a base station is the power amplifier (PA) [20, 21, 22], which is a key block of the RF transceiver that delivers high power to the antenna. The increase in data rates calls for increased signal bandwidths, in excess of $100\,\text{MHz}$. While new spectrum bands are being added to the standards, the spectrum is still a scarce resource. This has led to cell densification, whereby efficient spectral reuse is possible. Low-power small-cell base stations, namely, picocells and femtocells, have been emerging as a natural choice to increase the network capacity, with a low cost of installation and operation when compared to the microcells and macrocells.

In this chapter, an introduction to the basic background concepts in wireless communication systems and also the concepts related to the Power Amplifiers (PAs), are provided in Section 1.1 and Section 1.2, respectively. Conclusions for these two sections are provided in Section 1.3. The theory developed will be utilized in the subsequent section and

chapters. Finally, Section 1.4 provides an overview of the issues dealt in this work along with the scientific achievements.

1.1 Background on Wireless Systems

The radio standards have evolved from pre-cellular mobile radio telephone (or 0G), to cellular fourth generation (4G) Long Term Evolution (LTE), with each cellular generation lasting approximately for a decade. 4G networks have hit the theoretical data rate limits for the contemporary technologies and cannot address the growing demands for data rates in a sustainable manner. Hence they have evolved towards 5G.

1.1.1 5th Generation Mobile Networks

5G mobile radio networks are slated to be deployed beyond 2020 [23]. Though the standards are not yet released, the main features of 5G would be to provide ubiquitous and seamless communications all the time, not just between humans but also between machine to machine and human to machine. The 5G networks should also provide baseline data rates for each user of around 1 Gbps and peak rates of up to 10 Gbps. This requires larger signal bandwidth available for each user, over 100 MHz [24]. This is enabled by advanced Carrier Aggregation (CA). For example, five Component Carriers (CCs) of 20 MHz can be aggregated to provide 100 MHz of spectrum to a single user. While the available radio spectrum is getting crowded and is increasingly looking scarce, the target for 5G is to achieve 1000 times more system capacity. Spectral efficient modulation schemes will be introduced, with a target to improve the spectral efficiency by 10 times. The problem of spectrum availability could be solved by utilizing new bands in sub 6 GHz and exploring centimeter/millimeter spectrum (6 GHz to 100 GHz band). Spectral reuse and Multiple-Input Multiple-Output (MIMO) techniques for spatial multiplexing within smaller cells (picocells and femtocells) will result in denser deployment. Apart from providing high data rates, the networks should also be energy efficient, reliable, provide low latency services and should support multitude of low power devices of IOT (Internet Of Things) and hence the networks should be highly energy-scalable. Backward compatibility and co-existence with legacy radio access technologies are nevertheless needed.

1.1.2   Cellular Base Station Architecture

Base Transceiver Station (BTS) or simply Base Station (BS), also known as Radio Base Station (RBS), node B in 3G networks and evolved node B (eNB) in 4G networks are at the heart of cellular communication networks. They are usually stationary installations of Base-Band Unit (BBU) and radio equipment, known as Remote Radio Head (RRH), containing transceivers with necessary electronic circuitry, Power Amplifiers (PAs) and antennas to facilitate communication between User Equipments (UEs) [25, 26]. The fronthaul connects the BBU and RRH using Common Public Radio Interface (CPRI) or Open Base Station Architecture Initiative (OBSAI) optical links, of which CPRI is the most common one. The CPRI links are usually clocked at submultiples of 30.72 Mbps, since a basic CPRI frame rate is 3.84 MHz [25, 26]. The network backhaul provides the necessary data and control information from the mobile switching center or the core network for transmission and reception to the BS. The network backhaul might be either an optical or a microwave link providing sufficient data capacity. Free-space optical communications is also emerging as an option for the 5G technologies [27]. The BSs are also the most power hungry subsystem of the cellular network, amounting to more than 50% of the total power consumption [28].

Based on the minimum coupling loss between the BS and the UE, four classes of base stations are defined by 3rd Generation Partnership Project (3GPP) in the present communication standards, as mentioned in Release 15 [29].

Table 1.1 presents the four classes of BSs and the scenarios from which the classes are derived from. The $P_{rated,c}$ is the rated output power of the BS, which is defined as the mean power per carrier at the antenna connector port during the transmission. The BSs can operate in single carrier, multi-carrier, or carrier aggregation configurations.

TABLE 1.1: BS classification and their properties

BS class	BS scenario	Min. coupling loss (dB)	$P_{rated,C}$ (dBm)[1]	Approx. max. coverage radius
Wide area	Macrocell	70	No upper limit	≤ 35 km
Medium range	Microcell	53	≤ 38	≤ 2 km
Local area	Picocell	45	≤ 24	≤ 200 m
Home	Femtocell	—	≤ 20[2]	10m

[1] Nominal condition tolerance is ±2 dBm and extreme condition tolerance is ±2.5 dBm
[2] For one transmit antenna port. The rated power per antenna connector is accordingly scaled with the number of transmit antennas, for example, P_{rated} is ≤ 17 dBm for double transmit antenna ports and < 11 dBm for eight transmit antennas, which is the maximum number of transmit antennas.
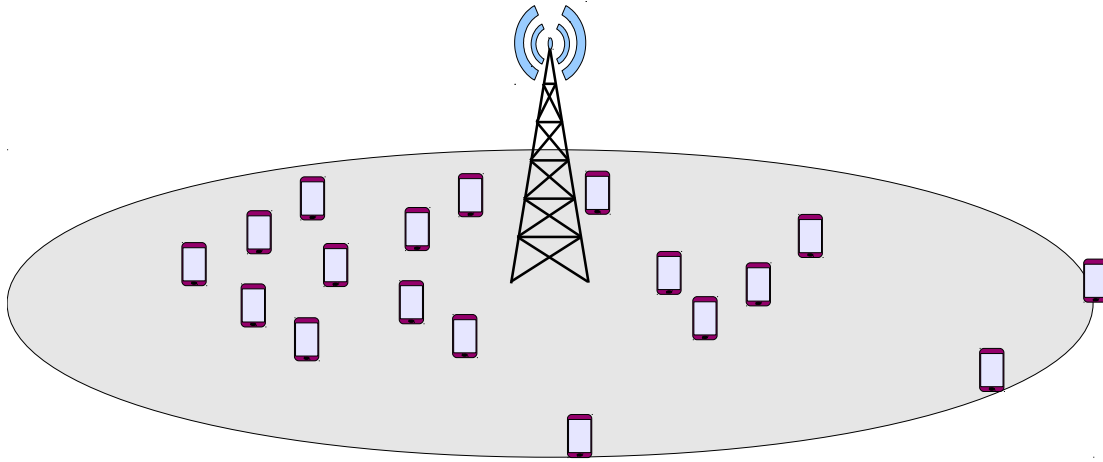
FIGURE 1.1: Illustration of cellular base station - a conventional macrocell network

Each of the BS class mentioned in Table 1.1 has its own purpose and properties. Wide area BS (macrocell) has been present since the inception of cellular communications and are commonly mounted on towers or rooftops of tall buildings for maximizing network coverage area, as shown in Fig. 1.1. The disadvantages of wide area BSs are high cost and power of operation, with the necessity for air-conditioning the High Power Amplifiers (HPAs), which by definition means any amplifier whose output power is greater than one Watt. Also, the users present at the edge of the cell have very weak signal strength. Later with the 2G the need for higher network capacity has called of microcells, which could serve densely populated localities needing extra network capacity, but with reduced radius of coverage and power. The aim is to utilize the spectrum efficiently by frequency reuse and reducing interference with proper frequency coordination.
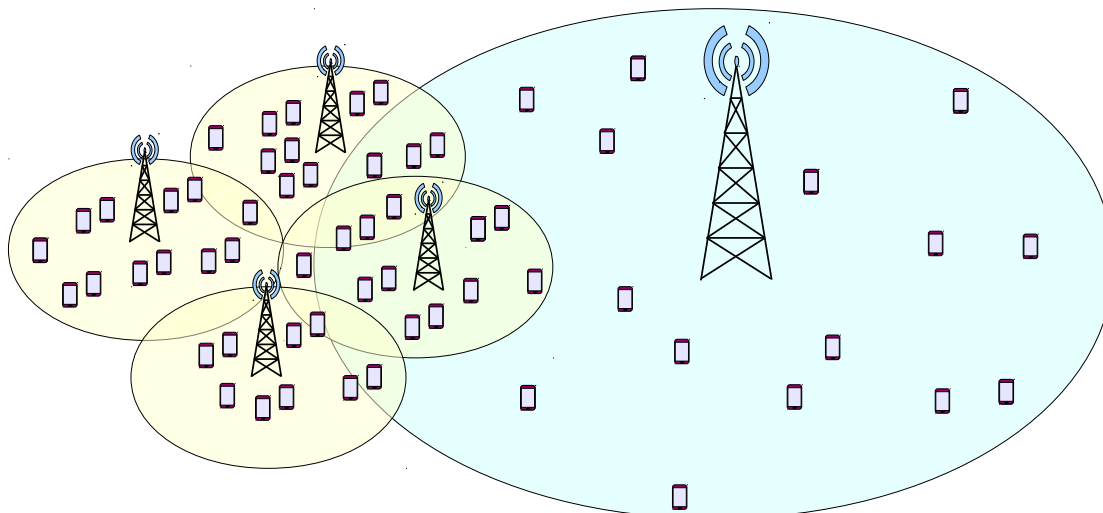
FIGURE 1.2: Illustration of cellular base station - Heterogeneous Wireless Network (HWN)

Small cells, namely, picocells and femtocells, have been emerging as a natural choice to increase the network capacity, with low cost of installation and operation and without costly air-conditioning equipment starting with the 3G technology. The picocell form factor could be used to give network coverage to usually large indoor areas, where the signal strengths from macrocells and microcells are poor, like a big shopping mall, railway station, or a stadium, etc,. The femtocell goes a step further with even smaller power and coverage area, but is designed to efficiently cover small office or a home. The small-cells are inexpensive and easy to deploy when compared to microcells and macrocells, which are usually mounted on towers. The denser deployment of small-cells also improves the reliability, for example, in the case of failure of a femtocell, other nearby picocell can possibly serve the users momentarily, which could not be the case for a single macrocell scenario. The future 5G network architecture will be a combination of all the four classes of BSs called HWNs, or simply, Heterogeneous networks (HetNets) [30], as illustrated in Fig. 1.2.

1.1.3 Radio Frequency Transceiver
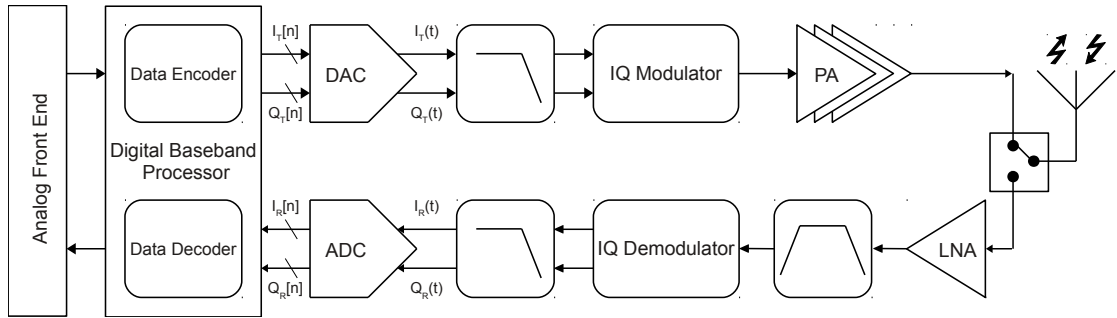
FIGURE 1.3: Simplified block diagram of an Radio Frequency (RF) transceiver

A simplified RF transceiver in the contemporary digital communications context is shown in Fig. 1.3. In the transmit path, the Analog Front End (AFE) senses or acquires electrical equivalent signals, filters in analog domain and converts it to digital signals with the help of an Analog-to-Digital Converter (ADC). The sensors could be anything ranging from temperature sensor, to Ultra High Definition (UHD) video camera in the case of an User Equipment, or a photo diode for an optical backhaul of a BS. The digitized data is then processed in the digital baseband processor, doing necessary signal processing such as filtering, Discrete Fourier Transform (DFT), source encoding, channel encoding, etc. The source encoding does the compression of the data, and the channel coding introduces controlled redundancy which reduces the probability of error when the data is transmitted through the channel to be received by the receiver. Based on the digital modulation scheme the Inphase and Quadrature (IQ) data, $I_T[n]$ and $Q_T[n]$, respectively, are sent to the respective Digital-to-Analog Converters (DACs) for the case

of a zero IF transmitter, as considered here. Zero-IF or direct conversion architecture is one of the most used transceiver architectures [31]. The output signals of the DACs, $I_T(t)$ and $Q_T(t)$, are then low-pass filtered using reconstruction filters also called anti-imaging filters, to remove unwanted out-of-band noise and harmonics. An IQ modulator upconverts the complex baseband analog signal into real RF signal, which is usually followed by a Band Pass Filter (BPF). The RF signal to be transmitted is amplified by a PA, which is discussed in greater detail in Sec. 1.2. The PA connects to the duplexer, which facilitates full-duplex operation avoiding the leakage of the transmitted signal into the receiver, and hence helps in avoiding two separate antennas, one for transmission and the other for reception. Finally, the antenna radiates the amplified RF signal into the free-space.

At the receiver end, the same antenna receives the desired reception signal with other unwanted signals at the same time. The received signal is processed in the reverse way, that is the signal is amplified, band selectively filtered, quadrature-downconverted to analog baseband and finally converted into digital domain, $I_R[n]$ and $Q_R[n]$, by the Low noise Amplifier (LNA), BPF, ADC preceding with an Anti-Aliasing Filter (AAF), respectively. The digitized received data is processed and decoded by the same baseband processor. The received digital data can be stored or can be sent to the actuators, for example on to a screen for displaying information, with the help of AFE.

1.1.4   Digital Modulation

Modulation is the modification of a carrier wave in accordance with the baseband data so that it is easy to transmit and receive properly. This can be accomplished by selectively modifying the sinusoidal carrier's parameters, namely amplitude, frequency, and phase. All the modern communication systems employ digital modulation techniques, which are robust in comparison with their analog counterparts.

The basic digital modulation schemes are Amplitude Shift Keying (ASK), Frequency Shift Keying (FSK) and Phase Shift Keying (PSK), which are analogous to Amplitude Modulation (AM), Frequency Modulation (FM) and Phase Modulation (PM) in analog modulation, respectively. Advanced digital modulation schemes use quadrature modulation, for example QPSK, whose constellation is as shown in Fig. 1.4.

Gray coding is often employed to minimize the number of bits differing between two adjacent symbols, thereby minimizing the error probability. Spectral efficiency is the most important metric of any digital modulation scheme, which is described as the number of bits per second that can be transmitted over a bandwidth of one Hertz. Spectral efficiency of various modulation schemes are presented in Table 1.2, [32, 33].
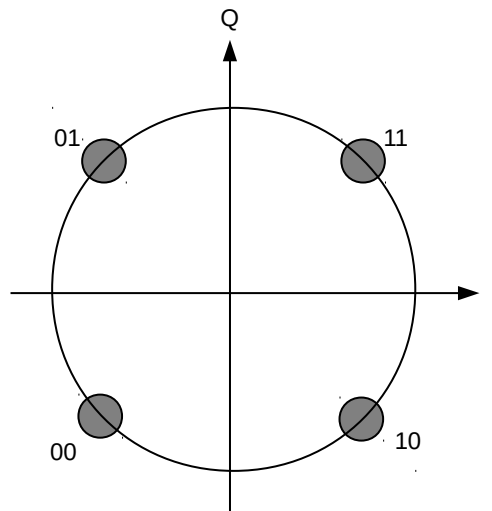
FIGURE 1.4: Constellation diagram for Quadrature Phase Shift Keying (QPSK)

TABLE 1.2: Spectral efficiency of various modulation schemes

Modulation scheme	Spectral Efficiency (bits/s/Hz)	PAPR (dB)
BPSK	1	0
QPSK	2	0
8PSK	3	3.3
64QAM	6	3.7
OFDM	≥ 10	~ 12

[1] Can reach as high as 30 bits/s/Hz, depending on the number of subcarriers and the modulation schemes for them

As can be seen in the Table 1.2, the OFDM is one of the most spectral efficient modulation schemes, which is used in 4G LTE. OFDM achieves this by employing multiple orthogonal subcarriers in a channel, each having its own modulation scheme (QPSK, 64QAM, etc.) and packing high amount of data in a given bandwidth. The orthogonality of the subcarriers avoids Inter-Symbol Interference (ISI). Orthogonal Frequency-Division Multiplexing (OFDM) is a synergistic combination of modulation and multiplexing technique. The other advantage of OFDM is immunity to multipath fading. For the case of 5G communications, other modulation formats such as Filter Bank Multi-Carrier (FBMC) and Universal Filtered Multi-Carrier (UFMC) are being looked at with advantages in comparison with OFDM [34, 35, 36].

1.2   Background on Power Amplifier

The power amplifier is the most important stage in the RF transmitter, and is also the most power-hungry circuit not just in the transmitter, but also in the entire transceiver chain and in the BS [22, 37]. It is the stage which prepares the low power radio signal coming form the IQ modulator by giving it as much power as possible from the DC supply, making it a high power signal to get transmitted into the free-space via antenna. The power gain of the PA is the ratio of the output power P_{Out} and the input power P_{In}, given as:

$$\text{Gain} = \frac{P_{Out}}{P_{In}}. \tag{1.1}$$

There are various classes of PAs with varying combinations of linearity and power efficiency [31, 38, 39]. Though from the definition perspective of an ideal PA, only linear gain is assumed, in reality, depending on the PA class of operation, various nonlinear effects comes into picture.

Coming to the choice of the RF power amplifier technology in the BSs, Gallium Nitride (GaN) and Gallium Arsenide (GaAs) were dominating the market till few years ago [39]. Now the silicon LDMOS technology is the leading choice, even for a small-cell BS requiring upto few Watts, for its better performance and lower cost. It is a variant of the CMOS technology but is capable of delivering far more output power when compared to that of a normal CMOS device.

1.2.1   Generic PA Metrics

Apart from the power gain given by Eq. 1.1, the following two metrics, namely, efficiency and Power Added Efficiency (PAE), are very important.

1.2.1.1   Efficiency

The foremost important performance measure of a PA is its efficiency, which is given by

$$\eta_{PA} = \frac{P_{Load}}{P_{Supply}}, \tag{1.2}$$

where P_{Load} is the power delivered to the load and P_{Supply} is the power that the PA draws from the supply. It is also known as the drain efficiency or collector efficiency (for

the respective CMOS or bipolar PA implementations [31]. Ideally, the efficiency should be unity, where the entire power supplied to the PA should be delivered to the load. But in reality, depending on the class of PA operation and other practical reasons of physical implementations, like the limited output swing originating from the operating region of the power device, the efficiency is always below unity. Also note that this metric doesn't take into account the gain of the amplifier.

1.2.1.2 Power Added Efficiency (PAE)

The other way to define the performance of the PA is by defining the PAE given by:

$$\text{PAE} = \frac{P_{Load} - P_{In}}{P_{Supply}}, \tag{1.3}$$

where P_{In} is the power at the PA input signal port.

PAE is important for a PA because the amplifier is mostly driven by another smaller amplifier in a tapering fashion for improved drivability and matching considerations. Hence we might note the following:

- The PAE is always smaller than the efficiency η_{PA}

- The better the PA input matching with the preceding stage, the higher the PAE.

- The higher the gain of the amplifier, the higher the PAE.

1.2.2 PA Behavior

PA input-output behavior can be categorized into three groups:

1. **Memoryless (static) nonlinearities** which is inherent to the power device. Taylor series expansion can be used to approximate the output behavior over some signal range:

$$y(t) = \sum_{n=1}^{N} a_n x(t)^n, \tag{1.4}$$

where $x(t)$ is the input signal, $y(t)$ is the output signal, a_n are the coefficients of the polynomial, N is the nonlinearity order considered.

2. **Linear memory effects** which are memory behaviors uncorrelated with the nonlinear response of the power amplifier arising from time delays or phase shift

in the matching networks and can be modeled as Finite-Impulse-Response (FIR) filters.

3. **Nonlinear memory effects** come from linear circuits, such as capacitors, for example, combining with the nonlinear behavior of the transistor results in a term in the output signal of the power amplifier that includes a nonlinear function of different samples of the input signal at different instances [37].The other sources of nonlinear memory effects are direct low-frequency dynamics, such as trapping effects and non-ideal bias networks [22].

1.2.3   PA Nonlinearity Characterization Metrics

As previously mentioned the amplifier gain is nonlinear which results in the in-band signal degradation as well as unwanted emissions in other channels. The following section describes briefly the metrics used to measure the PA nonlinearity.

1.2.3.1   Adjacent Channel Leakage Ratio
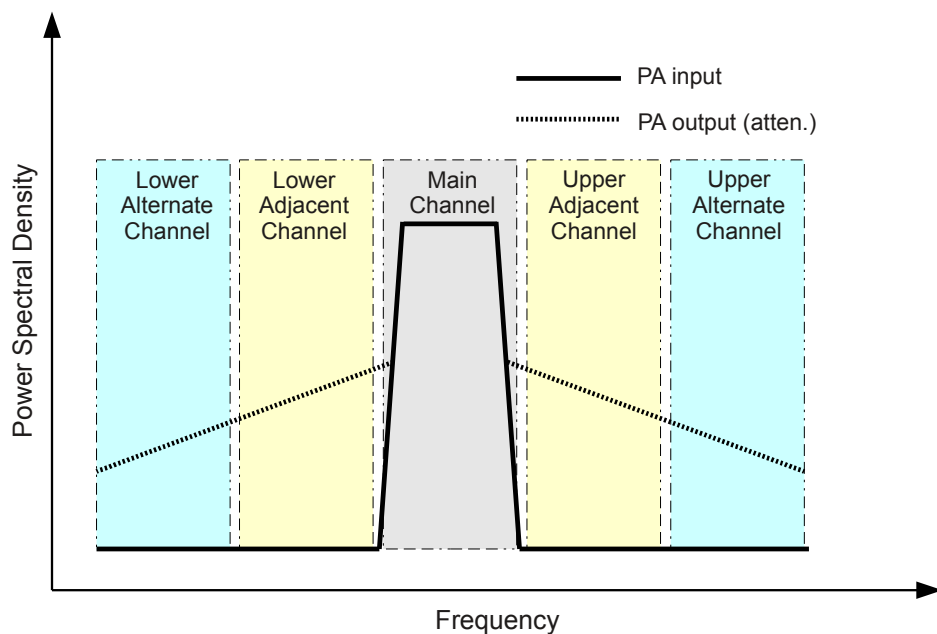
FIGURE 1.5: Illustration of PA input and output spectra

The PA nonlinearity can be characterized using Adjacent Channel Leakage Ratio (ACLR), also called as Adjacent Channel Power Ratio (ACPR), which gives us the measure of the extent to which the nonlinearly amplified modulated signal spreads to the adjacent and alternate channels in the frequency domain. Fig. 1.5 shows a simple illustration of a modulated input and attenuated output spectrum of a PA, attenuated with a factor

of PA linear gain. The main channel is surrounded by similar bandwidth lower and upper adjacent and alternate channels respectively with guard bands in between. The main channel is the desired channel and is considered as the reference channel when calculating the ACLR, which is the power ratio expressed in dBc, and is given by

$$\text{ACLR (dBc)} = 10 \log_{10} \frac{\int_{BW} P_{Main}(f) df}{\int_{BW} P_{Adjacent}(f) df}, \tag{1.5}$$

where $P_{Main}(f)$ and $P_{Adjacent}(f)$ are the Power Spectral Densities (PSD) in the main channel and the adjacent channel respectively. It could be measured with respect to the alternate channel as well. The usual notation is ACLR1_U and ACLR1_L, when referring to upper and lower adjacent channels and ACLR2_U and ACLR2_L, when referring to upper and lower alternate channels.

Characterizing the spectral regrowth with the calculation of ACLR is one of the most important requirement, as each radio communication standard defines limits on spectral emissions with the help of the spectral mask, which should be abided by anyone who wants to communicate wirelessly in a specified licensed spectrum.

1.2.3.2 Error Vector Magnitude (EVM)
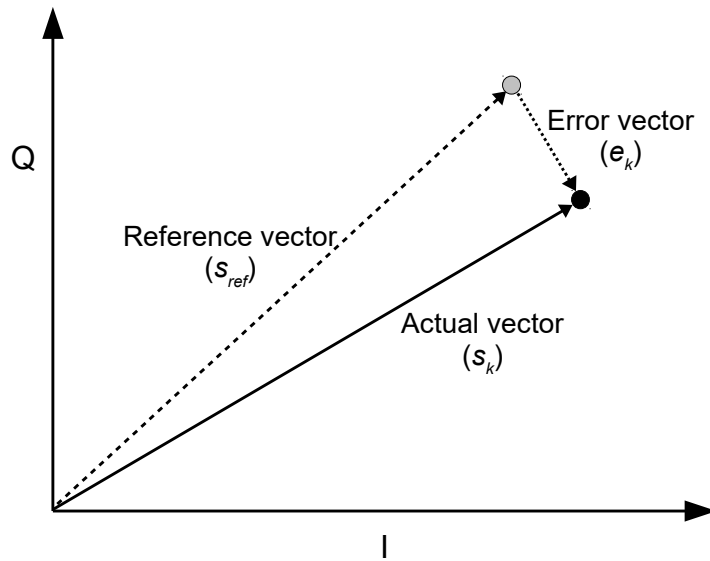
FIGURE 1.6: Illustration of EVM

The Error Vector Magnitude (EVM) is another nonlinearity measure which is used to quantify the PA nonlinearity. Contrary to ACLR, EVM describes the extent to which the nonlinearity of the amplifier degrades the inband quality of the modulated signal. The EVM is defined and calculated in the constellation domain, which because of nonlinearity

gets distorted and hence dispersed from its original position, as illstrated in Fig. 1.6. EVM is expressed in percentage and by definition is measured over one subframe in the time domain, which is 1 ms according to 3GPP [29]. The formula for EVM calculation is

$$\text{EVM } (\%) = \sqrt{\frac{\sum_{k=1}^{N} |e_k|^2}{\sum_{k=1}^{N} |s_{ref}|^2}}, \tag{1.6}$$

where

$$e_k = s_k - s_{ref}, \tag{1.7}$$

where s_{ref} is the reference vector, s_k is one of the N vectors present in one subframe, which is obtained after nonlinear PA amplification and e_k is the error vector.
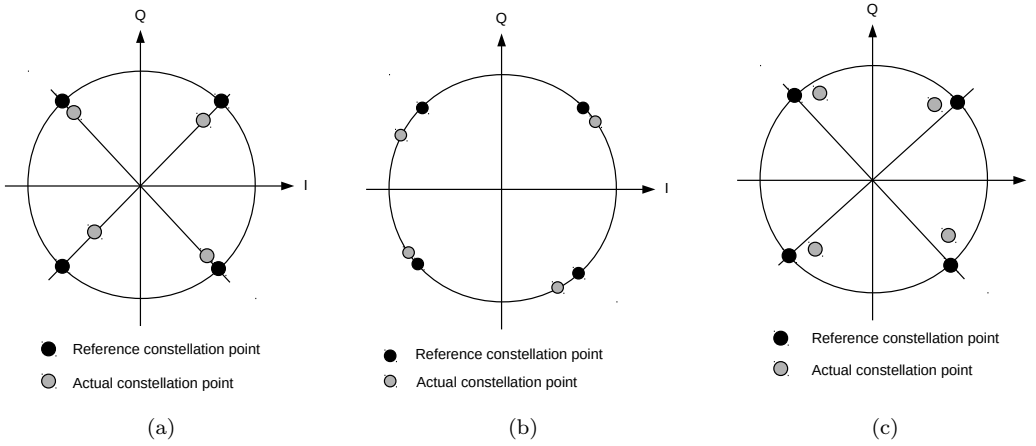
FIGURE 1.7: Effects of distortion on QPSK constellation: (a) amplitude distortions,(b) phase distortions, and (c) combination of phase and amplitude distortions

Fig. 1.7 illustrates the effects on PA nonlinearity on the signal constellation. Nonlinear PA can give amplitude distortion Fig. 1.7(b), phase distortion Fig. 1.7(a), or a combination of both amplitude and phase distortions Fig. 1.7(c).

1.2.4   Effect of Peak-to-Average Power Ratio (PAPR) and PA nonlinearity on the efficiency

Linear amplification is desired for the modulation schemes based on amplitude modulation, where the modulated signal envelope also carries information. New spectral-efficient modulation schemes, like OFDM for example. To achieve linear gain from the PA, the PA should be "backed-off", which means the whole input signal gets amplified only in the linear region of the PA without pushing it into nonlinear or saturation region. This significantly degrades the power efficiency of the PA. This is further exacerbated by the high PAPR, for example, in the case of OFDM, PAPR is very high around 12 dB
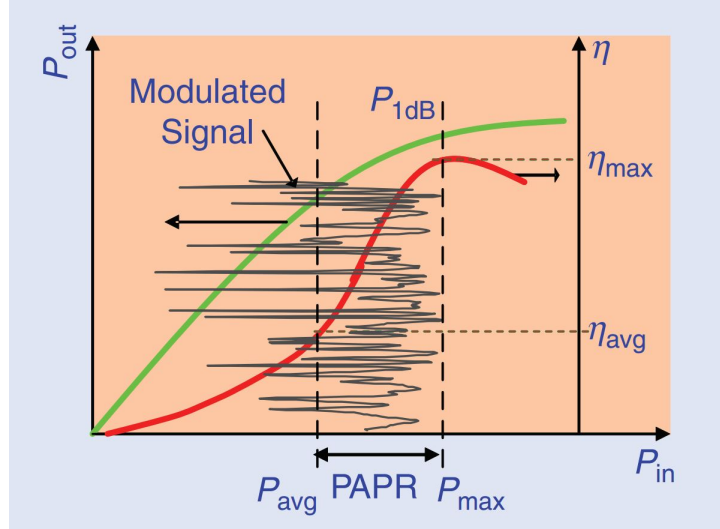
FIGURE 1.8: Illustration of effect of PAPR; output power and efficeincy vs. input power [1]

as shown in Table 1.2. As shown Fig. 1.8, there exists a trade-off between the linearity and power-efficiency of the PA, which also has a dependency on signal PAPR. This makes the efficiency go below 10%, with the rest of the 90% power being dissipated in the power device. This calls for an advanced thermal management like expensive packaging, large heat-sink and air-conditioning. Therefore, it is to be noted that PAPR plays a very crucial role in the efficiency of the power amplifiers. Techniques such as Crest Factor Reduction (CFR) is usually employed to reduce the PAPR, at the cost of EVM degradation [25].

1.3 Conclusion

This chapter has presented the scenario of 5G mobile networks, where increasing number of low power small-cell BSs are going to play a vital role in the realization of cost and energy efficient ubiquitous communications. Radio frequency transceivers in the context of modern day communications utilize robust digital modulation techniques, like Quadrature Amplitude Modulation (QAM) and OFDM, and the preferred transceiver architecture is the zero-IF architecture. PA being the important stage of the transmitters is also the most power hungry block in the entire BS, whose efficiency comes at the cost of linearity. The important generic and nonlinearity metrics of the PA were introduced and the effect of PAPR and nonlinearity on the efficiency of the PAs was discussed.

1.4   Specific Issues Dealt in This Work and Achievements

1.4.1   Problem Statement and Thesis Objective

The PA suffers from a strong linearity/efficiency trade-off. The nonlinearities result in intermodulation distortions at the PA output which when transmitted cause spectral pollution, i.e., leaking a portion of transmitted power into the adjacent and alternate channels. Also, with the increased signal bandwidths the problem of memory effects has increased considerably, which results in dynamic nonlinearities. Additionally, new types of modulation schemes, such as OFDM, generate modulated signals with non-constant envelope resulting in high signal PAPR, further degrading the PA characteristic. In order to break this trade-off and increase the efficiency of the PA without linearity degradation, predistortion is usually employed. Predistortion corrects the PA nonlinearity and memory effects by generating approximate PA inverse characteristic to generate a fairly linear output at the PA. Predistortion can be done in analog [40, 41], or digital [42, 43], or even in the analog RF domain [8, 10]. Owing to the robustness of digital signal processing, and benefits coming from the Complementary Metal Oxide Semiconductor (CMOS) technology scaling, Digital Predistortion (DPD) has become the *de facto* solution for the PA linearization [22, 44].

Even though a plethora of DPD solutions have been proposed in the literature based on behavioral modeling, ranging from simple memoryless look-up table methods to complex neural networks, as summarized in [22, 45], they are specifically targeted towards macro- and micro-cell BS PAs, where the DPD power consumption is negligible when compared to that of the PA. Hence all the research effort has been made to obtain highly performant and robust DPD. Also, with the increasing bandwidths, employing a DPD, which usually has to handle at least five times the bandwidth of the signal in order to cancel out the distortion products, becomes excessively power-hungry. In the context of small-cell base stations the use of conventional DPD solutions becomes prohibitively power-hungry, and hence no DPD solutions are generally used until very recently [21, 22, 46]. Without a DPD, the PA suffers from poor power efficiency as the PA is usually backed-off to operate in linear regime. New modulation schemes tend to show high PAPR, and hence worsening the aforementioned problems.

The objective of the thesis is to develop low-power predistorter solutions suitable to linearize the small-cell base station PAs in the context of high bandwidth input signals. In particular, we are interested in developing a simplified predistorter model that can be employed not just in DPD implementations but also in analog and mixed-signal based implementations, which are emerging as alternatives in the context of small-cell PA predistortion.

1.4.2 Thesis Contributions and Organization

The thesis is organized in various chapters. A brief outline of it is as follows:

- Chapter 2 presents a brief literature review up to various state-of-the-art DPD and ARFPD solutions, which are again grouped into memory unaware and memory aware techniques. The chapter culminates with a discussion on various advantages and disadvantages of both the predistortion techniques and provides a comparison of them.

- Chapter 3 presents the developement of the FIR-MP model starting with the predistorter modeling using conventional memory polynomial, detailing its shortcomings corroborated by MATLAB simulations. The digital implementation flow of the proposed FIR-MP algorithm in 28 nm Fully-Depleted Silicon-on-Insulator (FDSOI) CMOS technology and the simulation results obtained are presented.

- The architecture of the FIR-MP mixed-signal predistorter is presented in Chapter 4. A brief analysis of the various non-idealities to derive the requirements of the circuit to be implemented using the proposed architecture is provided along with simulations.

- Finally, Chapter 5 provides concluding remarks and directions for the future work.

1.4.3 Scientific Publications

The thesis has resulted in the following scientific publications:

1. **V. N. Manyam**, D.-K. G. Pham, C. Jabbour, and P. Desgreys, "A low-power high-performance digital predistorter for wideband power amplifiers," *Analog Integrated Circuits and Signal Processing*, pp. 1–10, Jun. 2018.

2. **V. N. Manyam**, D.-K. G. Pham, C. Jabbour, and P. Desgreys, "A Wideband Mixed-Signal Predistorter for Small-Cell Base Station Power Amplifiers," in *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, Florence, 2018, pp. 1–5.

3. P. Desgreys, **V. N. Manyam**, K. Tchambake, D.-K. G. Pham, and C. Jabbour, "Wideband power amplifier predistortion: trends, challenges and solutions," in *2017 IEEE 12th International Conference on ASIC (ASICON)*, Guiyang, 2017, pp. 100–103.

4. **V. N. Manyam**, D. K. G. Pham, C. Jabbour, and P. Desgreys, "An FIR memory polynomial predistorter for wideband RF power amplifiers," in *2017 15th IEEE International New Circuits and Systems Conference (NEWCAS)*, Strasbourg, 2017, pp. 249–252.

5. **V. N. Manyam**, D.-K. G. Pham, C. Jabbour, and P. Desgreys, "Filter Assisted Memory Polynomial Predistortion for Small-Cell Base Stations," presented at the *2017 12th National GDR SoC/SiP conference*, Bordeaux, 2017.

Chapter 2

State-of-the-Art Predistortion Techniques

2.1 Introduction

There exists a strong trade-off between linearity and power efficiency of the Power Amplifier (PA), as discussed in Chapter 1. Predistortion (PD) is the most preferred method with which this trade-off can be elegantly broken. There is an immense demand for low power predistortion system, which can be utilized for linearizing a PA in the context of small-cell Base Stations (BSs) and User Equipments (UEs). The aim of this thesis is to address the small-cell BS scenario PA PD implementation. The purpose of this chapter is to identify potential PD principles and architectures present in the existing literature.

Digital Predistortion (DPD) employed in digital baseband has dominated the predistortion scenario of PAs, because of the recent improvements in DSP and cost reduction and increased functionality coming from nanometer CMOS technologies. Most of the recent research is currently carried out on digital domain. This is clearly evident from the number of recent publications. Many researchers have done comparative analysis on behavioral modeling and predistortion techniques [45, 22, 47], they are predominantly composed of digital baseband modeling and DPD and there are only few analog and RF PD publications available in the existing literature.

Starting with the principle of predistortion and a brief outline on various predistorter classifications in Section 2.2, this chapter elaborates various types of predistortion techniques present in the literature. Memory-unaware and memory-aware DPD techniques, along with their advantages and disadvantages are presented in Section 2.3. In a similar

33

way ARFPD techniques are outlined with their advantages and disadvantages in Section 2.4. We then present the comparison of the two categories of predistortion methods in Section 2.5 and finally present the conclusions in Section 2.6.

2.2   Outline of the PA Predistortion

The basic principle of predistortion can be understood with the help of Fig. 2.1. The PA exhibits a compressive input-output transfer characteristic as shown in y vs. v curve. The goal of the predistortion system is to generate an expansive characteristic output, mimicking the PA inverse behavior as shown in v vs. x plot so that the overall output of the PD and PA becomes linear for a reasonable input range as depicted in y vs. x plot. PD system has a feedback path also known as observation path and an implementation
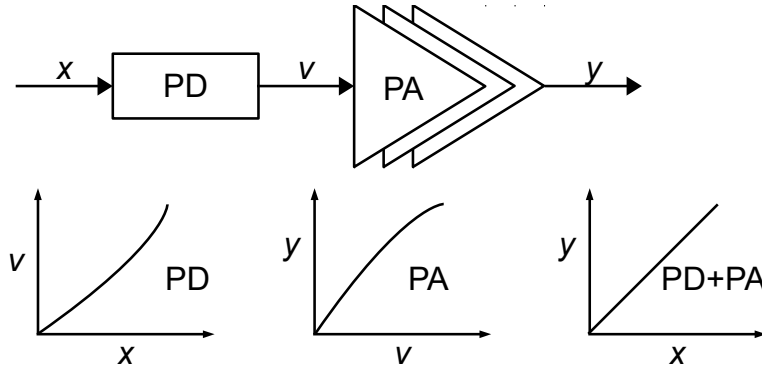
FIGURE 2.1: Illustration of the principle of PA predistortion

path.

Based on the domain in which the predistortion is performed, the predistortion methods can be broadly classified into two categories: digital and analog/RF, known as DPD and Analog Radio Frequency Predistortion (ARFPD), respectively. There are some hybrid predistorters which combines digital and analog RF techniques, but can be categorized as ARFPDs. Each of the categories can be further sub-classified into memory-aware and memory-unaware PD methods, depending on the PA memory-effects correction capability. An ideal memoryless nonlinear system can be described by its AM/AM (amplitude modulation to amplitude modulation) characteristic. But usually a memoryless PA exhibits not only AM/AM but also AM/PM (amplitude modulation to phase modulation) characteristic, hence quasi-nonlinear system. Hence by memoryless nonlinearity correction we mean to correct not only the nonlinear AM/AM characteristic but also the AM/PM characteristic of the PA. Also, the predistortion can be adaptive or non-adaptive. In the adaptive PD [48, 49, 50], the feedback path should always be present. Whereas in the non-adaptive predistorters, feedback path is used only

during the initial learning phase or when an update is needed. Adaptive predistortion can mitigate the changes in PA characteristics, originating due to aging, temperature variations and supply voltage variations, and other reliability dependent effects. The training or learning of the PD can be categorized into direct or indirect learning. Here we mainly focus on the implementation path and non-adaptive predistorters. Also, multi-band (dual-band [51, 52, 53, 54, 55], triple-band [56]), MIMO systems [57, 58] and low-rate ADC feedback (band-limited or undersampled) DPDs, which are usually adaptive systems [59, 60] are not addressed explicitly in the thesis.

BSs in the conventional macrocell/microcell scenarios use High Power Amplifiers (HPAs) which exhibit strong static nonlinearities and memory effects [61]. For any HPA the wider the input signal BandWidth (BW), the stronger the memory effects are. Further to that the constraints on Adjacent Channel Leakage Ratio (ACLR) are also very stringent on BS. On the other hand, handsets or UE PAs are generally of power around a watt [62], which coincides with the small call BS PA transmit powers, especially for picocells and femtocells as previously shown in Table. 1.1. The good thing with the UE PAs is that they have very less memory effects [61]. According to 3rd Generation Partnership Project (3GPP) the ACLR specification for a BS PA should be always greater than (except for Band 46) 45 dBc in adjacent and alternate channel, known as ACLR1 and ACLR2, respectively [29, 22]. For the UE PA it should be always greater than 33 dBc, 43 dBc at 5 MHz and 10 MHz offset, respectively, for Wideband Code Division Multiple Access (WCDMA) signals [7, 63, 62]. General observation regarding BS PA is that the ACLR specification for an uncorrected BS PA, i.e. without any PD is around 30 dBc, this is in order to achieve high efficiency at the cost of nonlinearity [64]. After correction using predistortion it becomes greater than 50 dBc, i.e., there is an improvement of at least 20 dB, with minimum 5 dB margin [64]. Margins are necessary especially if the predistorter is not adaptive. For the case of UE PAs the ACLR improvement is usually less than 10 dB, but the power constraint on PD is very stringent [65, 7].

2.3   Digital Predistortion Methods

In DPD, the digital baseband modulated signal is subjected to the inverse nonlinear transfer characteristic of the power amplifier, in the digital baseband itself. Fig. 2.2 shows a transmitter chain employing a DPD solution. To correct the inter-modulation distortion components of the PA, the spectrum regrowth occurs at the digital baseband itself, which is usually at least five times the input signal bandwidth.

In the DPD context, a proper behavioral model must be capable of characterizing the nonlinear distortion and memory effects. Since PD implements the inverse function of
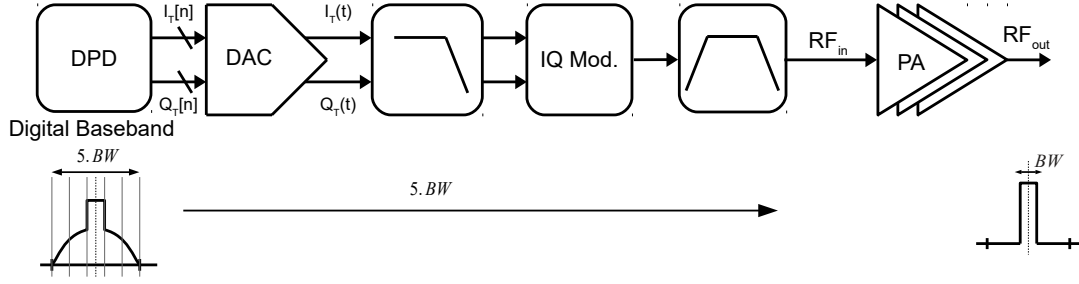
FIGURE 2.2: Illustration of a BS transmitter employing DPD system

the PA, PD modeling is also a major part of the implementation. This section presents an overview of various existing memory-unaware and memory-aware DPD techniques available in the literature in a chronological way, later discussing the advantages and disadvantages of it. Note that ≤5 MHz input signal BW systems are considered as narrowband systems here. Also, unless stated otherwise the techniques are mainly applied for the BS HPA.

2.3.1   Memory-Unaware DPD

Memory-unaware models can only correct static nonlinearities of the PA and not the dynamic nonlinearities or memory effects of the PA and are mainly used in narrowband predistortion. They are either Memory Less Look-Up Table (ML-LUT) based implementations or polynomial models based. ML-LUT implementations dominate the memory-unaware predistortion scenario. Firstly, BS memory-unaware DPD methods are presented and later UE PA memory-unaware DPD methods are briefly discussed in this section.

One of the earliest examples is as shown in [66], adaptive predistortion for a Traveling Wave Tube (TWT) PA based transmitter was implemented for 64 Quadrature Amplitude Modulation (QAM). The implemented system has predistorted values of in-phase and quadrature component voltages of each of the 64 QAM constellation symbols in a RAM. A memory-lookup encoder obtains each input data symbol and generates the RAM addresses of the desired signal point. The corresponding stored, predistorted baseband voltage values are used. This method is custom tailored to 64 QAM and has to be completely redesigned to address other modulation formats. So this method is not suitable for the current multi-mode communication systems, where different kinds of modulation formats are employed based on different requirements.

Gain based Look-Up Table (LUT) method in [2], unrestricted to modulation format, exploits the fact that the memoryless nonlinearity of the PA only depends on the envelope power of the input signal. As shown in Fig. 2.3, the envelope power $|x(t)|^2$ of the input

signal $x(t)$ is quantized and used as the indexing parameter of the LUT. The read LUT entry value is used to generate the predistorted signal $z(t)$ by modifying the input signal to obtain a linear output $y(t)$ at the PA.
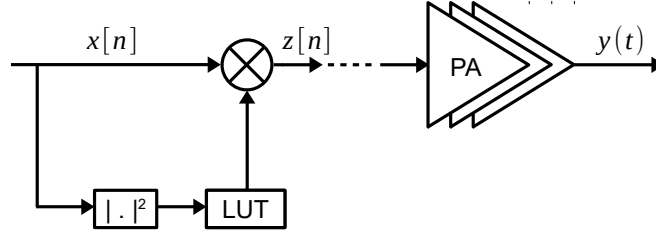
FIGURE 2.3: Gain based LUT DPD of [2]

Among various power series models polynomial model is one of the most popular model for a quasi-memoryless nonlinear PA correction. The predistorter output $z_{PD,Pol}[n]$ is given by:

$$z_{PD,Pol}[n] = x[n] \sum_{k=1}^{K} a_k |x[n]|^{k-1}, \tag{2.1}$$

where $x[n]$ is the input signal, a_k are the coefficients of the polynomial, K is the nonlinearity order of the predistorter.

In the case of UE PAs the memory effects are less pronounced [61] and hence memoryless models can suffice the PA modeling as well as predistortion, which is predominantly modeling nonlinearity [47]. Mostly this is accomplished by using simple ML-LUT.

There are two schools of thoughts contradicting each other, to employ or not to employ adaptive predistortion for UE PAs. On one hand, Asbeck group claims [63] that the DPD should be adaptive because the UE PA has large variations and mismatch in load when compared to BS DPD and is battery driven, which makes the PA distortion characteristics vary a lot. Also, in the context of Code Division Multiple Access (CDMA), the variation of the output power is in the excess of 70 dB which changes the PA linearity characteristics for different power modes [63]. On the other hand, many authors have done PD implementation using a simple ML-LUT and recent publications have done it in a open-loop fashion claiming that in the UE applications, it is difficult to reconstruct the real-time DPD because of the sizes of additional circuits and their power consumption as detailed in [3, 7].

In [63], a fast-real time adaptive DPD system (RT-ADPD) is shown and various issues associated with the UE PA are discussed. The DPD is based on LUTs, one for amplitude and one for phase. The main differences are the usage of IF ADC, to avoid IQ imbalance coming from RF IQ modulator and converting the digitally down converted IQ data into polar form by using CORDIC (COordinate Rotation DIgital Computer) algorithm. Similarly, the quadrature Baseband (BB) signal is converted into polar form and is

compared with the output after time alignment and the LUTs is adaptively adjusted. The adaptation takes less than $50\,\mu$s. It is implemented on an Field Programmable Gate Array (FPGA). The power consumption estimate of the FPGA implementation is not given as it is a prototype and has been mentioned that the optimized IC design realization will have very less power and the estimate is beyond the scope of the authors. It is also mentioned that usage of DPD reduces the PA power consumption by around 350 mW and increases the Power Added Efficiency (PAE) by 10% [65]. The performance details are presented in the Table. 2.1.

[3] presents a UE PA DPD with a simple ML-LUT as shown in Fig. 2.4, which consists of 128 entries, and it is extracted at the peak average output power level. The power level is scanned from 15 dBm to 31 dBm and a linear operation of PA is assumed below the level. Control signal based on the average output power is used to index the ML-LUT DPD, to address the large dynamic range in the UEPA scenario. The DPD improves the Adjacent Channel Power Ratio (ACPR) from -29 dBc to -37 dBc, by 8 dB, at an average output power of 28 dBm over the entire range of PA operation, i.e., 1.7–2.0 GHz. The PA delivers a gain of 16–18 dB, with the PAE of 41.1–42%.
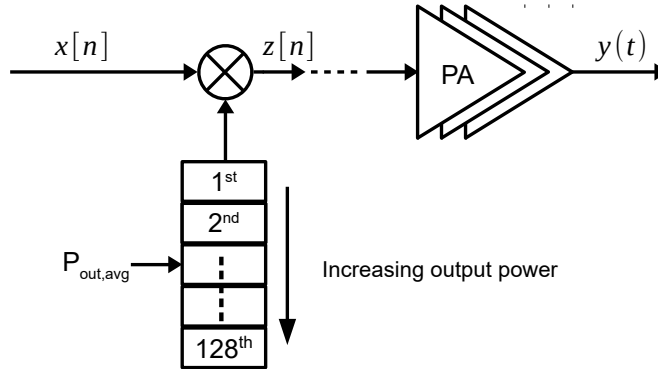
FIGURE 2.4: LUT DPD indexed by average output control signal [3]

2.3.2   Memory-Aware DPD

This section presents memory-aware DPD methods, which are usually used in the context of BS HPA linearization. Similar to memory-unaware DPD implementations, memory-aware DPDs also fall under LUT methods or models with nonlinear basis functions [67]. Volterra series proposed by Italian mathematician Vito Volterra in the year 1887 [68, 69] can perfectly model nonlinearity as well as linear and non-linear memory effects of any non-linear system [61, 70, 67], which in our case is a DPD. It uses polynomial basis functions to describe nonlinearity and memory effects. The output signal of the Volterra

series DPD is given by $z_{PD}[n]$:

$$z_{PD}[n] = \sum_{k=1}^{K} \sum_{m_1=0}^{M} \cdots \sum_{m_k=0}^{M} a_k(m_1, ..., m_k) \prod_{i=1}^{k} x[n - m_i], \qquad (2.2)$$

where $x[n]$ is the input signal, $a_k(m_1, ..., m_k)$ are the model coefficients, also called Volterra kernels, K is the nonlinearity order and M is the memory depth of the predistorter. The computational complexity of the model is very high for implementation of high memory order and nonlinearities and hence prohibitively high cost of computation and training in the context of linearization of small-cell BSs.

Hence less complex alternative implementations which are derivatives of Volterra series such as Memory Polynomials (MP) [71, 42] are employed in most of the recent works [72, 73]. Memory polynomial model is a baseband model, derived from Volterra using narrowband approximation. Narrowband approximation assumes the signal bandwidth is small compared to that of the RF signal carrier frequency. Also, it doesn't have the cross terms, when compared to Volterra series [43]. Cross terms refer to the product involving samples with different time shifts of their signal envelope samples. MP predistortion model's output $z_{PD,MP}[n]$ is given by:

$$z_{PD,MP}[n] = \sum_{k=1}^{K} \sum_{m=0}^{M} a_{km} x[n - m]|x[n - m]|^{k-1} \qquad (2.3)$$

where $x[n]$ is the input signal, a_{km} are the model coefficients, K is the nonlinearity order and M is the memory depth of the predistorter. The model is linear-in-parameters and its coefficients can be identified by indirect learning approach using least squares method [42] and can be made adaptive as well.

The MP simplification is effective but with expanding bandwidths, it has been found that the MP needs more Volterra cross terms to expand its capability. Generalized Memory Polynomial (GMP) is one such method where the delay terms adjacent to the diagonal in the matrix are also considered. The adjacent delay terms comes from the upper and lower diagonals and adds lagging and/or leading exponential envelope terms as explained in [43]. The output of GMP DPD is given by:

$$\begin{aligned} z_{PD,GMP}[n] = &\sum_{k=0}^{K_a-1} \sum_{l=0}^{L_a-1} a_{kl} x[n - l]|x[n - l]|^{k} \\ &+ \sum_{k=0}^{K_b} \sum_{l=0}^{L_b-1} \sum_{m=1}^{M_b} b_{klm} x[n - l]|x[n - l - m]|^{k} \qquad (2.4) \\ &+ \sum_{k=0}^{K_c} \sum_{l=0}^{L_c-1} \sum_{m=1}^{M_c} c_{klm} x[n - l]|x[n - l + m]|^{k}. \end{aligned}$$

In the above equation, there are three polynomial components. The first polynomial is based on time-aligned input signal and its envelopes, which is memory polynomial term with the nonlinearity order K_a and the memory depth L_a. The second one is based on the input signal and its lagging envelopes, with the nonlinearity order K_b and the memory depth L_b and lagging envelope cross-term depth of M_b. The third polynomial component is based on the input signal and its leading envelopes, with the nonlinearity order K_c and the memory depth L_c and lagging envelope cross-term depth of M_c. The advantage of GMP model is that it is still linear-in-parameters and indirect learning with least square estimation technique can be used to derive its coefficients. The disadvantage is that a large number of coefficients are required to see a noticeable linearization performance gain in comparison with MP DPD.

The paper [43] presents the effectiveness of the GMP for a 11-carrier CDMA input signal with a total bandwidth of approximately 15 MHz. Digital-to-Analog Converter (DAC) and ADC evaluation boards were employed with a 30 W PA at 2.14 GHz. With a MP configuration, an ACLR of about 52.5 dB and up to 54 dB is achieved with 20 coefficients and 40 coefficients, respectively. With addition of cross terms further ACLR improvement is shown possible up to 58 dB as the tap coefficient number increases to 68.

Similar simplifications of Volterra series has been recently presented in [74] known as Dynamic Deviation Reduction (DDR) by pruning the Volterra model in a systematic manner. Pruning is the process of keeping only the terms with noticeable impact and discarding other terms.

Other different methods exploit the Volterra series, like the memory fading Volterra series [70], which forces the memory depth to decrease with increasing kernel order. It has been found that the memory terms of higher order polynomial components of the Volterra series do not contribute as much as the lower order terms. Because of it, there is a 97% reduction in the terms when compared to a full Volterra model consisting of a 7th order kernel with memory depth 4. This is demonstrated in a very high power base-station LDMOS Doherty amplifier (350 W) for single and two-carrier WCDMA signal.
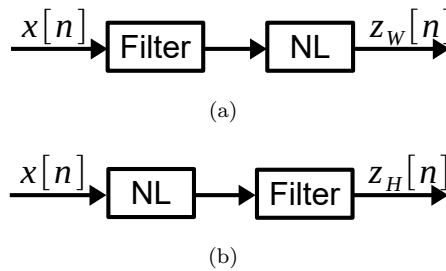
FIGURE 2.5: Two-box DPD models (a) Wiener model and (b) Hammerstein model

Simpler predistorters could be realized by a combination of linear filters (Finite-Impulse-Response (FIR) or Infinite-Impulse-Response (IIR)) and static nonlinearity such as polynomial or LUTs, commonly known as two-box models. LUTs or polynomial model the nonlinear distortion components and the filters model the memory effects. Wiener and Hammerstein models are based on such structure. As shown in Fig. 2.5(a), in Wiener model, the filter is followed by a nonlinearity block, for example, a LUT in [75] and in the Hammerstein model, shown in Fig. 2.5(b) the filter is preceded by a nonlinearity. The comparison between the two models show that the Hammerstein class of predistorters are superior in performance to that of the Wiener [76]. The system identification can be done iteratively [77] or in a two step method: first the LUT entries for each input is identified and then the filter coefficients are identified by de-embedding the input and output waveforms of the FIR filter as explained in [76]. The linearization performance of these two-box predistorters can be increased by combining them to form three-box models, as depicted in Fig. 2.6, called Wiener-Hammerstein and Hammerstein-Wiener. The aforementioned three-box models are still nonlinear-in-parameters and the estimation of parameters is harder than that for the two-box models [43].

(a)

(b)
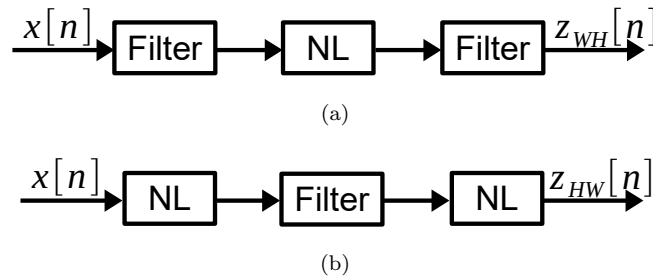
FIGURE 2.6: Three-box DPD models (a) Wiener-Hammerstein model and (b) Hammerstein-Wiener model

The Wiener and Hammerstein model's memory modeling accuracy could be augmented by adding one more parallel branch to the filter. The parallel branch implements multiplication of the filter block's input signal with its envelope, i.e., $x(n)|x(n)|$ and then the output of it is filtered with another filter and finally added to the main filter's output. This additional path to the existing filter can efficiently model the memory effects coming from bias circuits, impedance variations and harmonic loading [78, 79].

Filter base LUT (FLUT) is presented as a low cost alternative solution to MP DPD, whose structure is analogous to that of a Hammerstein predistorter [4], as shown in Fig. 2.7. The difference being the usage of filter codebook implementing multiple filters instead of a single one in the case of Hammerstein predistorter.

Twin Nonlinear Two-Box (TNTB) models are also proposed [80] which reduce the MP model dimension by upto 50%. This is achieved by adding ML-LUT to the MP. The paper presents three configurations: forward, reverse and parallel TNTB models. A
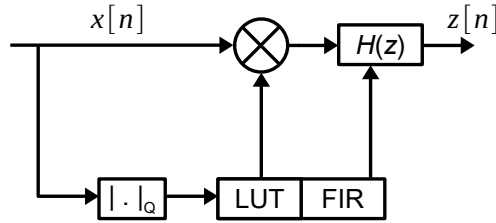
FIGURE 2.7: Block diagram of FLUT DPD of [4]

high power Doherty PA was linearized using a normal MP DPD having a 4-branch, with a 12th-order nonlinearity in each branch and a parallel TNTB model (12th-order static nonlinearity for the LUT and 4-branches with 3rd order nonlinearity for the memory polynomial block. The identification is slightly complex and involves two steps of identifying the highly nonlinear static and the mildly nonlinear dynamic behaviors, which is performed successively in the case of the TNTB models .

Artificial Neural Networks (ANN) could also be employed for a BS HPA as shown in [81], where linear and non linear neuron models are used hence making it not suitable for training the model in a linear way.

Bandlimited [82] and undersampled DPD [60] solutions are emerging, aimed to solve the problem associated with the observation path's limited bandwidth, and increasing ADC power consumption and cost because of the increasing bandwidths.

2.3.3 Advantages of DPD

DPD leverages the power of digital signal processing, which is immensely robust with ever reducing cost per operation coming from the shrinking of the transistor dimensions. From around three decades, DPD has been contributing to the PAs efficiency improvement in the BS scenario. Base-station has a very rigorous spectral mask requirements and can never be compromised at any cost. Though DPDs are power hungry, compared to rudimentary analog techniques like feedforward, DPDs perform immensely well, as explained in [43], 3%-5% power efficiency of a normal PA employed for WCDMA system could have 6%-8% with feedforward and an efficiency and 8%-10% with DPD. With the technology scaling and smarter implementation of memory polynomial based DPD through look-up-table approach, the cost of employing DPD becomes significantly cheaper. For example, the DPD presented in [46] consumes a meager $40\,\text{mW}$, for predistorting a $20\,\text{MHz}$ LTE signal.

2.3.4   Disadvantages of DPD

All the components present in the implementation path starting with the digital baseband, where the DPD is implemented must be able to handle the wide signal bandwidths of the predistorted signal. This is due to the added distortion products, a general rule of thumb is that the bandwidth is at least five times the signal bandwidth in the entire loop. Though shrinking dimensions of transistors has bought cost advantage to the DPD, the ever increasing signal bandwidths of communication systems pose significant challenges to the design of the implementation path and their associated power consumption, essentially becoming a major bottleneck. The digital dynamic power consumption is proportional to the clock frequency and hence, the power overhead because of the DPD in the digital baseband scales up according to the frequency. Not only the power overhead increases with the increasing bandwidths but also gives rise to timing closure challenges in its digital implementation [67]. A typical zero-IF transmitter consists of DACs in I and Q path, and the subsequent components in the signal chain, namely, reconstruction/anti-imaging filters in I and Q path, IQ modulator and the subsequent RF band-pass filter. All of these components must be now able to support the wide instantaneous bandwidth, along with the DPD in the baseband. Hence, for the case of the LTE-advanced system with bandwidth in excess of 100 MHz, each of the I and Q paths should be able to handle at least 250 MHz. Each of the DACs should be clocked above 500 MHz instead of 100 MHz. So, because of the increasing bandwidths, DPD solutions power overhead can become prohibitive for a small-cell BS PA. And hence the above points have to be carefully considered to reduce the power overhead during all the levels of the system design.

TABLE 2.1: Survey of various DPD systems with various performance metrics

Reference	BW (MHz)	|ACLR| (dBc)	Complexity[1]
Morgan 2006 [43]	15	58	7
Wood 2010 [70]	15	60	8
Liu 2005 [78]	10	50^2	6
Mkadem 2011 [81]	20	50	10
Hammi 2009 [80]	20	50^2	5
Cho 2014 [3]	3.84	37	1
Presti 2012 [63]	3.84	40	3

[1] Approximate computational complexity on a scale of 1 to 10, where 1 being the lowest and 10 the highest.
[2] Approximate value for comparison purpose.
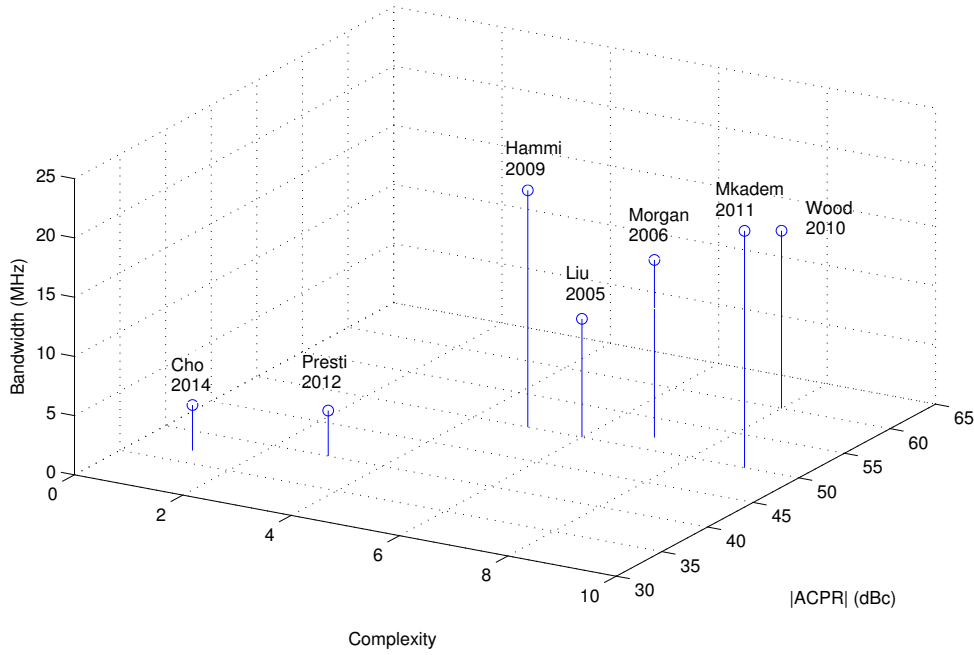
FIGURE 2.8: 3D plot of various DPD systems

2.3.5   Conclusions on DPD

Digital Predistortion has been de facto the industry standard for the PA linearization
in high power BS's and is also being used in narrow-band UE PAs. There exits wide
varieties of DPDs classified into memory-unaware or memory-aware DPDs, with varying
computational complexity and ACLR correction performance. Advantage of memory-
unaware DPDs such as LUTs are, their implementation easiness. Disadvantages are
the resulting limited accuracy since the memory effects of the PA are unaccounted as
the name suggests. They additionally suffer from quantization effects because of the
finite size of LUT entries. In the case of narrowband UE and small-cell BS context
LUTs could be sufficient for predistortion, since the PA has very less memory effects and
there is a stringent power constraint. Memory-aware DPD can model and correct both
nonlinearities and memory-effects. Though full Volterra series is capable of correcting
very strong nonlinearities, the computational complexity is high and the implementation
becomes costly.

A simple comparison of the DPD methods is shown as a 3D plot in Fig. 2.8 and the
corresponding data is presented in Table. 2.1. We can observe that the complexity of
the predistorter increases with the increasing linearization performance over wideband.
Two-box models such as TNTB models proposed in [80] could be promising technique,
which can achieve better performance, with comparatively less computational complexity.

2.4   Analog Radio Frequency Predistortion

Though analog predistortion has been around since the generation of TWT power amplifiers [83, 84], but because of the improvements in digital domain in the past few decades and the advantages in DPD, analog PD has been overshadowed by DPD. This section presents various ARFPD methods available in the literature.

For the case of a generic ARFPD the correction signal is synthesized in the analog baseband using the envelope of the PA RF input and the predistortion is performed in the RF domain using the input and corrected signals. Fig. 2.9 shows a transmitter chain employing a generic ARFPD solution. For the case of ARFPD, to correct the inter-modulation distortion components of the PA, the spectrum regrowth of 5X the input signal bandwidth occurs just before the PA, and hence the transmitter chain can be left mostly unaltered, supporting only the signal bandwidth. This makes ARFPD a very attractive alternative to DPD solutions. This is the reason why the predistortion is preferred in the RF section rather than analog baseband, though analog baseband and IF predistorters exist in the literature.
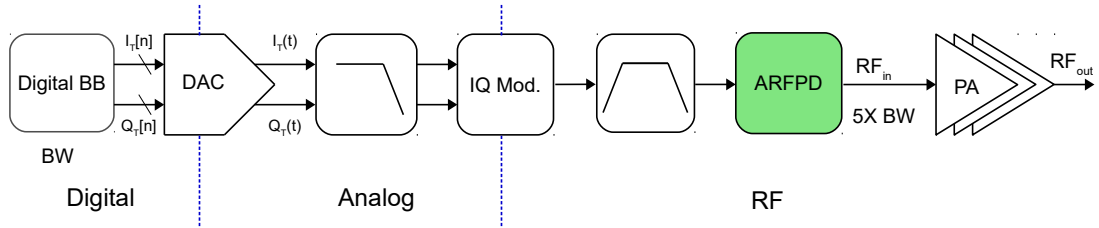
FIGURE 2.9: Illustration of a BS transmitter employing ARFPD system

Appendix A presents the basic principle of canceling-out the nonlinearity using an ARFPD, with the help of a two-tone RF signal, when a polynomial predistorter is employed. A further detailed ARFPD system block diagram is shown in Fig. 2.10. The envelope of the RF signal $X_{RF}(t)$ is extracted and a complex analog predistortion (APD) is applied, which generates the required nonlinearity. Vector modulator functions as a complex gain adjuster, modifying the gain and phase of the input RF signal. The τ_{delay} usually obtained with a delay line compensates for the delay that occurs as the signal envelope traverses through the work function. The obtained predistorted RF signal $Z_{PD}(t)$ produces linear signal $Y_{RF}(t)$ at the PA output. This is because with an ideal predistorter output when fed as the input to the PA produces IMD components whose amplitude is equal but in anti-phase (180° of phase difference) to that of the IMD components when $X_{RF}(t)$ is directly input to the PA.

Along with the basic ARFPD method explained before, there are three ways of implementing predistortion in the RF domain, which are as follows:
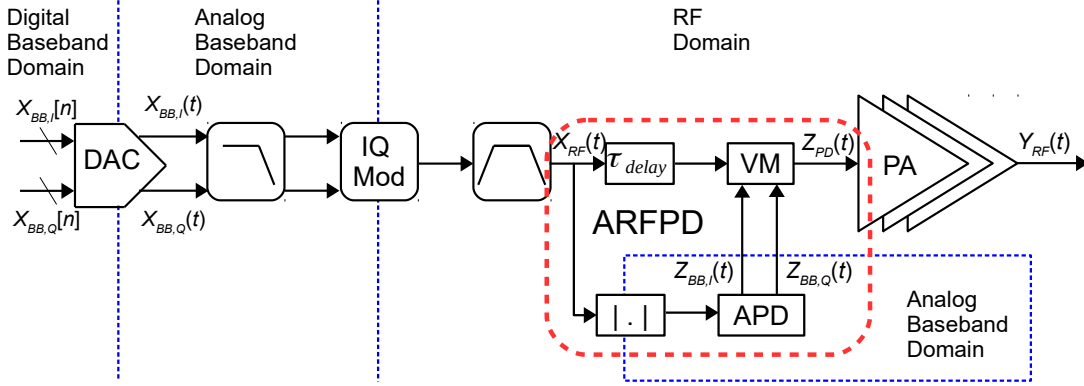
FIGURE 2.10: Block diagram of transmitter with ARFPD system

- LUT method: LUT based predistorter providing quadrature correction signals [5, 6] or amplitude and phase correction signals [85]

- *Work function* method implementing polynomial basis function in Analog baseband as discussed before [86, 87, 41, 8] or in or IF domain [88]

- Nonlinearity generation through usage of a pair of diodes connected in anti-parallel [84, 89, 90]

2.4.1   Memory-Unaware ARFPD

Fig. 2.11 shows an illustration of a LUT based predistorter providing quadrature correction signals [5, 6]. Digital control words are output from the LUTs in I and Q paths, which are converted to analog signals with the help of DACs followed by reconstruction filters. The analog signals are the correction signals, which are multiplied with the original undistorted RF signals using a RF vector multiplier to generate the predistorted RF signals. An RF vector multiplier is the combination of the polyphase filter, multiplier and an adder. Polyphase filter generates the in-phase and quadrature RF signals from a modulated RF signal. DACs and reconstruction filters in the I and Q channels should support at least 5X input signal bandwidth. Also, the LUT is indexed using the quantized RF signal envelope,which requires an envelope detector and an ADC. Hence, the power overhead of utilizing this architecture can be high.

Simple ML-LUT is presented in [7], where the linearization is performed in RF domain, by modifying the driver stage of the PA by using variable gain amplifiers, with programmable bits (m by n) for selecting the binary weighted cells with the contents of the ML-LUT, which are 5 bit control words, as shown in the Fig.2.12. This approach eliminates the need of DACs in I and Q paths, which demand at least 5X bandwidth requirement as explained in the previous architecture.
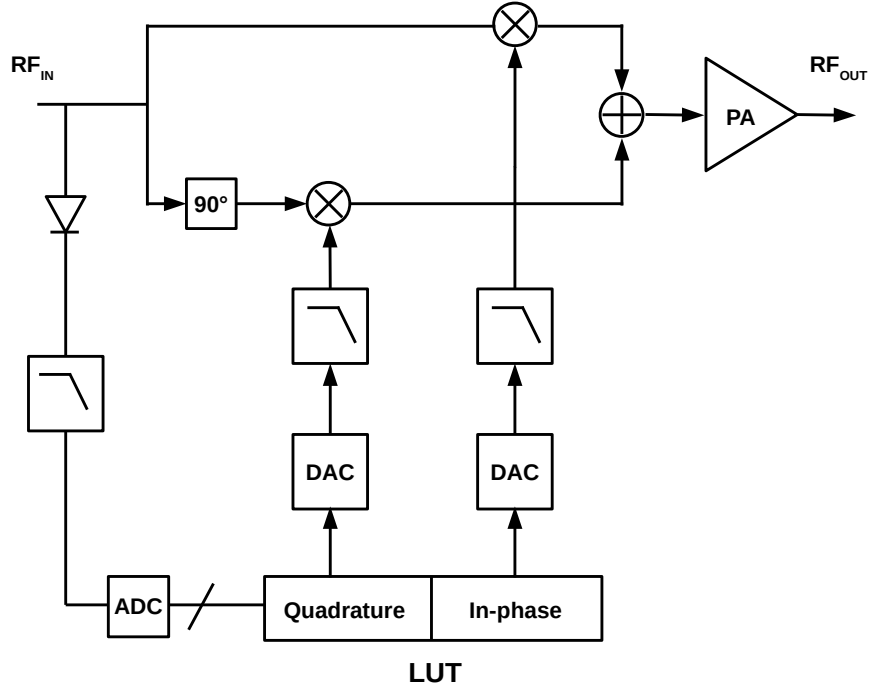
FIGURE 2.11: LUT based RF predistorter of [5, 6]
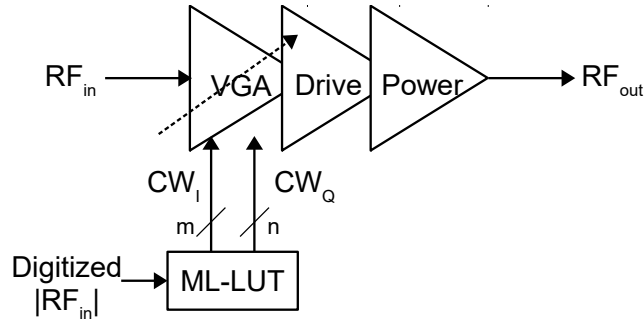
FIGURE 2.12: RFPD based PA driver stage of [7]

An analog 5th order polynomial predistorter with programmable coefficients has been presented at ESSCIRC 98 in [88] where a 0.8 μm BiCMOS process IC is shown, which mitigates the IMD3 by 20–25 dB at an IF frequency of 20.4 MHz. The PA was already fairly linear class A. In [91] over 1–3 MHz signal BW could be linearized for class A-C PA in BB or IF. The major drawback was that these predistorters could not model memory effects as good as digital techniques and hence only suitable for narrowband smoothly compressing AM-AM curves. Also they were demonstrated with just two tones, which is much less sensitive to memory effects compared to a modulated signal. Similar realizations were demonstrated, without incorporating memory effects but at very low power consumption, core working with just 2 mA on 2.7 V supply for linearizing IS-95 and WCDMA signals at 200 MHz IF [41].

2.4.2   Memory-Aware ARFPD

To address the memory effects, envelope memory polynomial (EMP) model [92] has been practically used to implement the ARFPD predistorter IC in 180 nm CMOS in [8]. EMP model is a further simplified formulation of MP and hence has a limited correction performance, even after using a higher number of coefficients (nonlinearity order and memory depth) [10]. The output of an EMP predistorter in digital baseband is given as follows:

$$z_{PD,EMP}[n] = x[n] \sum_{k=0}^{K-1} \sum_{q=0}^{Q} a_{kq}|x[n-q]|^k \qquad (2.5)$$

where $x[n]$ is the baseband complex input signal, a_{kq} are the model coefficients (complex), K is the nonlinearity order and Q is the memory depth of the EMP predistorter. As can be seen from (2.5), the predistorter now only needs the current sample and just the magnitude or envelope information of current and past samples, according to the memory depth Q.

The output of an EMP predistorter in analog baseband is given as:

$$z_{PD,EMP}(t) = x(t) \sum_{k=0}^{K-1} \sum_{q=0}^{Q} a_{kq}|x(t-t_p)|^k \qquad (2.6)$$

The EMP predistorter output in RF can be obtained from baseband by the following equation:

$$z_{PD,EMP,RF}(t) = \text{Re}\left\{z_{PD,EMP}(t)e^{jw_ct}\right\}, \qquad (2.7)$$

$$= \text{Re}\left\{\left[x(t) \sum_{k=0}^{K-1} \sum_{q=0}^{Q} a_{kq}|x(t-t_p)|^k\right] e^{jw_ct}\right\} \qquad (2.8)$$

Let us assume:

$$a_{kq} = a_{kq,R} + ja_{kq,I} \qquad (2.9)$$

where $a_{kq,R}$ and $a_{kq,I}$ are real and imaginary parts of a_{kq}. Substituting (2.9) in (2.8) gives:

$$
\begin{aligned}
z_{PD,EMP,RF}(t) = &\text{Re}\left\{x(t)e^{jw_ct} \sum_{k=0}^{K-1} \sum_{q=0}^{Q} a_{kq,R}|x(t-t_p)|^k\right\} \\
&+ \text{Re}\left\{jx(t)e^{jw_ct} \sum_{k=0}^{K-1} \sum_{q=0}^{Q} a_{kq,I}|x(t-t_p)|^k\right\}
\end{aligned}
\qquad (2.10)
$$

$$z_{PD,EMP,RF}(t) = \sum_{k=0}^{K-1} \sum_{q=0}^{Q} a_{kq,R} \text{Re} \left\{ x(t) e^{jw_c t} \right\} |x(t - t_p)|^k$$
$$+ \sum_{k=0}^{K-1} \sum_{q=0}^{Q} a_{kq,I} \text{Re} \left\{ x(t) e^{j(w_c t + \frac{\pi}{2})} \right\} |x(t - t_p)|^k \quad (2.11)$$

From (2.11), we can see that the predistortion signal in RF can be obtained from the current RF signal as well as current and the lagging envelopes of RF signal.
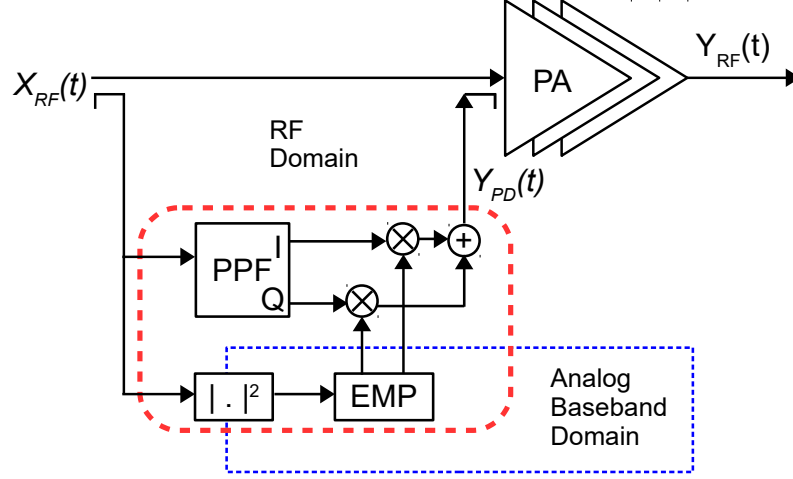
FIGURE 2.13: Block diagram of the ARFPD system of [8]

In a practical implementation of ARFPD as shown in [8], a fourth order memory and 11th order nonlinearity EMP was implemented. A simplified block diagram of the predistorter is as shown in Fig. 2.13. The necessary quadrature RF signal was obtained by using analog polyphase filter (PPF), the EMP coefficients were given with the help of current steering DACs, which also performs multiplication with the signal envelope and its delayed versions. Finally, inside the EMP block, the in-phase and quadrature components are separately generated and multiplied back with the corresponding PPF outputs and added back, which makes the RF predistorted signal. The IC is intended for base station RF Power Amplifiers. It achieves a power consumption of 200 mW for implementing the predistorter. The signal path occupies 4 mm^2, and the power supply is 1.8 V. It was mentioned that 65% of that power is consumed by the RF circuitry.

The EMP based ARFPD has a major shortcoming that is to not be able to properly address linear memory effects of the PA. To improve the performance of the EMP based ARFPD system, EMP model can be further assisted by the help of an FIR filter in digital baseband, as proposed in [10, 9, 93], which is as shown in Fig.2.14. The FIR filter is used to compensate for the linear memory distortion of the PA, which is poorly modeled by the original EMP model, since EMP is an oversimplification version of MP, which in turn is derived from Volterra series through simplification. The digital baseband

equivalent of the ARFPD based on FIR-EMP is given by:

$$z_{PD,FIR-EMP}[n] = \sum_{l=0}^{L} h_l x[n-l] \times \sum_{k=0}^{K-1} \sum_{q=0}^{Q} a_{kq} \left| \sum_{l=0}^{L} h_l x[n-l-q] \right|^k \qquad (2.12)$$

where L is the FIR filter order and h_l are the filter coefficients, the rest of the variables being the same as mentioned in the former equations. The results of the predistorter performance are detailed in Table 2.2. Addition of an FIR filter improves the ACLR in high bandwidth case, by about 8.5 dB for 80 MHz case, when compared with EMP modeling [9].
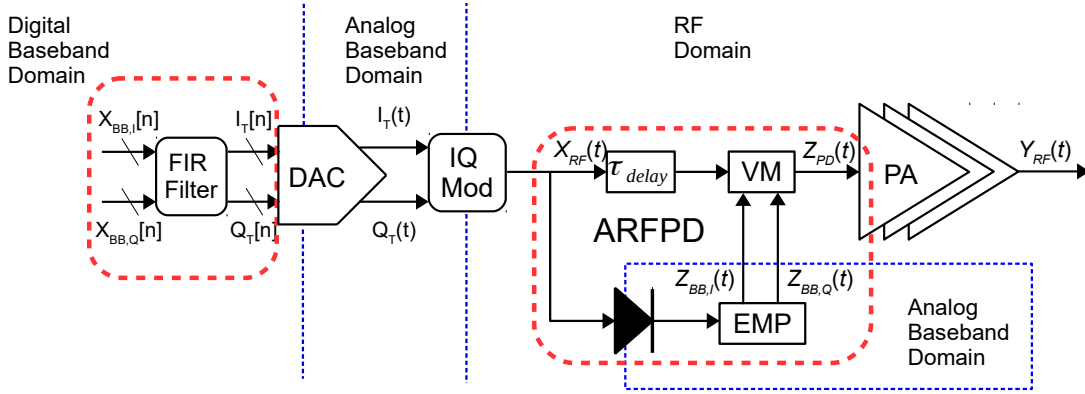
FIGURE 2.14: Block diagram of the FIR-EMP ARFPD [9, 10]

The coefficients of the FIR filter can be calculated using nonlinear estimation methods such as Newton iterative method, as used in for example [94] or a two-step algorithm known as, small signal assisted parameter identification (SSAPI), as shown below, [10]:

- Firstly, the forward model of the PA is estimated using the output of the PA and its input modulated signal

- The forward model is fed with small signal training data and the output is measured

- Based on the small signal data, the coefficients of the FIR filter are derived using the least square estimation algorithm

- With the help of FIR filter output data the EMP model can be derived

The state-of-the-art ARFPD systems employ Envelope Memory Polynomial (EMP) model for predistortion and not the MP model. The reason for not using memory polynomial in ARFPD is because the implementation path demands multiple RF delay elements and the same number of RF vector modulators, which is equal to the considered memory depth for the correction. The RF components are very power hungry and are tough to

design for higher accuracy because of their sensitivity to process, voltage and temperature (PVT) variations. One such component in the ARFPD is the RF vector multiplier, which tends to show severe nonidealities. Especially, with the increasing signal bandwidths, it not only exhibits narrowband nonidealities, such as mismatch and offsets between I and Q paths but also frequency dependent wideband nonidealities. The nonidealities call for very complex wideband compensation techniques as discussed in [93].

2.4.3 Advantages of ARFPD

There are several advantages in employing ARFPD for PA linearization. ARFPD implementation relaxes the overall requirements on the transmitter chain since the digital baseband signal does not experience bandwidth expansion, the digital baseband can be clocked at normal clock rates, the same goes for the DAC and the subsequent filters. The relaxed specifications of the entire transmit chain results in a very low overall power consumption. The other benefit of ARFPD systems, where the linearization is performed in the RF domain, is that the ARFPD can be used to linearize any existing base station PA without predistorter, since the inputs to the ARFPD are only the RF input and attenuated RF output of PA and the output of the ARFPD can be used as the PA input.

2.4.4 Disadvantages of ARFPD

There are also several disadvantages. Though theoretically complex models such as MP and GMP models can be used in ARFPD, from the implementation point of view, only EMP model has been used in ARFPD [10, 92]. This results in limited performance when compared to state-of-the-art DPD performance. Also, since ARFPD is implemented in analog way, various kinds of problems inherent to analog implementations like noise, mismatch, offsets, PVT variations and so on has to be addressed carefully. This results in challenging circuit design. Because of the inherent nonlinear structure of the predistorter, there exists signal amplitude expansions and compressions and also bandwidth expansions internally, which needs to be carefully handled. While the PVT variations and signal expansions could be kept under control using replica biasing stages, it is not straight forward [8, 95].

2.4.5 Conclusions on ARFPD

Similar to the case of DPD, ARFPD can be memory-unaware or memory-aware. Predistorter implementation in analog domain paves way for overall low power consumption.

Except for the design complexity and challenges involved, ARFPD has significant benefits when compared to DPD. With the recent publications showing decent correction performance, ARFPD can be a good candidate to achieve low power PD in small-cell BS and UE PA scenarios.

TABLE 2.2: State-of-the-art survey of predistortion systems

Reference	Context	Power Amplifier		Predistorter		Fc (GHz)	BW (MHz)	Signal	PAPR (dB)	ACLR1 (dBc)
	BS/UE	Type	Power	Domain	Model			Type		
Morgan 2006 [43]	BS		30 W	DPD	GMP	1.9 –2.1	15	11C CDMA		58
Wood 2010 [70]	BS	LDMOS Doherty	350 W	DPD	Memory Fading Volterra	2	3.84, 15	WCDMA	10	60
Liu 2005 [78]	BS	GaAs	60 W	DPD	Augmented Wiener	1.96	10	2C WCDMA		
Mkadem 2011 [81]	BS	Doherty	250 W	DPD	ANN	2.14	20	4C WCDMA	7.24	50
Hammi 2009 [80]	BS	Doherty	300 W	DPD	TNTB	2.14	20	4C WCDMA		
Roger 2013 [8]	BS	Doherty		ARFPD	EMP	2.14	10	2C WCDMA	7.5	52
							10	6C-GSM	6.5	62
							20	LTE	9	49.9
Huang 2015 [9]	BS	GaN Doherty	20 W	ARFPD	FIR EMP	2	20	LTE	9.3	51.5
							40	3C WCDMA +LTE	8.4	46
							80	4C WCDMA +LTE	8.3	46.9
Cho 2014 [3]	UE	CMOS	28 dBm	DPD	ML-LUT	1.7 –2.0	3.84	WCDMA	3.24	37
Presti 2012 [63]	UE		30.9 dBm	DPD	RT-ADPD	1.95	3.84	WCDMA	3.3	40
Son 2012 [7]	UE	CMOS	29.1 dBm	ARFPD	ML-LUT	1.96	3.85	WCDMA	3.4	36

2.5   Comparison of DPD and ARFPD

Table 2.3 shows the summary of comparison between DPD and ARFPD. From the small-cell base station perspective, the overall system power consumption is the major determinant. Though DPD provides best linearization performance, ARFPD provides good linearization with less overall system power consumption. Also, it is worth noting that the design of high performance ARFPD, especially IC implementations are highly challenging when compared to DPD, which is already a challenging task in itself. It is not straight-forward to conclude which among the two methods of predistortion is the most suitable candidate in the context of future small-cell 5G wireless nodes, since many factors are involved.

	DPD	ARFPD
Transmit chain BW requirement	🙁	🙂 🙂
Robust to PVT	🙂	🙁
Predistorter Power consumption	🙂 🙂	🙂
Overall Power consumption	🙁	🙂
ACLR correction Performance	🙂 🙂	🙂
Design challenges	😐	🙁

TABLE 2.3: Comparison of DPD and ARFPD in the context of small-cell BS with high bandwidth (≥ 100 MHz)

2.6   Conclusion

In this chapter, various digital and analog RF predistortion techniques have been reviewed. The two parts of the chapter on DPD and ARFPD techniques were subdivided into memory-aware and memory-unaware correction techniques. Though memory-unaware predistorters tend to obtain low-cost and low-complexity, their performance is severely impeded when the PAs tend to show increased amounts of memory affects, in the context of future wideband communications (BW ≥ 100 MHz). In the context of small-cell base stations, it is prerogative to have wideband linearization along with low-cost and low-complexity, which are contradictory requirements. Though Volterra series simplifications, such as MP, and box-oriented models such as TNTB DPD and FIR-EMP ARFPD tend

to show promising capabilities, there is still a strong need for new predistortion models in the low-power small-cell base station context.

On the other hand, it is harder to decide among DPD and ARFPD implementations. DPD banking on the robust and well-proven DSP techniques and CMOS device scaling advantages has been leading the way in the linearization scenario. The disadvantage of DPD is that the bandwidth expansion starts very early in the transmit chain resulting in a transmit chain design complexity and power consumption. Advancements in low-power ARFPD techniques have shown less overall transmitter power consumption and complexity while the performance is still limited and with open design challenges. In the Chapter. 3, we show a novel predistortion model that augments the performance of the MP and provide its ASIC implementation. Based on the developed novel predistorter model, Chapter. 4 provides an interesting mixed-signal solution that combines the advantages of DPD and ARFPD.

Chapter 3

Algorithm Level Design and Digital Implementation

3.1 Introduction

As introduced previously in Chapter 2, numerous predistortion models have been proposed in the literature. Most of the models addressing memory effects are simplifications of full Volterra model, such as memory polynomial model [71]. Memory polynomial model's complexity is highly reduced when compared to that of a full Volterra model and hence its accuracy. Nonetheless, it is still one of the most attractive predistortion models and a potential predistorter for small cell base stations, providing significant performance, usually with a very few number of coefficients. In this chapter, we show that the memory polynomial model needs higher nonlinearity order and memory depth to significantly linearize a small-cell base station PA driven with high bandwidth input signals. We propose an elegant way to mitigate and further improve its performance.

This chapter describes the proposed new low-complexity digital baseband predistorter with an FIR filter preceding a memory polynomial, called as FIR-MP. The predistorter is targeted towards wideband small-cell base stations. We first present the conventional memory polynomial predistorter and its shortcomings in Section 3.2, and then the proposed FIR-MP predistorter is presented, to mitigate the shortcomings of the memory polynomial predistorter. A commercial small cell 1 W GaAs HBT PA (ADL5606) model has been used to perform the assessment of the predistortion algorithms. We show the methodology used to extract the model of the PA in Section 3.3. Section 3.4 presents detailed explanation of the procedure used to estimate the coefficients of the FIR-MP predistorter. With the help of predistortion simulations performed on the ADL5606 PA's extracted MP models, we present an optimal dimensioning of the predistorter

and the subsequent DACs for MP and FIR-MP DPDs in Section 3.5. The DACs are used to convert the baseband digital signal to analog in the I and Q paths. The digital implementation methodology used to realize the predistorter in CMOS process is shown in Section 3.6. Conclusions and summary are presented in Section 3.7.

3.2   Predistorter Modeling

The choice of the model used for predistorter plays a key role in determining the implementation complexity and the linearization performance, as explained in Chapter 2. From the low-power practical implementation perspective, memory polynomial is one of the best candidates for the predistorter model.

3.2.1   Conventional memory polynomial predistorter

We recall from Section 2.3.2 the output of a conventional memory polynomial predistorter [71, 42] as:

$$z_{PD,MP}[n] = \sum_{k=0}^{K_{PD}-1} \sum_{q=0}^{Q_{PD}} a_{kq} x[n-q]|x[n-q]|^k, \tag{3.1}$$

where $x[n]$ is the input signal, a_{kq} are the model coefficients, K_{PD} is the nonlinearity order and Q_{PD} is the memory depth of the predistorter. Note that for simplicity in the notation we would like to, later in the manuscript use K and Q instead of K_{PD} and Q_{PD}, respectively. For a PA nonlinearity and memory depth, the subscripts are explicitly mentioned as K_{PA} and Q_{PA}, respectively. We have considered that the predistorter and the PA model contains only even-order terms of k, since the odd-order nonlinearities are the dominant ones [96, 43, 97]. MP model can be derived from generalized Hammerstein model [43], which is given as:

$$z_{PD,GH}[n] = \sum_{k=0}^{K_{PD}-1} \sum_{q=0}^{Q_{PD}} a_{kq} x^k[n-q]. \tag{3.2}$$

As explained in [43], when we assume that the input signal bandwidth is small compared to the carrier frequency we can approximate Eq. 3.2 to only contain the terms of the form $x[n-q]|x[n-q]|^k$, which is the MP model shown in Eq. 3.1. Hence, MP model can be considered as a special case of the generalized Hammerstein model using the aforementioned narrowband approximation.
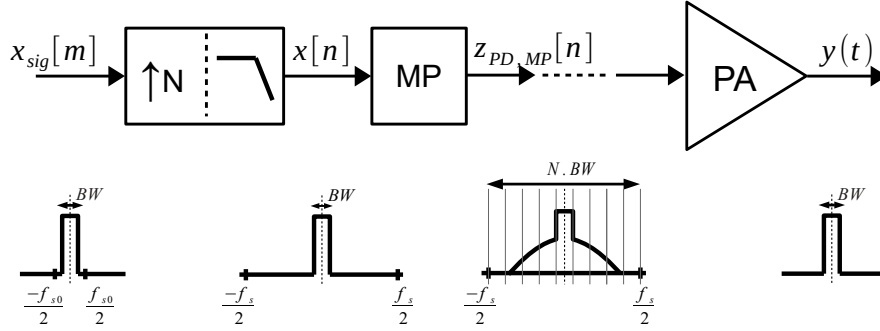
FIGURE 3.1: Illustration of MP predistorter

It should be noted that the sampling frequency of the data should be at least five times the signal bandwidth. As shown in Fig. 3.1, the original complex baseband signal $x_{sig}[m]$ with a bandwidth BW is sampled with f_{s0}. f_{s0} can be as low as the frequency defined by Nyquist criterion, i.e., BW. With the help of interpolation, which is upsampling the baseband signal $x_{sig}[m]$ by a factor N, usually five times the original signal sampling frequency and then low-pass filtering the resultant data to obtain $x[n]$, at a sampling frequency of f_s. The signal after predistortion expands in bandwidth by a factor of K times. The predistorted signal when input to PA will ideally give a linear output which occupies a bandwidth of BW, the original signal bandwidth.
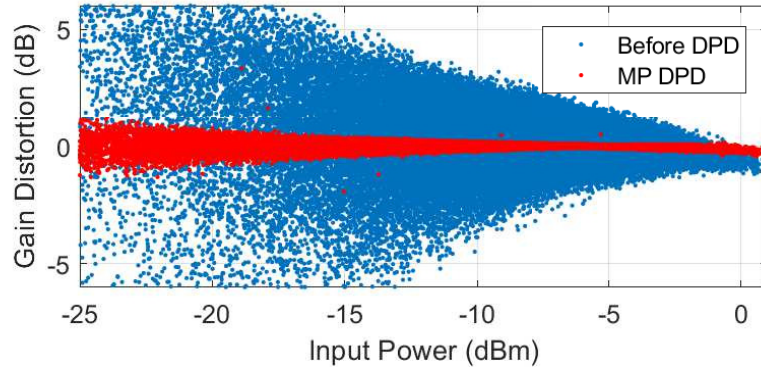
FIGURE 3.2: AM/AM plot without and with MP predistorter for a 4 carrier modulated signal with a total bandwidth of 80 MHz and PAPR of 8.4 dB

The predistortion performance of the memory polynomial is evaluated with an extracted MP model of ADL5606 PA. The results obtained, as will be discussed in detail in Section 3.5, show that the memory polynomial predistorter can linearize up to 20 MHz of input signal bandwidth, reaching an ACLR beyond 45 dBc with margin, in adjacent and alternate channels denoted as ACLR1 and ACLR2, respectively [29]. But for a higher bandwidth case of 80 MHz, the predistorter performance degrades and just meets the ACLR specification, with a very small margin. This is a consequence of the narrowband approximation, as explained previously. When the input signal bandwidths are comparable to the PA carrier frequency (2 GHz) the narrowband approximation does
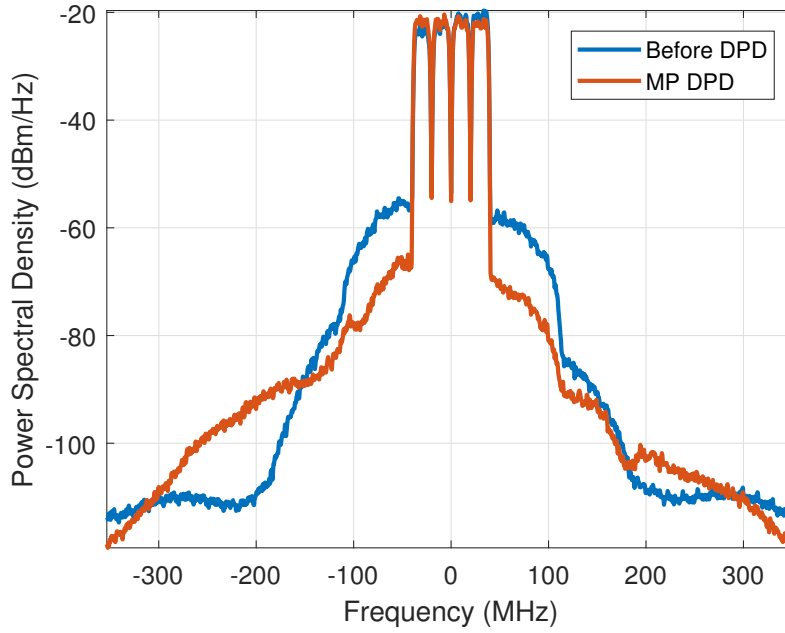
FIGURE 3.3: Power spectra of the output without and with MP predistorter for a 4 carrier modulated signal with a total bandwidth of 80 MHz and PAPR of 8.4 dB

not hold. The AM-AM (gain distortion) plot and the power spectra of the PA output, before and after predistortion, using memory polynomial predistorter are shown for a 4 carrier modulated signal with an initial bandwidth of 80 MHz in Fig. 3.2 and Fig. 3.3, respectively. It can be seen that the PA exhibits an increased amount of linear memory distortion at low powers, which in turn manifests itself as nonlinear dynamic distortion and hence degrading the predistorter performance. The ACLR1 and ACLR2 before predistortion are 34.4 dBc and 36 dBc, respectively, and after predistortion are 45.6 dBc and 48.7 dBc, respectively, when considering the DPD and DAC bandwidth equal to nine times the input signal bandwidth, as will be explained later in Section 3.5.

3.2.2   FIR Memory Polynomial Predistorter

We propose to add a linear FIR filter before the memory polynomial to mitigate the aforementioned shortcoming of the memory polynomial predistorter in addressing linear memory distortion, which is the dominant source of distortion at low input powers. As seen briefly in Section 2.4.2, the usage of FIR filter to increase the memory correction performance of an analog RF predistorter based on envelope memory polynomial (EMP) has been presented in [10]. Envelope memory polynomial is a further simplified version of memory polynomial, and hence has very limited linearization performance when compared to the conventional memory polynomial. Here we propose to use memory polynomial

instead of envelope memory polynomial to have the best linearization performance of the predistorter.

The output of the FIR filter which is used as the input to the memory polynomial predistorter is given as:

$$x_{FIR}[n] = \sum_{l=0}^{L} h_l x[n-l], \tag{3.3}$$

where $x[n]$, L and h_l is the input signal, FIR filter order and filter coefficients, respectively.

By substituting the FIR filter output (3.3) in (3.1), we obtain the FIR memory polynomial (FIR-MP) given as:

$$z_{PD,FIR-MP}[n] = \sum_{k=0}^{K-1} \sum_{q=0}^{Q} a_{kq} \sum_{l=0}^{L} h_l x[n-l-q] \\ \times \left| \sum_{l=0}^{L} h_l x[n-l-q] \right|^k. \tag{3.4}$$
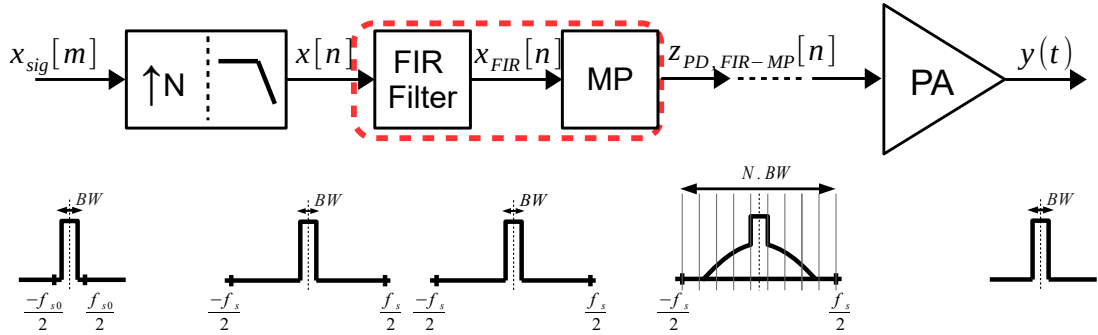
FIGURE 3.4: Illustration of FIR-MP predistorter

FIR-MP can be considered as a special case of a Wiener-generalized Hammerstein, which can have potential to augment the performance of the conventional MP model.

As shown in Fig. 3.4, even here, we note that the sampling frequency f_s of the data input of the FIR filter and the subsequent MP block should be at least five times the signal bandwidth BW. But, the FIR filter acts only on the region inside signal bandwidth and hence the spectrum does not produce any intermodulation products. For the MP block the intermodulation distortion correction terms produced will increase the signal bandwidth accordingly, usually considered five times the bandwidth.

3.3 PA Model Extraction Procedure

In order to assess the linearization performance of the proposed FIR-MP predistorter, computer simulations are performed as a proof-of-concept. As mentioned previously,

ADL5606 PA's extracted models are used for the simulations. ADL5606 PA is a commercial 1 W Gallium Arsenide (GaAs) Heterojunction Bipolar Transistor (HBT) PA from Analog Devices, Inc. (ADI), suitable for small-cell base station applications [98]. The PA has a wideband of operation from 1.8 GHz to 2.7 GHz. The extraction of the models are done for two single carrier LTE downlink signals of 20 MHz and 80 MHz bandwidths of LTE signals centered at 2.0 GHz carrier frequency. The measured PAPR is 8.75 dB and 9.3 dB, respectively for 20 MHz and 80 MHz signal cases. Though the aforementioned signals are not 5G candidates but mimic the behavior in terms of PAPR of a 5G candidate waveform, such as Filter Bank Multi-Carrier (FBMC) [99].
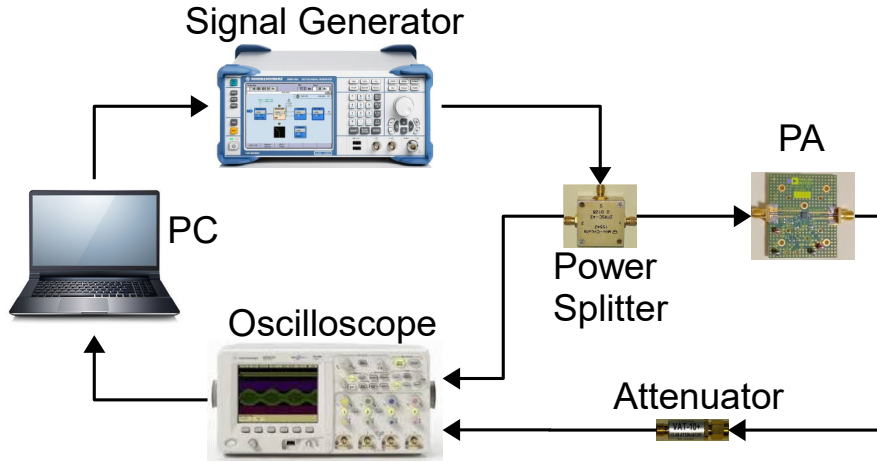
FIGURE 3.5: Measurement setup used for PA characterization

Firstly, the PA models are extracted for the aforementioned two input signal cases. Fig. 3.5 illustrates the PA measurement setup. With the baseband data generated in the computer, Rohde & Schwarz SMBV100A Vector Signal Generator is used to generate the PA input RF signal at the desired carrier frequency. The PA model extraction is done at the RF carrier frequency of 2.0 GHz. A ZFRSC-42 power splitter/combiner from Mini-Circuits [100] is used to split the input RF signal into two identical signals, one signal for the PA input and the other for the oscilloscope. The power splitter has an insertion loss of about 6 dB in the frequency of PA operation, while the coaxial cables incur around 0.4 to 0.6 dB of attenuation.

From the datasheet of the PA [98], we can observe that the PA 1 dB compression output power is 30.2 dBm at a frequency of operation at 1960 MHz, which is close to our carrier frequency of 2.0 GHz . With a PA power gain of 24.7 dB, the output 1 dB compression point translates to an input power of 6.5 dBm, while it should have been only 5.5 dBm input power if not for the gain compression arising due to nonlinearity. Any input signal beyond this power tends to undergo severe compression.

Taking into account the insertion loss and cable attenuation, the total attenuation up to the PA input amounting to around 7 dB. So the PA sees an input power of around 6 dBm when the power at the output of the signal generator is 13 dBm. This is the average power that we have provided for the two modulated signals. With a PAPR of 8.75 dB and 9.3 dB, respectively for 20 MHz and 80 MHz signals, the peaks of the signals at the PA input reaches 14.75 dBm and 15.3 dBm, respectively. This amounts of peak power drives the PA into a very strong nonlinear regions in both the signal cases. The PA models obtained at this power can be used to model the PA behavior when excited with lower input signal powers but not vice versa, hence our motivation to choose such amount of input signal power.

The attenuated output of the PA along with its input replica from the power splitter is captured using two separate channels of the Agilent technologies 54853A DSO Infiniium Oscilloscope. The oscilloscope has a capturing bandwidth of 2.5 GHz, a maximum sample rate of 20 GSPS and can handle 30 dBm maximum input average power. The PA output is attenuated by 20 dB, the reason being not to exceed the 30 dBm maximum input power at the oscilloscope and also not to excite any nonlinearities arising from the oscilloscope's internal circuitry.
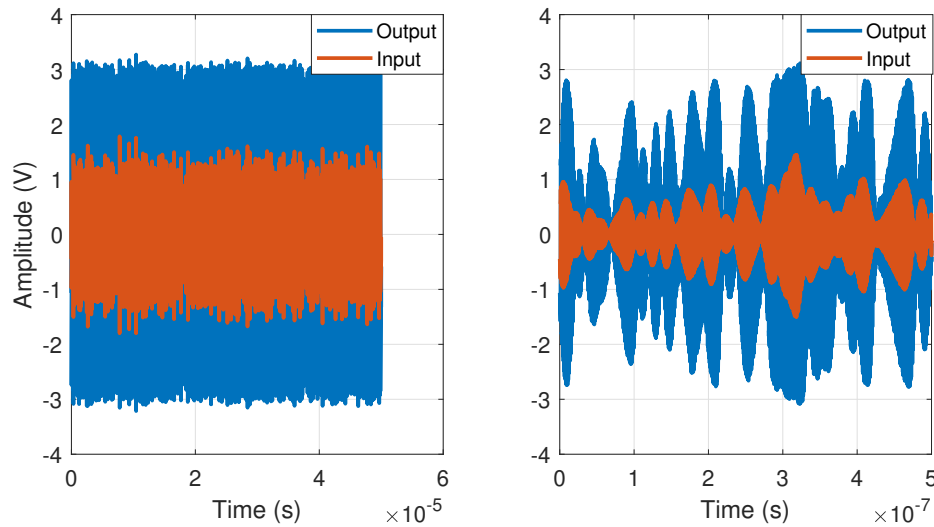
FIGURE 3.6: Measured PA input and output RF data from the oscilloscope sampled at 20 GSPS. Plots on the left is for the total captured duration, i.e., 50 μs and on the right is the time-magnified data for 50 ns duration

Over a million samples at a sampling rate of 20 GSPS were captured for each of the input and output signal at the RF frequency. The total duration of the captured signals is for 50 μs. The input and output signals, along with their time axis magnified versions (50 ns duration) from the oscilloscope for the 80 MHz signal case is shown in Fig. 3.6. The spectra are shown in Fig. 3.7. Harmonics and spurious tones are noticeable in the spectra on both the input and output signals, which is because of the nonidealities of

the test and measurement equipment, degrading the measurement performance. The nonidealities can be because of offsets and mismatches in the high-speed data converters, especially that of the ADCs present in the oscilloscope. Though the signal generator is capable of generating modulated signals at an output power of 30 dBm, it starts to show palpable amount of nonlinearity at a power above 10 dBm. This can be observed as the intermodulation distortion in the input signal itself.

Note that, in this section, we will provide figures to illustrate each of the signal processing steps that were carried out to obtain the PA model. For the sake of avoiding redundancy, we provide figures for the case of 80 MHz input signal. Until unless mentioned explicitly, the figures correspond for the case of 80 MHz bandwidth signal. Since most of the figures are in theory similar in both the cases.
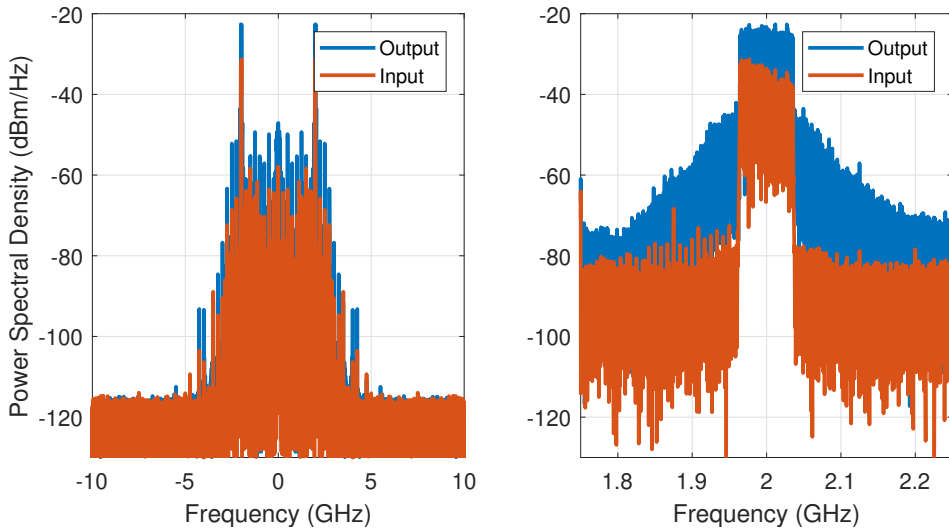
FIGURE 3.7: Spectra of the measured PA input and output RF data captured from the oscilloscope. Spectra on the left is for the total captured frequency, i.e., from -10 GHz to 10 GHz and on the right is the frequency-magnified data from 1.75 GHz to 2.25 GHz

In MATLAB, the captured PA input and output RF signals are digitally downconverted to baseband and then power aligned at 20 GSPS. The process of digital downconversion converts the real signal in RF to baseband complex signal composed of in-phase and quadrature signal components. The input signal samples were normalized based on the mean power of the downconverted input data. The inband signal power of the output downconverted signal is made equal to that of the downconverted input signal, known as power alignment. The inband power aligned downconverted spectra of input and output signals are shown in Fig. 3.8.

The digitally downconverted and power-aligned signals are then time-aligned using cross-correlation. Time alignment between input and output signals are necessary to compensate for the finite delay that the input signal undergoes through the PA to become
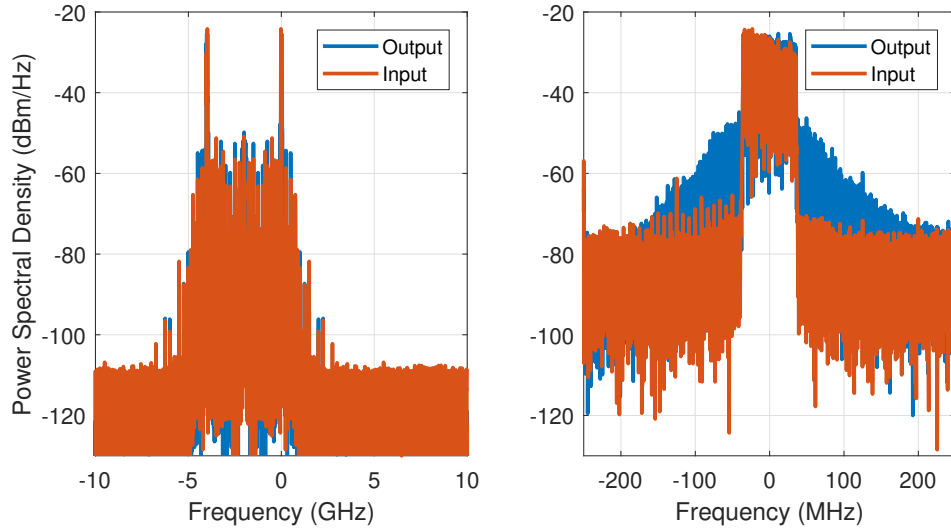
FIGURE 3.8: Spectra of the measured PA input and output downconverted and power-aligned data. Spectra on the left is for the total captured frequency, i.e., from $-10\,\text{GHz}$ to $10\,\text{GHz}$ and on the right is the frequency-magnified data from $-250\,\text{MHz}$ to $250\,\text{MHz}$

the output signal. We perform this at the oscilloscope sampling rate of $20\,\text{GSPS}$ to obtain very fine delay adjustment between the input and output sampled signal data, the time-alignment error is lower than the sample period of the oscilloscope, which is $50\,\text{ps}$. Fig. 3.9 shows the output of the cross-correlation that was performed between the input and output signals. The PA output lags by $4.3\,\text{ns}$, which corresponds to 86 samples.
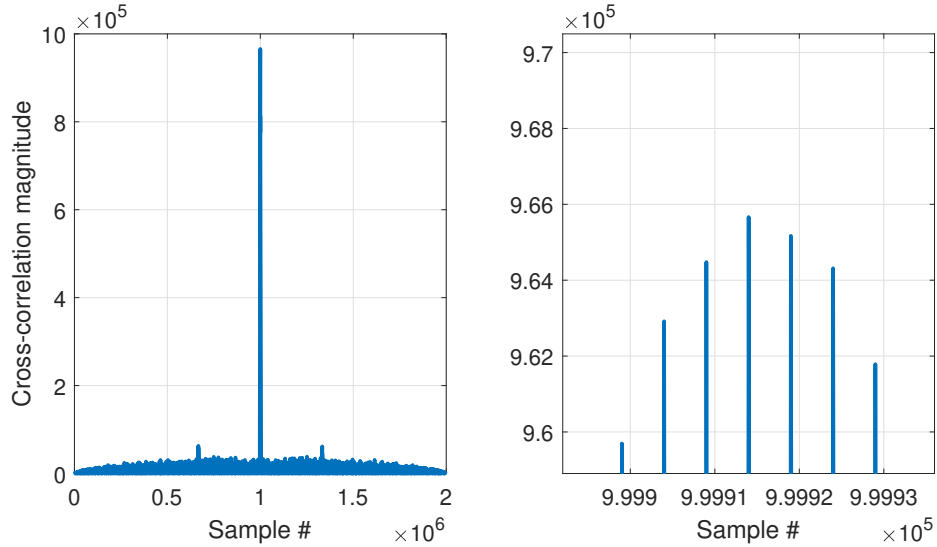
FIGURE 3.9: Cross-correlation output plot. On the left is for the total correlation data samples, i.e., a sample less than two million samples (1999999) and on the right is the data obtained by magnifying around the center peak, from 999892 to 999936 cross-correlation output samples

The decimation of the time-aligned high sample rate data is then performed to obtain a sampling rate reduction, which after decimation is equal to $125\,\text{MHz}$ and $500\,\text{MHz}$,
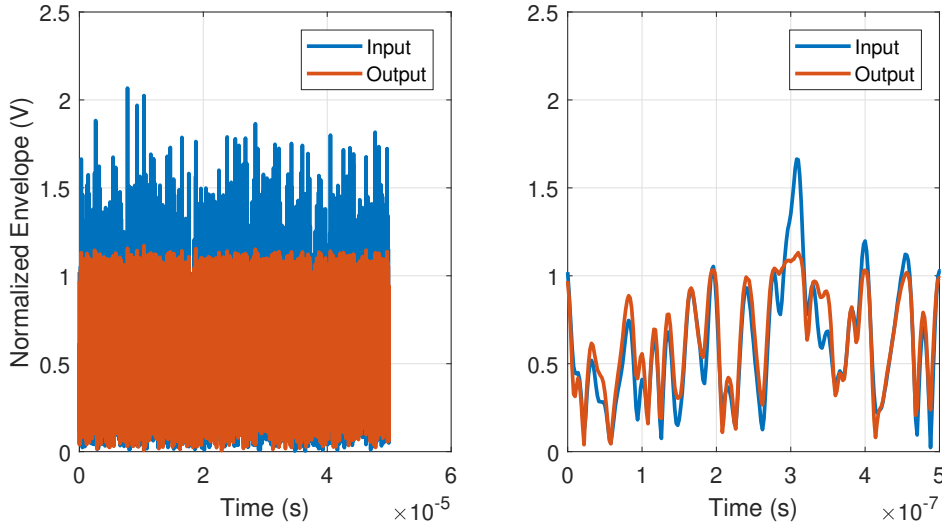
FIGURE 3.10: PA input and output baseband data envelope plots. Plots on the left is for the total captured duration, i.e., $50\,\mu$s and on the right is the time-magnified data for $50\,$ns duration
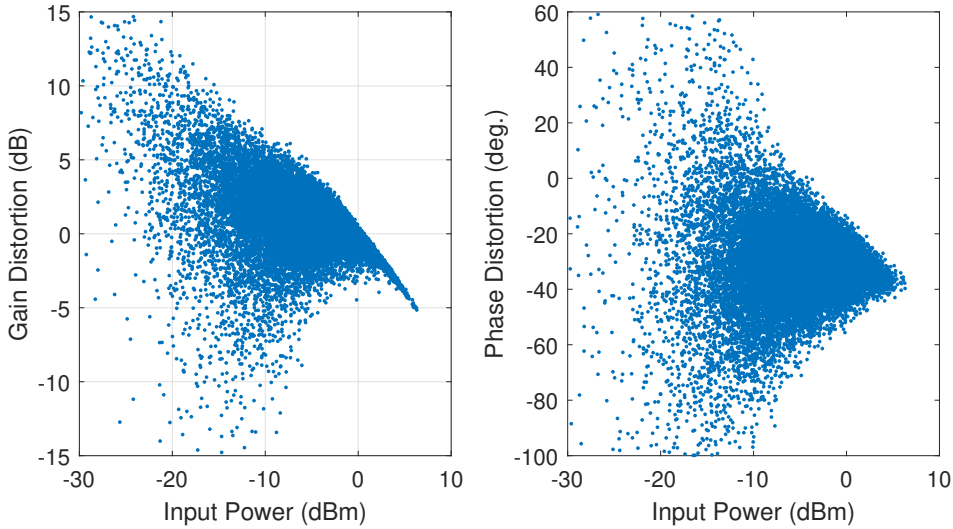
FIGURE 3.11: AM/AM (left) and AM/PM plots (right) of the baseband signal data

respectively, for 20 MHz and 80 MHz signals. The baseband sample rate is chosen such that it is more than five times the input signal bandwidth and hence, can capture all the intermodulation distortion components of the PA output. We obtain 6250 and 25000 baseband samples for 20 MHz and 80 MHz signals, respectively, at 125 MHz and 500 MHz respective sample rates. The signals are now ready to be used for PA modeling. Fig. 3.10 shows the input and output baseband signal envelopes obtained after decimation. The AM/AM and AM/PM plots of the baseband data are shown in Fig. 3.11, where gain and phase distortion of the PA are respectively depicted. It can be observed that the PA gain compresses by $-5\,$dB, showing a huge amount of nonlinearity, this is also clearly evident

in the Fig. 3.10, where the PA normalized input envelope peaks beyond $2\,\text{V}$ while that of the output saturates around $1\,\text{V}$. Also, we can observe very high amount of dispersion in AM/AM and AM/PM plots, proving that the PA suffers from memory effect.
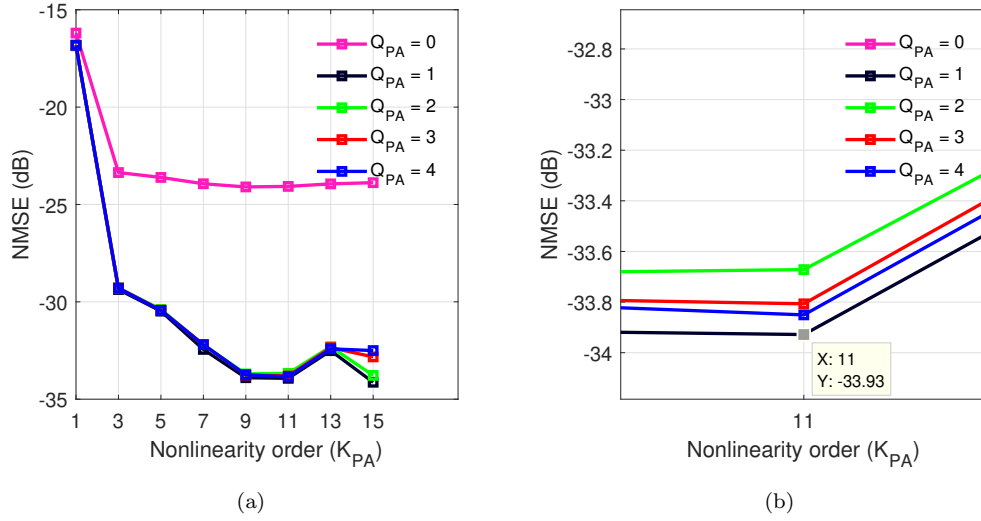
FIGURE 3.12: NMSE (dB) vs. K_{PA}, for different Q_{PA} for the 20 MHz bandwidth signal (a) K_{PA} from 1 to 15, Q_{PA} from 0 to 4 and (b) magnified around $K_{PA} = 11$
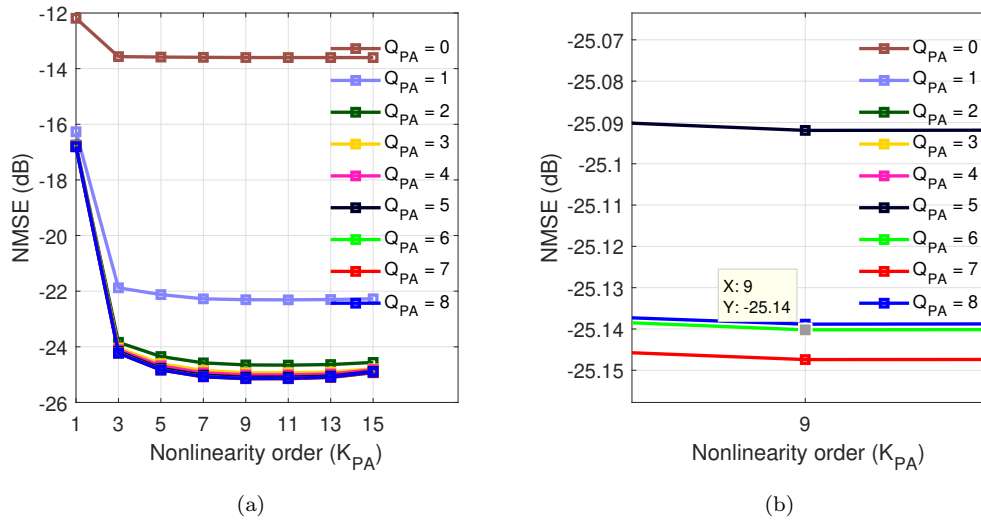
FIGURE 3.13: NMSE (dB) vs. K_{PA}, for different Q_{PA} for the 80 MHz bandwidth signal (a) K_{PA} from 1 to 15, Q_{PA} from 0 to 8 and (b) magnified around $K_{PA} = 9$

The behavioral model of the PA is extracted using the input and output decimated baseband complex data. Memory polynomial model has been chosen as the PA model candidate. The choice is motivated by three major reasons, it is widely used to model the PA in the literature, even in the recent studies [72, 73], it is easy to model with less number of coefficients and hence less computational complexity and finally, it attains a good modeling accuracy.

With the help of iterative sweeps, the nonlinearity order (K_{PA}) and the memory depth (Q_{PA}) at the input average power level of 6 dB are assessed to be $K_{PA} = 11$ and $Q_{PA} = 1$, $K_{PA} = 9$ and $Q_{PA} = 6$, respectively, for the case of 20 MHz and 80 MHz signals. The modeling accuracy is measured using Normalized Mean Square Error (NMSE) criteria. We have used 1500 and 6000 points out of 6250 and 25000 baseband samples for the PA model's coefficient identification, 20 MHz and 80 MHz signal cases, respectively. The rest of the points are used to evaluate the PA model. The NMSEs obtained are -33.93 dB and -25.14 dB, respectively, for 20 MHz and 80 MHz signals, as depicted in Fig. 3.13 and Fig. 3.14. The choice of K_{PA} and Q_{PA} was mainly based on least possible NMSE attainable and we have also considered the model complexity, which increases with the increasing number of coefficients. We need $(Q_{PA} + 1) * (K_{PA} + 1)/2$ number of coefficients for the memory polynomial model, i.e., 12 and 35 for 20 MHz and 80 MHz signals. Note that we have used only odd-order nonlinearity for the PA model, and hence half the number of coefficients in comparison with that of the MP model having both even and odd order coefficients.
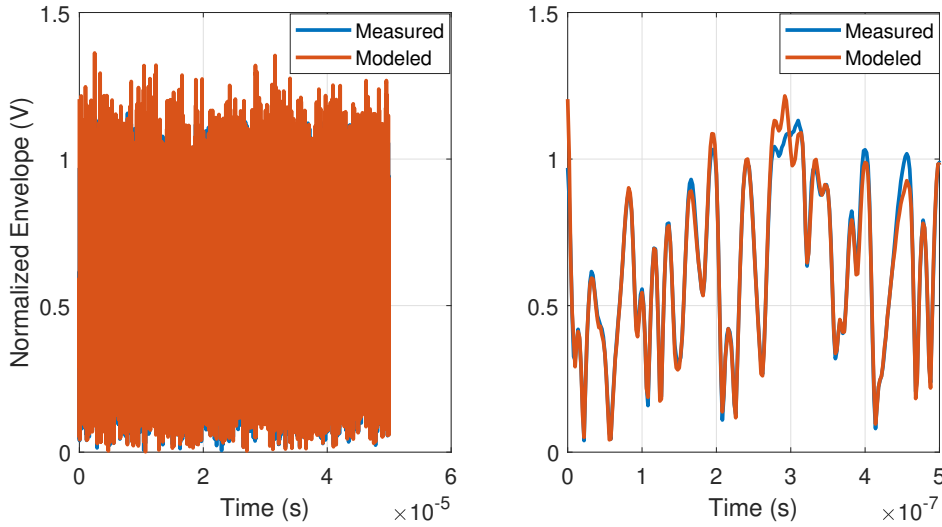
FIGURE 3.14: Plots of the measured and modeled PA output envelope. Plots on the left is for the total signal measurement duration of 50 μs and on the right is the time-magnified data for 50 ns duration

Fig. 3.14 shows the measured and modeled envelope of the PA output signal. It can be observed on the time-magnified version of the plot, shown on the right for the first 50 ns, that at high peak powers the MP model output is higher than that of the actual measured data.

The spectra of the measured and modeled output is plotted as shown in Fig. 3.15. The spectra on the left is obtained with a single spectrum on the 25000 samples each for both measured and modeled data and the spectra on the right is obtained by averaging over 20 spectra of 1250 samples each, using the same data of total number of 25000 samples.
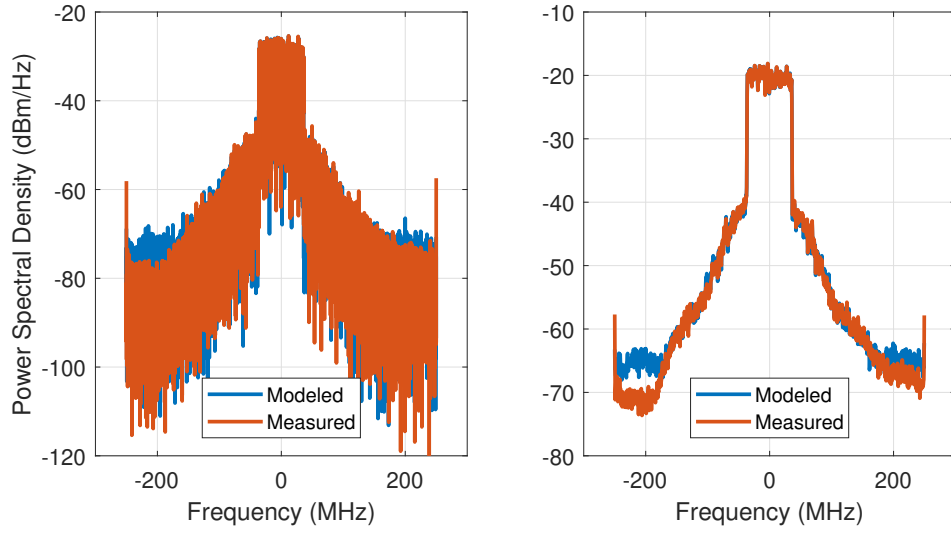
FIGURE 3.15: Spectra of the measured and modeled PA output. Spectra on the left is obtained with a single spectrum for both measured and modeled data and the spectra on the right is obtained by averaging
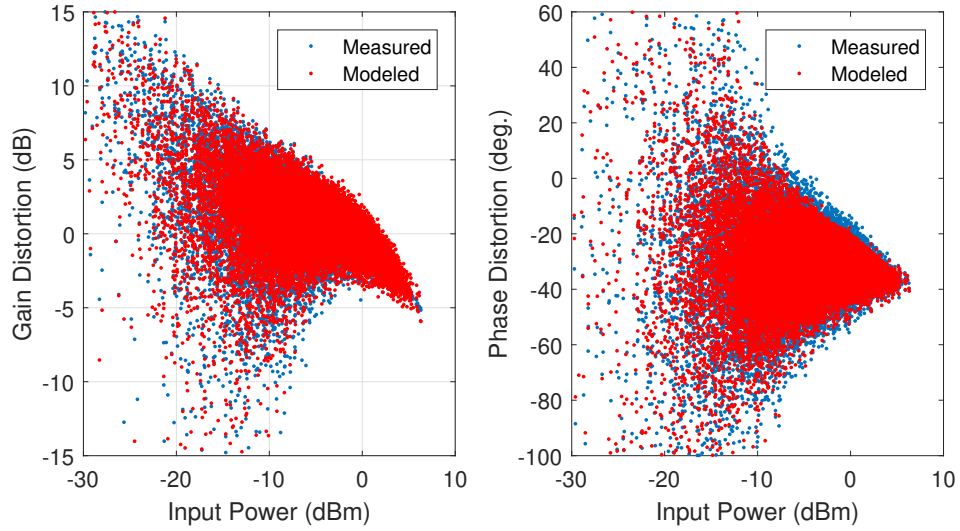
FIGURE 3.16: AM/AM (left) and AM/PM plots (right) of the measured and modeled PA output signal data

Spectral averaging reduces the amount of uncorrelated white noise that obscures the visual perception of the spectra. The spectra on the right shows that the spectra obtained from the modeled output coincides well with that of the measured output signal spectra, showing the visual validity of the model.

Fig. 3.16 shows the AM/AM and AM/PM plots of the measured and modeled PA output signal data similar to the one shown in Fig. 3.11, where gain and phase distortion of the PA are respectively depicted.

The modeling accuracy can be improved if the power of the input signal is reduced, which reduces the static and dynamic PA nonlinearities or by using better models such as Volterra, which results in high computational complexity which is unnecessary and impractical especially in the context of small-cell base station applications. It will be shown in the following section, the FIR-MP predistorter coefficient extraction algorithm requires PA model to be learned as the first step. This is opposed to the indirect learning architectures, where the predistorter coefficients are learned without the necessity of learning the PA model.

3.4   FIR-MP Coefficient Identification Methodology

In this section, we illustrate the method used to identify the coefficients of the FIR-MP predistorter. As explained in Section 2.3.2, Memory polynomial model is a linear-in-parameter model, i.e., there exists a linear relationship between the coefficients and the output signal. And hence, the coefficients can be extracted using any linear training method. For example, least square estimation is used in [42]. But for the case of FIR-MP predistorter model, it can be seen from (3.4) that this is no more the case; and hence the parameters have to be deduced using a nonlinear estimation algorithm, such as Newton iterative method, as used in for example [94]. The usage of iterative methods to estimate parameters of nonlinear systems is often discouraged since they are resource-intensive and hence require more processing power to converge in a reasonable amount of time. Hence, we propose to use another simple non-iterative method based on
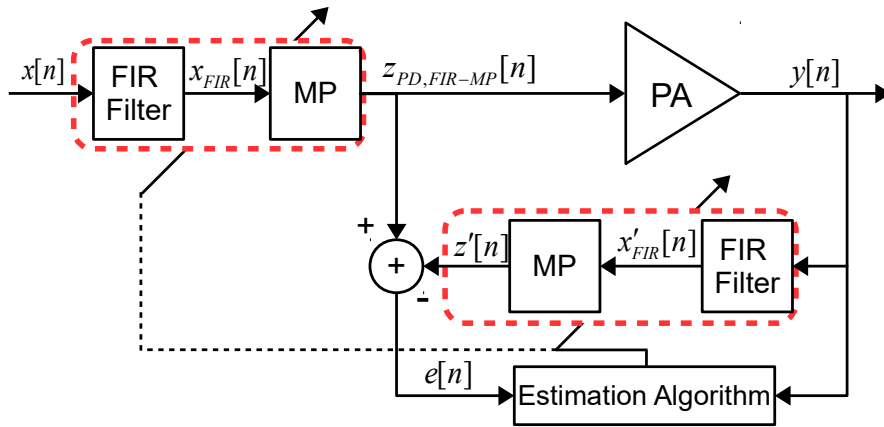
FIGURE 3.17: Illustration of the coefficient estimation methodology for FIR-MP DPD

learning approach, using small-signal-assisted parameter identification (SSAPI) algorithm as proposed in [10]. The SSAPI algorithm was used in the context of determining the coefficients of the FIR-EMP model. The SSAPI algorithm uses a two-step approach to estimate the parameters of both the FIR filter and the envelope memory polynomial

one after another, using least-squares estimation method. Here instead of EMP we use MP in the case of FIR-MP model. As explained in [10], the basic idea is to use small-signal training data as the input to the PA, in order to excite the PA's linear memory distortion while static nonlinearity and nonlinear memory effects are avoided, and hence facilitating the learning of FIR filter coefficients. The small-signal training data cannot be directly obtained from the measurements, since it requires the PA to be backed-off significantly changing the PA's linear memory behavior altogether. This approach requires very sensitive observation path with high dynamic range. Also, this method is inherently an offline-training method and hence unsuitable for the real-time on-field practical applications. Hence, the small-signal training data is derived with the help of PA's behavioral model in the software. As explained in Section 3.3 we have used MP model as the PA model. The small-signal input used for training is a linearly scaled version of the original input training data, we resorted to a linear scaling factor of 100. The learning methodology is illustrated in the Fig. 3.17, where the inverse model of the PA is first obtained during the training and it is then used as the predistorter. The process of obtaining the inverse model is as summarized below:

- Firstly, the behavioral model of the PA is estimated using the measurement input and output data of the PA, which are $x[n]$ and $y[n]$, respectively. The measurement data is obtained in the absence of the predistorter.

- Then the PA model is fed with the small-signal training data $x_{ss}[n]$, and the output $y_{ss}[n]$ is obtained as depicted in Fig. 3.18.

- Based on $x_{ss}[n]$ and $y_{ss}[n]$, the coefficients of the FIR filter that minimizes the linear distortions are derived by using the least square estimation method.

- As shown in Fig. 3.19 with the help of large signal PA input $z_{PD,FIR-MP}[n]$, which initially in the absence of predistorter is $x[n]$ and the FIR filter output data $x'_{FIR}[n]$, which is obtained when the filter is excited with the PA large signal output $y[n]$, the MP model coefficients can be derived by using least square estimation method. Instead of measurement data, we have used the large signal data generated by the PA model for the fed input large signal modulated data as part of simulations, which will be explained in the following section and in the later part of the thesis.

3.5 Simulation Results, Optimal Dimensioning of DPD and DAC

The baseband input signals used for all the later simulations are:
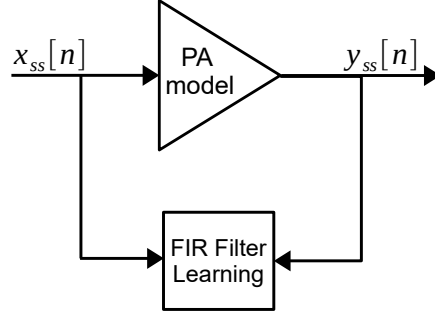
FIGURE 3.18: Illustration of FIR filter coefficients learning
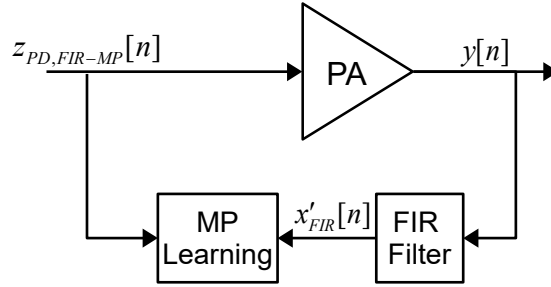
FIGURE 3.19: Illustration of MP block coefficients learning

- 4-carrier WCDMA signal with a total bandwidth of 20 MHz and PAPR of 8.4 dB;

- 4-carrier signal with a total bandwidth of 80 MHz and PAPR of 8.5 dB. This signal is obtained by increasing the WCDMA signal bandwidth by four times, in order to check the PA and predistorter performance.

Though these signals are different from that of the signals used to perform the PA model extraction, the model should be suitable to be used as the PA model for the two signal bandwidth cases. It is because the aforementioned two signals have similar PAPR and bandwidths to that of two LTE signals used for the PA model extraction.

As detailed in Section 3.4, the coefficients of the MP and FIR-MP predistorters can be derived using LSE and SSAPI algorithms, respectively. It is necessary to determine the nonlinearity order (K) and the memory depth (Q) of the MP block of the FIR-MP DPD as well as for the MP DPD. This can be iteratively chosen based on the ACLR metrics or special algorithms, such as genetic algorithm can be used to reduce the number of iterations [101], we chose the former. Along with the nonlinearity order and memory depth, FIR-MP requires the order of the FIR filter (L) to be determined. Determining (L) also follows an iterative simulation approach, which is found equal to the memory depth of the PA (Q_{PA}).

Fig. 3.20 presents the simulation testbench used to evaluate the linearization performance of the MP and FIR-MP DPD. We have used the MP models of the PA, as derived in
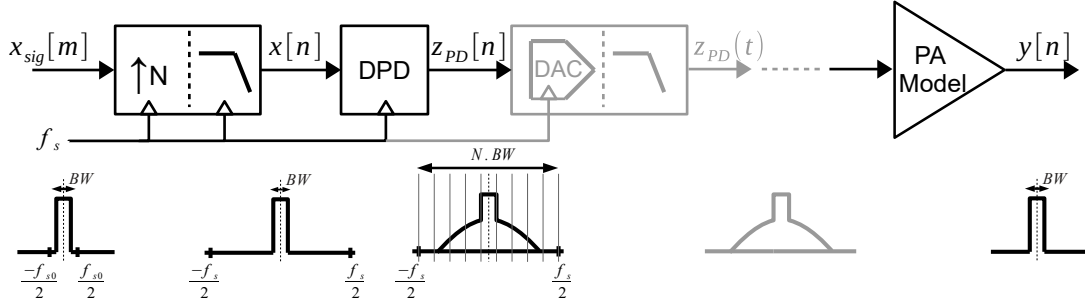
FIGURE 3.20: Simulation testbench for the DPD

Section. 3.3 for the 20 MHz and 80 MHz input signals. For both the models of DPD, there is also a necessity to determine the optimal sampling rate for f_s, which is obtained by interpolating the original baseband signal $x_{sig}[m]$ present at sample rate, as low as $f_{s0} = BW$. Generally, $f_s = N \times BW$, N must be at least five times that of the signal bandwidth for the predistorter to effectively cancel out 3rd-order and 5th-order intermodulation terms. The DPD which outputs complex baseband digital data is followed by a DAC and reconstruction filter in each of the in-phase and quadrature paths, as previously explained in Section. 1.1.3. The simulations of DPD are done in digital baseband, with the baseband PA model, and hence we don't have to consider the subsequent Tx components. But in the real implementations, the subsequent nonidealities of the entire Tx chain upto the PA poses DPD performance degradation.

We notice that the higher the sampling rate, the better the performance of the DPD, as can be seen by comparing Fig. 3.21(a) and Fig. 3.21(b), showing the ACLR1 vs. nonlinearity order (K) for different memory depths (Q), when the sampling rate is increased from five times (5X) to nine times (9X) for the 20 MHz input. In a similar way, Fig. 3.22(a) and Fig. 3.22(b) can be used to compare the ACLR1 when the sampling rate is increased from 5X to 9X for 80 MHz input signal case.

To give a hardware comparison, later in Section 3.6 we choose similar values of K and Q, and the DAC bandwidth for both MP and FIR-MP DPDs. For the case of 20 MHz, it is enough to use 5X sampling rate for the predistorter and DAC, with K=11 and Q=1, giving an ACLR1 of 52.4 dBc and 59.6 dBc, respectively, for MP and FIR-MP predistorters (from an ACLR1 of 36.2 dBc before predistortion). The value of L=1 was used for the FIR-MP. This meets the 3GPP ACLR specification [29] of 45 dBc with a margin. The improvement of using FIR-MP over MP is 7.2 dB. And for the case of 80 MHz signal, 9X sampling rate has been chosen so that the 3GPP ACLR specifications are met for both the MP and FIR-MP predistorter. With K=9 and Q=4, an ACLR1 of 45.6 dBc and 61.2 dBc, respectively are obtained for MP and FIR-MP predistorters from an ACLR1 of 34.3 dBc before predistortion. The value of L=6 was used for the FIR-MP. The improvement of using FIR-MP over MP is 15.6 dB. Note that with the 5X

(a)

(b)
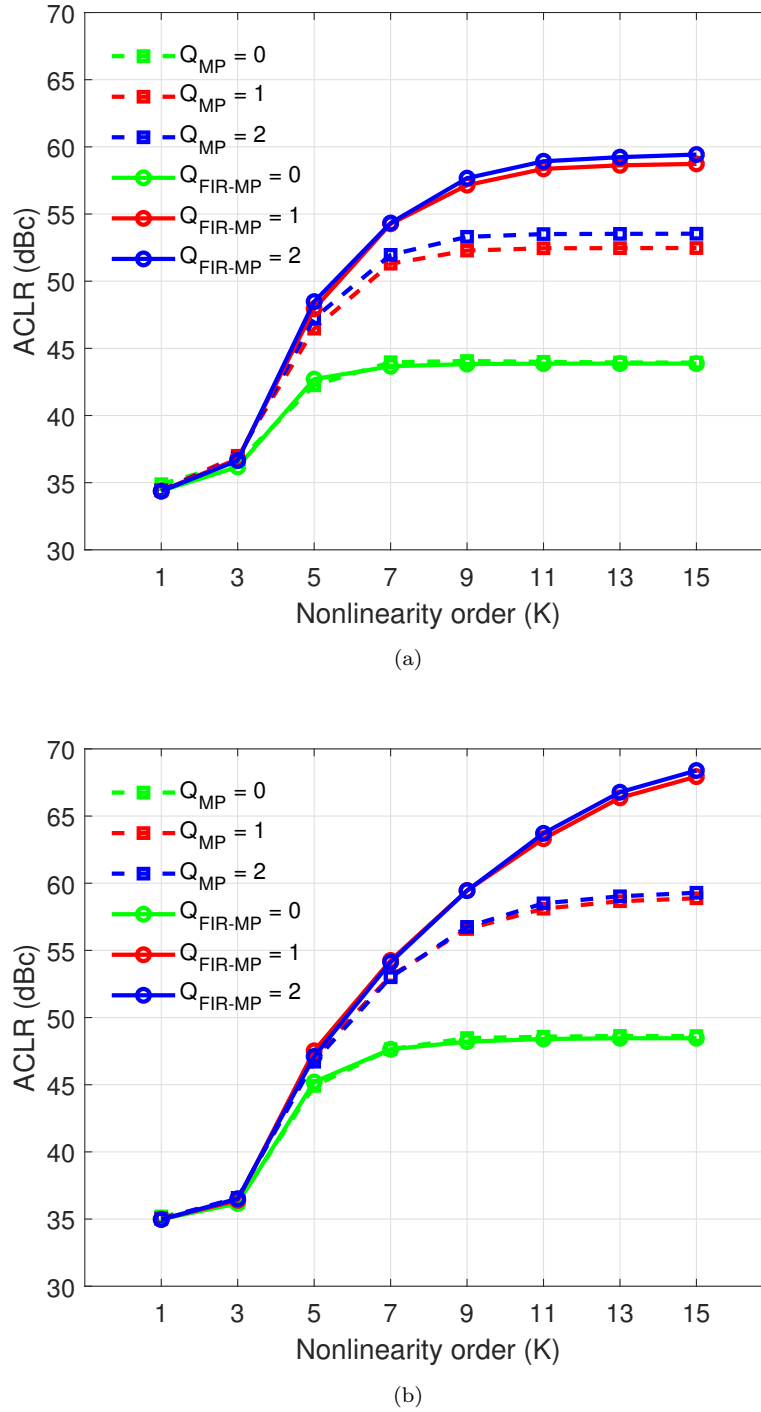
FIGURE 3.21: ACLR1 (dBc) vs. K, for different Q in the 20 MHz bandwidth signal case with MP and FIR-MP (L=1) DPDs, clocked at (a) 5X and (b) 9X signal bandwidth

sampling, FIR-MP obtains ACLR1 of 54 dBc, while MP obtains only 40.5 dBc and hence, fails to meet 3GPP specification. It is also worth noting that with Q=0, i.e., when the polynomials are memoryless, both the MP and FIR-MP behave similarly and cannot sufficiently linearize the PA, in spite of FIR filter being present in the case of FIR-MP DPD.

(a)

(b)
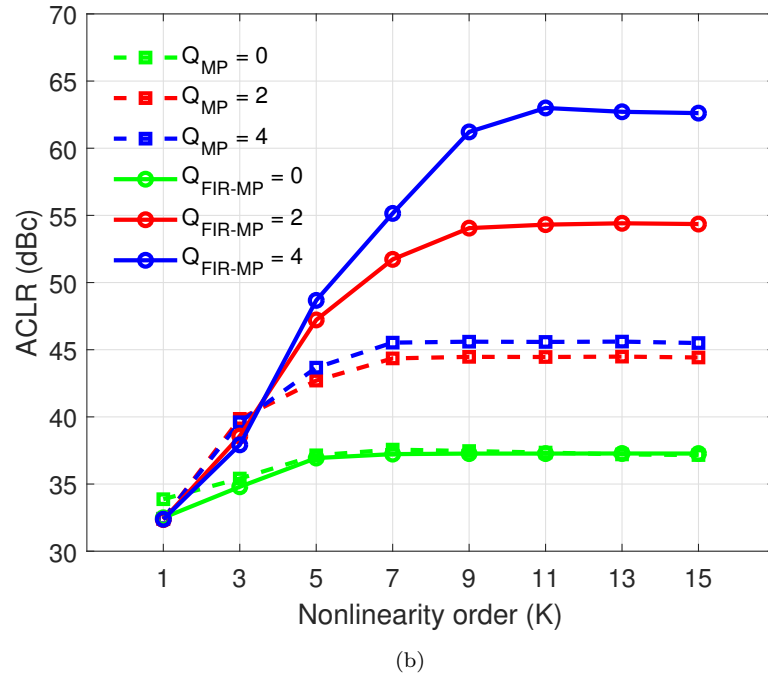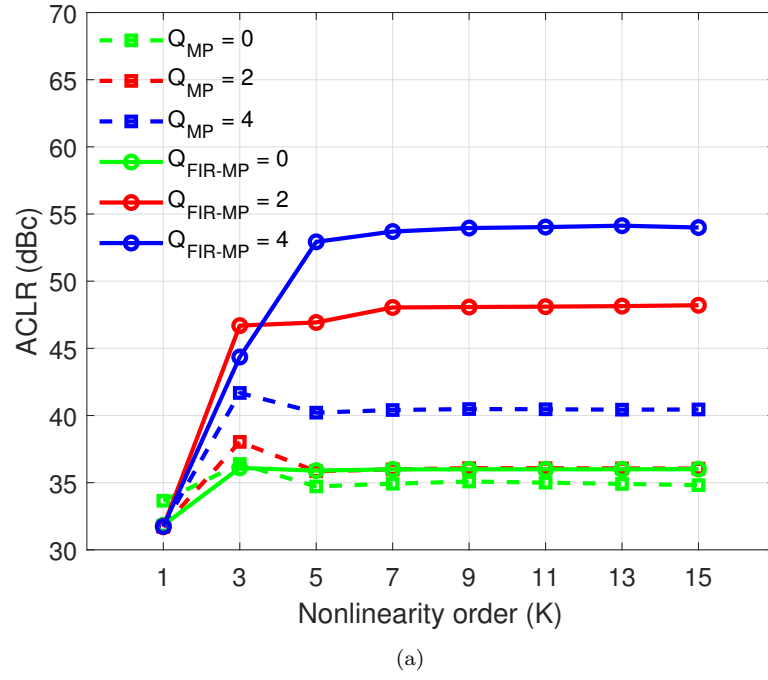
FIGURE 3.22: ACLR1 (dBc) vs. K, for different Q in the 80 MHz bandwidth signal case with MP and FIR-MP (L=6) DPDs, clocked at (a) 5X and (b) 9X signal bandwidth

Fig. 3.23 and Fig. 3.24 show the power spectra of the PA output before and after linearization using memory polynomial and FIR-MP predistorter, for the considered two input signal cases of 20 MHz and 80 MHz bandwidth, respectively. Table 3.1 encapsulates DPD linearization performance on the considered PA models driven by the two signal test cases, in terms of ACLR1 and ACLR2. The table also presents the comparison

of the nonlinearity order (K), memory depth (Q), FIR filter order (L) and the total number of coefficients used for the two predistorter models. The AM-AM plot of the 80 MHz signal depicted in Fig. 3.25 shows the improvement in distortion correction with FIR-MP DPD when compared to the conventional MP DPD. It can be inferred that the proposed FIR-MP DPD outperforms MP DPD in each of the signal cases.

TABLE 3.1: Performance summary of MP and FIR-MP DPDs

	K	Q	L	No. of coeff.	ACLR1 (dBc)	ACLR2 (dBc)
BW = 20 MHz						
Before	-	-	-	-	36.2	38.7
MP	11	1	-	12	52.4	58.7
FIR MP	11	1	1	14	59.6	61.4
BW = 80 MHz						
Before	-	-	-	-	34.3	36
MP	9	4	-	25	45.6	48.7
FIR MP	9	4	6	32	61.2	58.4
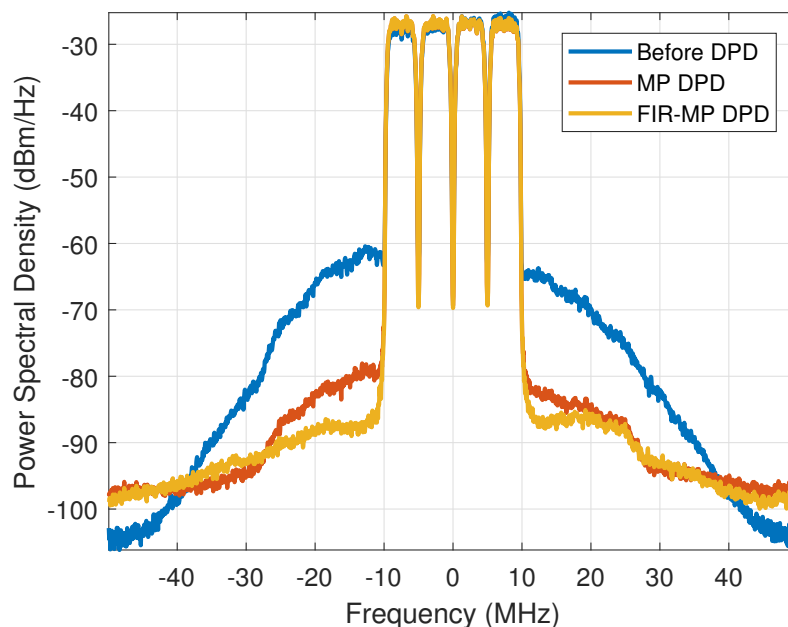
FIGURE 3.23: Power spectra of the output before and after linearization for a 4 carrier WCDMA signal with a total bandwidth of 20 MHz and PAPR of 8.4 dB. ACLR1 improvement of 7.2 dB is obtained using FIR-MP.

The increase in sample rate, which results in improved linearization can be made possible only if there exists DACs that finally convert the DPD's I and Q channel outputs, for the case of zero-IF transmitter architecture, as we have considered here. This also results in increased power consumption of the DACs and hence the overall cost of linearization. In
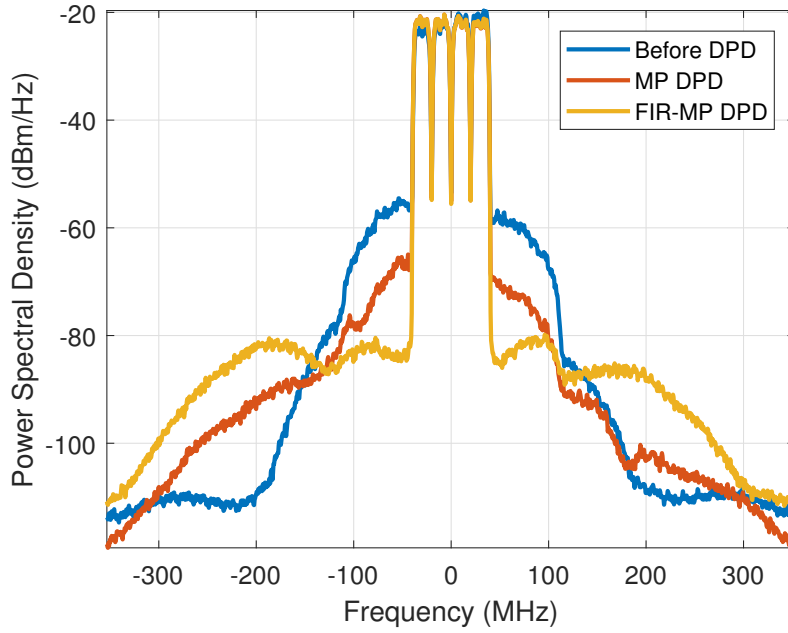
FIGURE 3.24: Power spectra of the output before and after linearization for a 4 carrier modulated signal with a total bandwidth of 80 MHz and PAPR of 8.5 dB. ACLR1 improvement of 15.6 dB is obtained using FIR-MP.
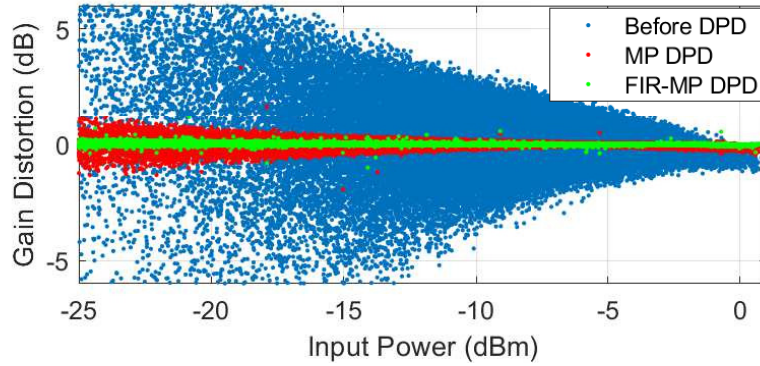
FIGURE 3.25: AM/AM plot showing the gain distortion, without and with the different predistorters for the 80 MHz input signal.

the context of low-power small-cell base stations, the selection of DPD as well as DAC bandwidth is an important step, as it dictates the overall power consumption and the performance of the DPD. Based on the linearization performance, we would like to use 5X and 9X clocking, respectively for the two signal cases of 20 MHz and 80 MHz. Hence, for the 80 MHz bandwidth case, the DACs should be clocked above 720 MHz. Note that we need two similar DACs for I and Q paths. In the case of real IF the DAC sample rate will be double, i.e., 1440 MHz, but only a single DAC will be needed [67]. Recent realizations of low-power, high bandwidth, and medium resolution DACs are available in the literature. For example, one of the low-power state-of-the-art DAC that might be suitable for the current small-cell base station application is presented in [102]. The

DAC is a 12-bit, 1.6 GSPS DAC, implemented in a 40 nm CMOS process, and achieves more than 70 dB SFDR and 50 dB SNR for signals over the 800 MHz Nyquist bandwidth. It has an active die area of $0.016 \, \text{mm}^2$ and consumes 40 mW.

3.6   Digital Implementation of the Predistorter

The proposed FIR-MP DPD differs form the classical MP DPD by the addition of an FIR filter. While it provides very good performance improvement in comparison with a similar order standalone MP DPD, the cost overhead of FIR filter needs to be quantified. Previously published works on the digital implementation of DPD has been mostly limited to field-programmable gate array (FPGA) based implementations [103, 104, 105]. Though FPGA DPD solutions are highly flexible, they also demand various commercial off-the-shelf components such as digital signal drivers, DACs and other precision clocking components, adding up to the increased power and cost. The cost and power consumption of using FPGA DPD solution in the context of wideband small-cell base station power amplifier linearization is prohibitively large, and hence unacceptable. Application specific integrated circuit (ASIC) implementation in deeply scaled CMOS technologies provides a low-power alternative. Though at the level of functionality both FPGA and ASIC are the same, but ASICs are inherently low-power and tend to be cheaper when mass produced, while FPGAs are highly programmable. Also, the DPDs can now be implemented as an integrated solution along with the baseband processor and other transmitter components, such as filters, DACs and mixers [46].

In this section, we explain the digital implementation methodology, the steps and the trade-offs necessary for low-power DPD implementation in 28 nm FDSOI CMOS process. We have used the HDL coder from MATLAB [106] to aid in the RTL code and testbench generation. The entire digital flow for the ASIC implementation is summarized in Appendix. B. The input to the HDL coder is the MATLAB code with persistent variables along with its testbench. An example code is shown in Appendix. C.1, where the FIR filter's MATLAB code is given for the case of 20 MHz.

3.6.1   DPD fixed-point implementation

Simulation tools such as MATLAB by default use double precision floating-point (64 bit), and hence produces the best possible result obtainable by DPD algorithms. But in the low-power implementations, be it FPGAs or ASICs, the first thing that has to be done is the conversion of the signal and all the DPD coefficients from floating-point to fixed-point representation. The fixed-point designer [107] is invoked by the HDL coder

TABLE 3.2: FIR-MP DPD performance summary for floating-point and various datapath wordlengths

	ACLR1 (dBc)	ACLR2 (dBc)
BW = 20 MHz		
No DPD	36.2	38.7
Floating point	59.6	61.4
WL = 16 bits	55	55.9
WL = 14 bits	54.9	55.8
WL = 12 bits	52.9	54
BW = 80 MHz		
No DPD	34.3	36
Floating point	61.2	58.4
WL = 16 bits	61.1	58.5
WL = 14 bits	60	58.6
WL = 12 bits	52.5	53.8

in MATLAB to accomplish the task of floating-point MATLAB code to its fixed-point counterpart. Appendix. C.2 shows the code generated by MATLAB fixed-point designer for the case of aforementioned 20 MHz FIR filter.

The cost of an ASIC implementation, i.e., the power consumption and area is directly proportional to the wordlength precision, i.e., the number of bits considered in the implementation datapath. We quantify the finite wordlength effects on the DPD performance using ACLR1 and ACLR2 metrics. Table 3.2 summarizes the ACLR performance obtained when the DPD is implemented in floating-point and fixed-point with resolutions of 16 bits, 14 bits and 12 bits for 20 MHz and 80 MHz signals. With 14 bits of wordlength, we can get ACLR performance similar to the floating-point implementation for 80 MHz signal, but with an ACLR1 degradation of 4.7 dB for 20 MHz. For 12 bits of fixed-point implementation, the ACLR1 starts to degrade considerably, by around 6.7 dB and 8.7 dB, respectively, for 20 MHz and 80 MHz signals.

Fig. 3.26 and Fig. 3.27 show the spectra depending on the chosen wordlength for 20 MHz and 80 MHz signals. Based on this quantitative analysis, we propose to implement the predistorter with 14 bit fixed-point implementation.
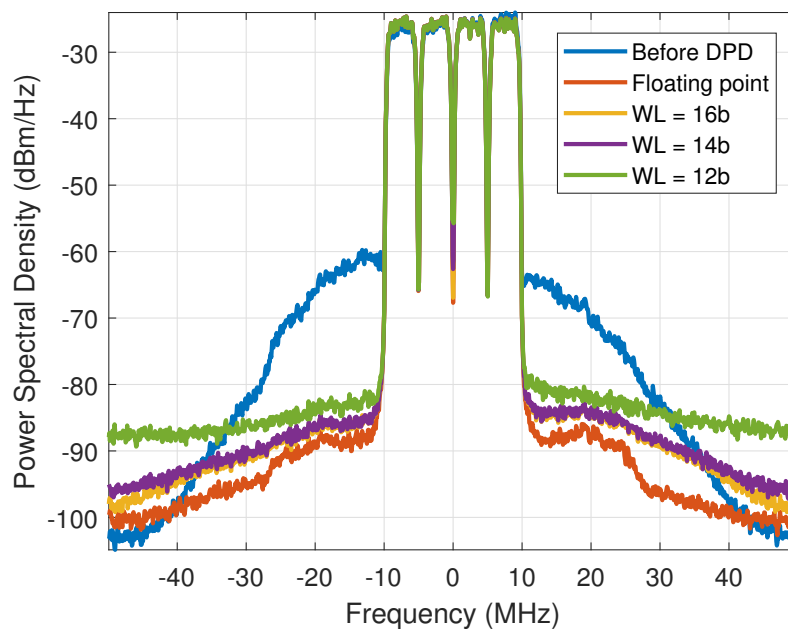
FIGURE 3.26: Spectra of the output signal without DPD and with floating-point and fixed-point representations of 16 bits, 14 bits and 12 bits for 20 MHz signal
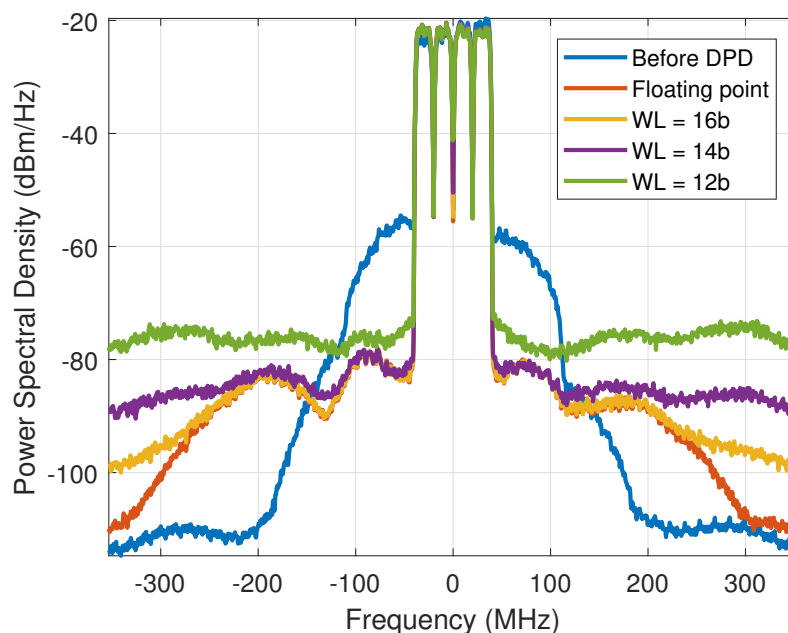
FIGURE 3.27: Spectra of the output signal without DPD and with floating-point and fixed-point representations of 16 bits, 14 bits and 12 bits for 80 MHz signal

3.6.2   Hardware synthesis

The DPD algorithms in fixed-point representation are then converted into verilog RTL code. The RTL code is used to synthesize the digital circuits, using standard cells in 28 nm FDSOI CMOS technology from STMicroelectronics. Pipelining and retiming are

performed to minimize the power consumption, and also to meet the timing constraints for FIR filter and memory polynomial sections for both the signal cases, except for the FIR filter of 20 MHz bandwidth signal. Four pipeline stages were inserted for the FIR filter section of 80 MHz system. Automatic placement and routing (PR) for the synthesized circuits are performed, along with the clock tree synthesis, to obtain the physical layout of the DPD system. The clock tree is used for the distribution of the clock signal to the sequential logic cells. The power of the circuits is estimated based on the post-layout simulations, using the corresponding modulated signals. The total power consumption of each of the FIR and MP section is estimated, which is given as:

$$P_{\text{total}} = P_{\text{leakge}} + P_{\text{dynamic}}, \tag{3.5}$$

where P_{total}, P_{leakge} and P_{dynamic}, respectively, are the total, leakage and dynamic power consumption. Similarly, the die area estimate is also obtained. Table 3.3 summarizes the FIR-MP synthesized metrics, namely the die area, number of standard cells used and the power consumption, for 20 MHz and 80 MHz bandwidth signals. The number of coefficients mentioned in the table are double that of the number mentioned in Table 3.1 since each complex coefficient has a real and an imaginary part. The overall power consumption of the FIR-MP DPD stands at 9.18 mW and 116.2 mW, respectively, for 20 MHz and 80 MHz signals, while the respective die areas are $39,374\,\mu\text{m}^2$ and $106,930\,\mu\text{m}^2$. It can be inferred that the cost of employing FIR filter is low in terms of power consumption, which is 1.276 mW and 30.55 mW, or 13.9% and 26.3% of the overall power, respectively, for 20 MHz and 80 MHz signal cases. It is because the FIR section compared to MP has few coefficients and only performs linear operations. Hence, the critical path is much smaller when compared to the memory polynomial section of the DPD.

For the 20 MHz signal bandwidth case the FIR-MP DPD solution in 28nm FDSOI CMOS consumes about 4.4X lower power in comparison with the state-of-the-art integrated solution of DPD (40 mW) [46] and 22X lower than ARFPD IC (200 mW) [8]. But in the latter case of ARFPD of [8] since the predistorter is present just before the PA, the whole Tx chain has to support just the signal bandwidth instead of at least 5X bandwidth. Hence, the total power consumption of the Tx chain when ARFPD solution of [8] is employed might be not be exactly 22X higher when compared to the total Tx power with our solution.

TABLE 3.3: Digital implementation summary for the FIR-MP DPD

	BW = 20 MHz		BW = 80 MHz	
	FIR filter	MP	FIR filter	MP
No. of Coeff.	4	24	14	50
Clock (MHz)	100	100	720	720
Die area (μm^2)	5,058	34,316	22,646	84,284
Std. Cells	3,916	24,456	15,921	61,076
P_{leakge} (mW)	0.065	0.37	0.27	0.92
$P_{dynamic}$ (mW)	1.211	7.53	30.28	84.75
P_{total} (mW)	1.276	7.9	30.55	85.67

3.7 Conclusions

Beginning with the discussion on the shortcomings of the memory polynomial predistorter while linearizing high bandwidth signals, this chapter has presented FIR-MP predistorter, which augments the linearization performance of the MP predistorter. The linear memory distortion that is dominant in the case when high bandwidth signals are applied to the PA is mitigated with an addition of an FIR filter before the memory polynomial predistorter. The simulations performed on the extracted models of a potential small-cell base station PA (ADL5606) showed that the proposed FIR-MP predistorter achieves better linearization when compared to MP predistorter, with slightly higher number of coefficients or similar performance achievement with less number of coefficients. DPD and DAC selection strategy was outlined based on systematic simulation based analysis.

A 28 nm FDSOI CMOS implementation of the MP and FIR-MP based digital predistorters has been performed. The digital implementation flow used to translate the algorithm to CMOS circuit has been presented, along with the optimal choice for internal wordlengths in the fixed-point representation. With a wordlength of 14 bits FIR-MP DPD obtains ACLR beyond 45 dBc, with a margin of 10 dB to meet 3GPP specification. Thanks to the advanced CMOS process technology, the synthesized FIR-MP DPD even with a higher nonlinearity order MP, clocked at high frequency, meets the specifications at very low cost. With an overall power consumption of 9.18 mW and 116.2 mW, respectively, for 20 MHz and 80 MHz signals, the FIR-MP DPD proves to be a suitable candidate for small-cell base station PA linearization. For the 20 MHz signal bandwidth case the FIR-MP DPD solution in 28nm FDSOI CMOS consumes about 4.4X lower power in comparison with the state-of-the-art integrated solution of DPD (40 mW) [46] and 22X lower than ARFPD IC (200 mW) [8].

When compared with the conventional memory polynomial with a fractional cost overhead, very large improvements in ACLR can be observed. Further power reduction of MP section can be possible if high performance is not needed, as outlined. The MP section might also be moved altogether into the analog baseband to further reduce the power, as will be explained in Chapter 4. Thus, the proposed FIR-MP predistorter finds suitability for wideband RF power amplifier linearization, especially in the context of future small-cell base stations.

Chapter 4

Mixed-Signal Predistorter System

In the previous chapter, we have presented the digital implementation of the proposed FIR-MP predistorter. We have shown the advantage it brings in terms of the ACLR improvement when compared with the conventional MP predistorter. Thanks to the CMOS technology scaling, Digital Predistortions (DPDs) can reach very high linearization performance with low-power. But the DPD implementation path, as well as all the subsequent components of the transmitter (Tx) chain, which consists of DAC, reconstruction/anti-imaging filters and mixers, now has to handle at least five times (5X) the signal bandwidth. This is needed in order to suppress the distortion components and hence increasing the total power consumption and complexity of the Tx chain by a large factor. This problem exacerbates with increasing signal bandwidths. On the other hand, existing Analog Radio Frequency Predistortion (ARFPD) system though simplifies the Tx chain, need power-hungry RF components which are challenging to design . As explained in Chapter 2, the RF section of ARFPD IC in [8] consumes 65% of the total power consumption, they include the components of the RF signal processor, which are a polyphase filter, envelope detector, multipliers, adder and a Variable Gain Amplifier (VGA). Also, the state-of-the-art ARFPD systems employ Envelope Memory Polynomial (EMP) model for predistortion and not the MP model. The reason for not using MP in ARFPD is because the implementation path demands RF delay elements and the same number of RF vector modulators, which are equal to the considered predistorter's memory depth. The RF components are very power-hungry and are tough to design for higher accuracy and high bandwidth requirements because of their sensitivity to process, voltage and temperature (PVT) variations, as explained in Chapter 2.

In this chapter, we propose to explore a new hybrid solution that can significantly reduce the hardware complexity and challenges involved in the design of existing ARFPD system by employing a mixed-signal predistorter (MSPD), which combines the advantages of both

digital and analog predistorters. The MSPD is based on FIR memory polynomial (FIR-MP) DPD presented in Chapter 3. By developing on the theory presented in previous chapters, this chapter starts with a brief comparison of various low-power predistorters, along with their linearization performance at system level. The architecture of the proposed mixed-signal predistorter is presented in Section 4.2. The simulation results and a brief analysis of the various non-idealities to derive the requirements of the circuit to be implemented using the proposed architecture is provided in Section 4.3. Potential architectures for the subsystems are discussed in Section 4.4. Section 4.5 concludes the chapter.

4.1   Predistorter Modeling and Performance Comparison

MP predistortion model was previously presented in Chapter 2 and Chapter 3. We recall the MP predistorter given by:

$$z_{PD,MP}[n] = \sum_{k=1}^{K} \sum_{m=0}^{M} a_{km} x[n-m] |x[n-m]|^{k-1} \tag{4.1}$$

where $x[n]$ is the input signal, a_{km} are the model coefficients, K is the nonlinearity order and M is the memory depth of the predistorter. Also, as introduced previously in Chapter 2, in the conventional memory-aware ARFPD systems the correction signal is generated in the analog baseband domain based on Envelope Memory Polynomial (EMP), whose digital equivalent is given by [92]:

$$z_{PD,EMP}[n] = x[n] \sum_{k=1}^{K} \sum_{m=0}^{M} a_{km} |x[n-m]|^{k-1}. \tag{4.2}$$

As can be seen from Eq. (4.2), the major advantage of EMP predistorter is that it only needs the current sample and just the magnitude or envelope information of current and past samples, according to the memory depth M.

The EMP based ARFPD has a major shortcoming of not being able to properly address linear memory effects of the PA. To improve the linearization performance of the EMP based ARFPD system, FIR-EMP was proposed in [10]. As described in Section 2.4.2, the digital baseband equivalent of the ARFPD based on FIR-EMP is given by:

$$\begin{aligned} z_{PD,FIR-EMP}[n] &= \sum_{l=0}^{L} h_l x[n-l] \\ &\times \sum_{k=1}^{K} \sum_{m=0}^{M} a_{km} \left| \sum_{l=0}^{L} h_l x[n-l-m] \right|^{k-1} \end{aligned} \tag{4.3}$$

where L is the FIR filter order and h_l are the filter coefficients, the rest of the variables being the same as mentioned in the former equations. It has been shown in [10] that the FIR-EMP outperforms MP model.

From Chapter 2, we can conclude that FIR-MP model predistorter outperforms the classical MP model. Similar to FIR-EMP, FIR-MP predistorter output is given by:

$$z_{PD,FIR-MP}[n] = \sum_{k=1}^{K} \sum_{m=0}^{M} a_{km} \sum_{l=0}^{L} h_l x[n-l-m] \times \left| \sum_{l=0}^{L} h_l x[n-l-m] \right|^{k-1}. \tag{4.4}$$

The linearization performance of the aforementioned four predistorters at algorithm level simulation are assessed using the MP PA model derived in Section 3.3. The simulation testbench is the same as previously illustrated in Fig. 3.20 and is depicted again for convenience in Fig. 4.1. The signal used is a 4-carrier modulated signal with a total bandwidth of 80 MHz and PAPR of 8.3 dB. Both the FIR-MP and FIR-EMP models use small-signal assisted parameter identification (SSAPI) algorithm to identify the coefficients, as explained in Section 3.4. We have considered the sampling rate equal to nine times (9X) the signal bandwidth, as described in Section 3.5. Fig. 4.2 shows the spectra and the Table 4.1 summarizes the performance in terms of adjacent and alternate channel leakage ratio, ACLR1 and ACLR2, respectively. Nonlinearity order K, memory depth M, FIR filter order L and the total number of coefficients needed are also presented in the table.
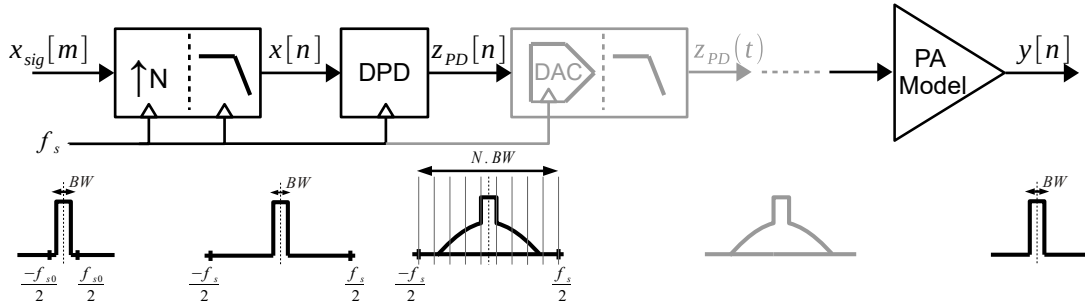
FIGURE 4.1: Simulation testbench for the DPD

Though the FIR-EMP performs better than the conventional MP predistorter, it can be clearly seen that FIR-MP model outperforms all the other three variants of predistorters including FIR-EMP. An improvement of 14.3 dB and 13.5 dB, in ACLR1 and ACLR2, respectively is obtained in comparison with FIR-EMP. This shows that the FIR-MP predistorter has high linearization performance for the considered scenario.
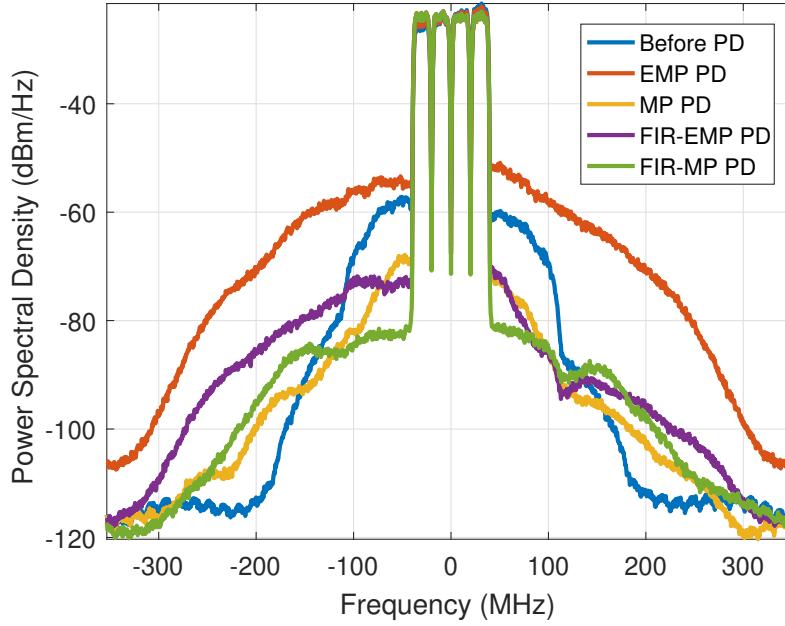
FIGURE 4.2: Power spectra of the output before and after linearization for a 4-carrier signal with a total bandwidth of 80 MHz and PAPR of 8.3 dB, using various predistorters

TABLE 4.1: Comparison of predistoter performance

	K	M	L	Number of coeff.	ACLR1 (dBc)	ACLR2 (dBc)
Before	--	--	--	--	34.5	36.1
MP	9	6	--	35	45.7	49.7
EMP	9	6	--	29	28.3	29.3
FIR-EMP	9	6	6	36	47.6	50.4
FIR-MP	9	6	6	42	61.9	63.9

4.2 Mixed-Signal Predistorter Architecture

The Tx chain with the proposed mixed-signal predistorter is shown in Fig. 4.3. Signal spectra at different stages of the Tx chain is also sketched in the figure. Based on the FIR-MP model, FIR filter is used in the digital baseband, while the MP is implemented in analog baseband domain, depicted as MP APD. FIR filter greatly improves the correction performance of the MP predistorter. When compared with a conventional ARFPD based on EMP, MP implementation provides higher linearization performance and also altogether eliminates the power-hungry RF section. MP APD can eliminate the usual 5X bandwidth overhead on the digital section, DAC and the reconstruction filter as required by a full DPD. Note that since the signal bandwidth in the digital section is only BW, and resultantly the FIR can be clocked at a rate equal to BW. But since
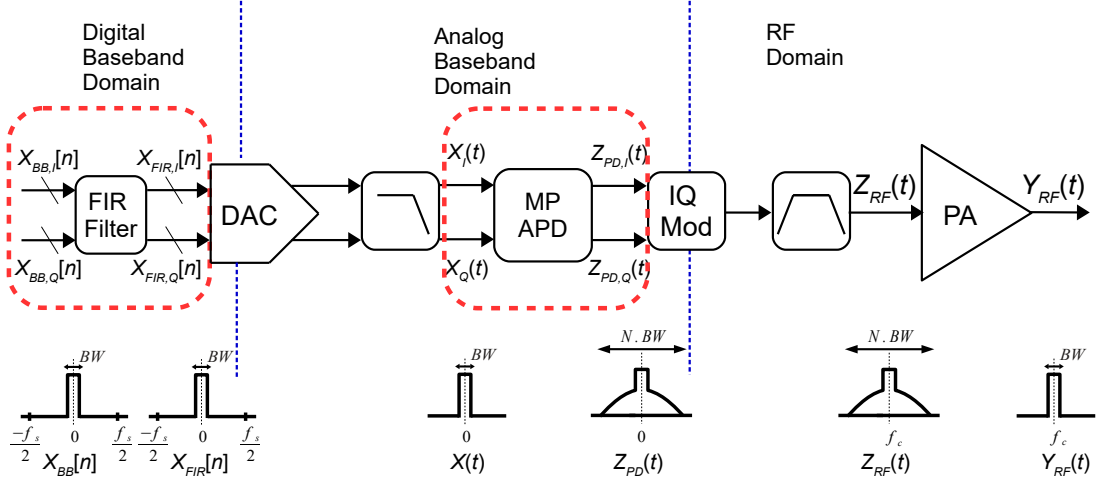
FIGURE 4.3: Transmitter with proposed mixed-signal predistorter (MSPD), along with the signal spectra at different stages

we did not use a multi-rate system for the FIR-MP, as explained in Section. 3.2.2, we have used the FIR filter clocked at $f_s = N \times BW$. Also, since SSAPI algorithm has been used to learn the parameters, we recall that from Section. 3.4, we learn the MP coefficients by sending large signal training data through the FIR filter. This requirement imposes high sample rate requirement on the FIR filter. We would like to address the $f_s = N \times BW$ sample rate requirement as part of future work, as will be presented in Chapter. 5. Though the signal from FIR filter $x_{FIR}[n]$ is clocked at $f_s = N \times BW$, the signal bandwidth is only BW. Hence we can use oversampling DACs in the in-phase and quadrature paths supporting only signal bandwidth, BW, but clocked at $f_s = N \times BW$. This greatly simplifies the analog reconstruction filter that follows the I and Q path DACs. As a matter of fact for an ideal system when $f_s = BW$, we should use brickwall filters with 3-dB cut-off frequency at $BW/2$.

Since the MP is in analog baseband, the complex output from the reconstruction filters given as

$$x(t) = x_I(t) + j x_Q(t), \tag{4.5}$$

where $x_I(t)$ is the in-phase and $x_Q(t)$ is the quadrature signals, which are outputs from the respective reconstruction filters. $x(t)$ is the input to the MP APD. The output of the MP APD is given by

$$z_{PD}(t) = z_{PD,I}(t) + j z_{PD,I}(t). \tag{4.6}$$

We would like to implement the MP described by Eq. 4.1 in analog baseband. We can rewrite Eq. 4.1 in analog domain as follows:

$$z_{PD,MP}(t) = \sum_{k=1}^{K} \sum_{m=0}^{M} a_{km} x(t - m t_d) |x(t - m t_d)|^{k-1} \qquad (4.7)$$

Rewriting Eq. 4.7 as

$$
\begin{aligned}
z_{PD,MP}(t) =& \sum_{k=1}^{K} a_{k0} x(t) |x(t)|^{k-1} \\
&+ \sum_{k=1}^{K} a_{k1} x(t - t_d) |x(t - t_d)|^{k-1} \\
&+ \dots \\
&+ \sum_{k=1}^{K} a_{kM} x(t - M t_d) |x(t - M t_d)|^{k-1}.
\end{aligned}
\qquad (4.8)
$$

Again rewriting Eq. 4.8 as follows:

$$
\begin{aligned}
z_{PD,MP}(t) =& a_{10} x(t) + a_{30} x(t) |x(t)|^2 + \dots + a_{K0} x(t) |x(t)|^{K-1} \\
&+ a_{11} x(t - t_d) + a_{31} x(t - t_d) |x(t - t_d)|^2 + \dots + a_{K1} x(t - t_d) |x(t - t_d)|^{K-1} \\
&+ \dots \\
&+ a_{1M} x(t - M t_d) + \dots + a_{KM} x(t - M t_d) |x(t - M t_d)|^{K-1}.
\end{aligned}
\qquad (4.9)
$$

Using Eq. 4.9 we can further go ahead and split each of the baseband complex signal
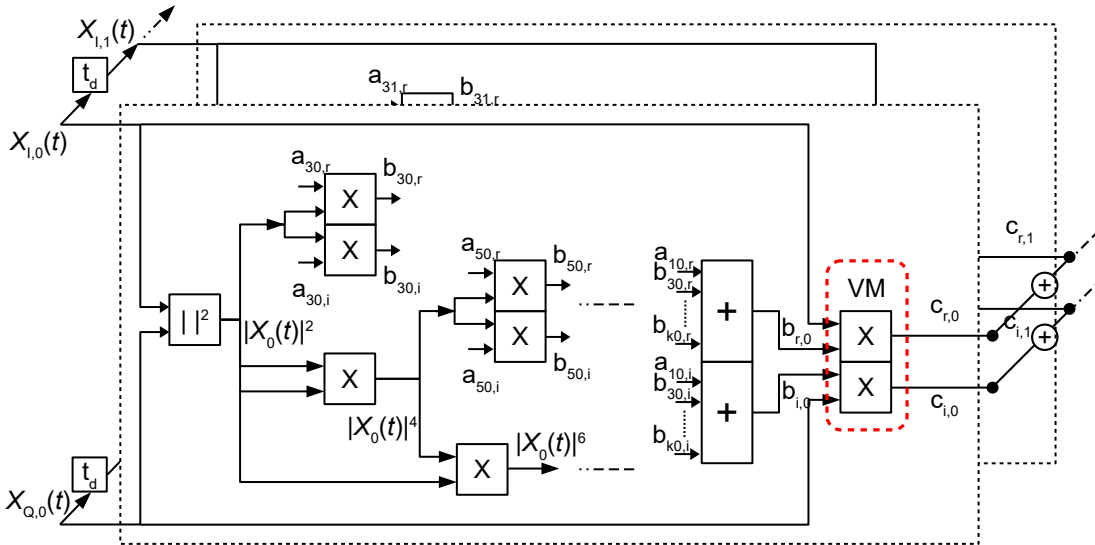
FIGURE 4.4: MP analog predistorter architecture implementation (VM: Vector Multiplier)

and also the complex coefficients into their constituent real and imaginary components. The implementation of the MP in the analog baseband is as shown in Fig. 4.4, obtained with the help of the aforementioned splitting up of complex signals and coefficients. The predistorter consists of memory kernels, depicted as dashed rectangular slices in black color. For the first kernel, the input in-phase and quadrature components are directly used. For the subsequent memory kernels, the appropriate delayed versions of the input generated using analog delay elements, which delay the respective signals by t_d are used. In each kernel, the in-phase and quadrature components of the kernel's input signal are used to generate the envelope power. The envelope power for the first memory kernel is given as:

$$|x_0(t)|^2 = x_{I,0}^2(t) + x_{Q,0}^2(t). \tag{4.10}$$

Two squarers (multipliers) and an adder accomplish this task. The envelope powers are used to generate their harmonics using squarers (multipliers).

The generated harmonics are multiplied with the real $(a_{km,r})$ and imaginary values $(a_{km,i})$ of the appropriate coefficients. The coefficients are generated by the DACs. The resulting real $(b_{km,r}(t))$ and imaginary $(b_{km,i}(t))$ values are added to obtain real and imaginary outputs, $b_{r,m}(t)$ and $b_{i,m}(t)$, respectively. For simplicity, t is omitted in Fig. 4.4. The complex-valued outputs $b_{r,m}(t)$ and $b_{i,m}(t)$, are now *vector multiplied* with the in-phase and quadrature signal components or their delayed versions, $x_{I,m}(t)$ and $x_{Q,m}(t)$, respectively, accordingly to obtain the in-phase and quadrature components of each memory kernel, $c_{r,m}(t)$ and $c_{i,m}(t)$, respectively, as given by:

$$c_{r,m}(t) = b_{r,m}(t) \cdot x_{I,m}(t) - b_{i,m}(t) \cdot x_{Q,m}(t), \tag{4.11}$$

$$c_{i,m}(t) = b_{i,m}(t) \cdot x_{I,m}(t) + b_{r,m}(t) \cdot x_{Q,m}(t). \tag{4.12}$$

The in-phase and quadrature outputs of all the memory kernels are added to obtain the predistorter's in-phase and quadrature components, respectively, as given by:

$$z_{PD,I}(t) = \sum_{m=0}^{M} c_{r,m}(t) \tag{4.13}$$

$$z_{PD,Q}(t) = \sum_{m=0}^{M} c_{i,m}(t). \tag{4.14}$$

4.3 Simulation with Major Non-idealities of APD

Since the MP part of the proposed MSPD system is in the analog baseband domain, the predistortion performance is governed by various non-idealities of the APD. The effect

of non-idealities should be studied in order to obtain the specifications of the various subsystems of the APD for an IC implementation.

The real and imaginary coefficients of the memory polynomial, $a_{km,r}$ and $a_{km,i}$, respectively can be generated with the help of Digital-to-Analog Converters (DACs). This is necessary in order to have the coefficients programmable. Time-delays can be implemented as all-pass filters, as will be explained in brief detail in Section. 4.4. Apart from delays and coefficient DACs, it can be observed from Fig. 4.4 that the signal path mainly constitutes of Multipliers and adders. The major non-idealities of the APD are produced by the aforementioned three subsystems and can be encapsulated as follows:

- Finite resolution of the coefficient DACs

- Presence of noise in the signal path, quantified by signal to noise ratio (SNR) value

- Time-delay, t_d mismatch

Each of the multiplier and adder stages are modeled such that they add uncorrelated white noise as the signal propagates through the predistorter. The noise floor of each element is with respect to the highest power component of the input signal, that is normalized to unity. This scheme is similar to the peak power normalization scheme used in [8]. The advantage we obtain is that the signal power never exceeds unity. The disadvantage is that the lower power components of signals tend to get smaller and smaller with each squaring operation, essentially burying them into the noise floor.

The performance of the predistorter is assessed with the non-idealities of limited signal-path SNR and finite coefficient DAC resolution. The simulations are performed in MATLAB using a 4-carrier modulated signal with a total bandwidth of 80 MHz and PAPR of 8.3 dB. This signal is used as the input to the predistorter to linearize the baseband MP PA model derived in Section 3.3. Fig. 4.5 shows the ACLR vs. coefficient DAC resolution, (N bits) for various values of the signal path SNR when the nonlinearity order, $K = 9$ and memory depth, $M = 6$ of the predistorter. The baseband FIR filter order, $L = 6$ is used. It can be seen that at least N of 6 bits and 60 dB of SNR are required in the signal path to achieve an ACLR value more than 45 dBc, to meet the 3GPP specification [29]. Though seven or eight bits of N is preferred in order to have some design margin.

In the context of small-cell base stations, where low-power linearization is an absolute requirement, the case of reduced nonlinearity order and memory depth can be used. From the simulation results presented in Fig. 3.22(b) we can observe that we can obtain ACLR value greater than 45 dBc even when $K = 5$ and $M = 2$. This reduces the number
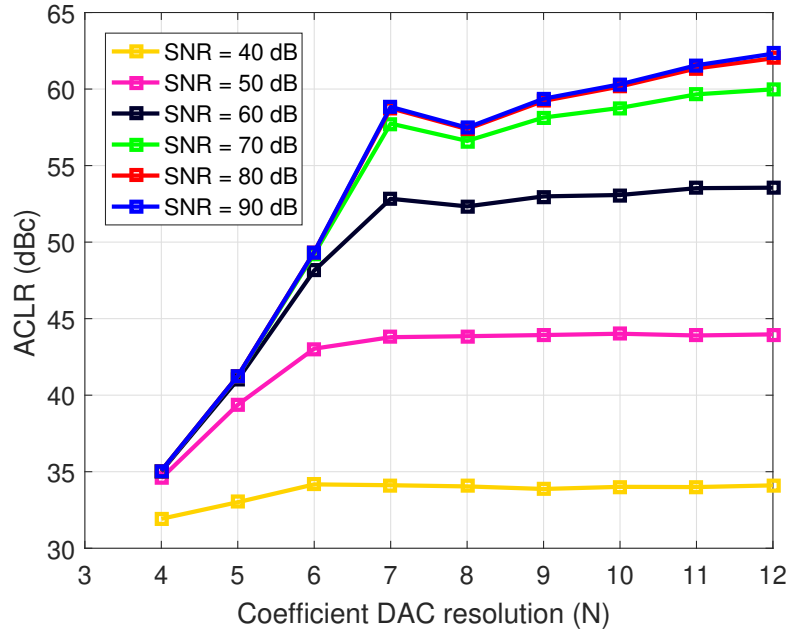
FIGURE 4.5: ACLR1 vs. coefficient DAC resolution, for various signal path SNR values, for the case of $K = 9$ and $M = 6$
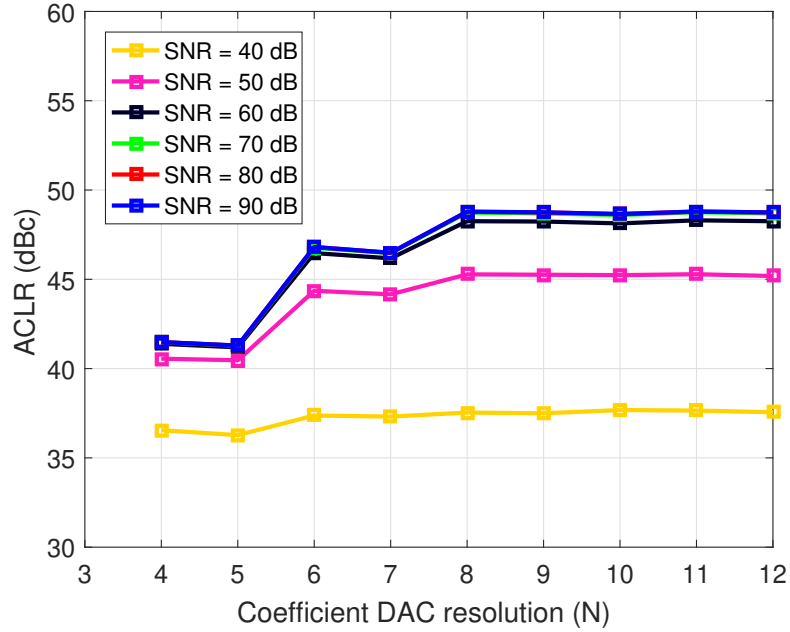
FIGURE 4.6: ACLR1 vs. coefficient DAC resolution, for various signal path SNR values, for the case of $K = 5$ and $M = 2$

of complex coefficients in the APD to just 9 ($\frac{(K+1)}{2} \times (M+1)$). The digital FIR filter order is kept unchanged, i.e., $L = 6$. As shown in Fig. 4.6, now with only 9 complex coefficients that have to be implemented in the analog domain, an ACLR value in excess of 45 dBc can be obtained. Again, a minimum of 60 dB of SNR and N of 6 bits are required. Fig. 4.7 shows the spectra before and after the predistortion with N of eight
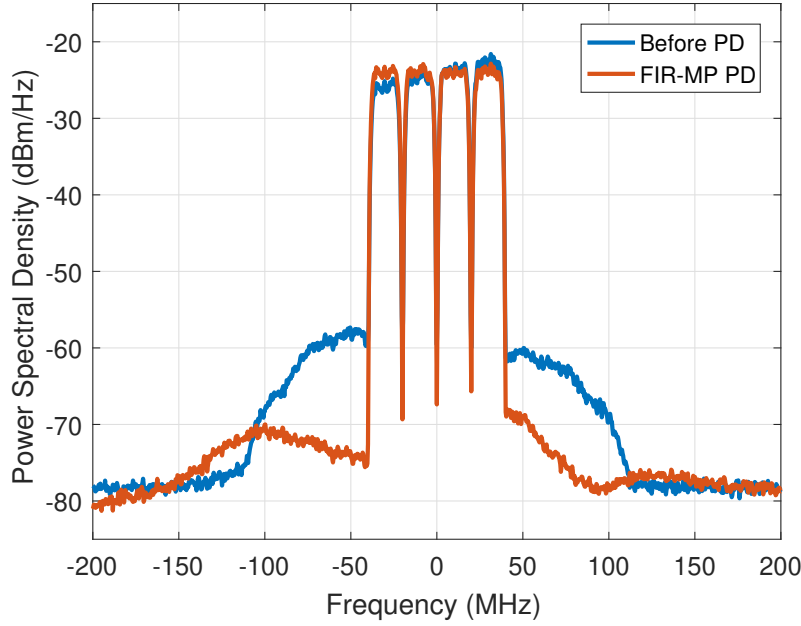
FIGURE 4.7: Power spectra of the output before and after linearization using non-ideal FIR-MP predistorter, with $K = 5$ and $M = 2$

bits. An improvement of ACLR1 from 34.6 dBc to 47.9 dBc and ACLR2 from 36.2 dBc to 48.4 dBc is obtained.

The effect of mismatch on analog time-delay, t_d is simulated using behavioral level Monte Carlo simulations for the predistorter with $K = 5$ and $M = 2$. The nominal value of t_d is about 1.4 ns for the case of input signal with 80 MHz of bandwidth, an amount equal to inverse of nine times the sampling frequency. We have performed multiple sets of Monte Carlo simulations and found that the predistorter can tolerate time delay mismatch of 5%. Fig. 4.8 shows the histogram of ACLR1 obtained from 1000 simulation runs with a 5% t_d mismatch, along with the aforementioned non-idealities. The value of 3-sigma obtained is 1.55 dBc, achieving a minimum of 46.1 dBc. The robustness is due to the less number of delay elements needed; two in each in-phase and quadrature path. In comparison, EMP based ARFPD of [8] has a memory depth of four. For addressing the PVT variations replica biasing circuits can be employed as in [8]. Hence, even with major non-idealities, the MSPD meets the 3GPP ACLR specifications along with margin.

4.4   Subsystems Architecture and Specifications

From the previous section, with the help of behavioral level electrical simulations, we have obtained a first-level approximation of the impact of major non-idealities. For the
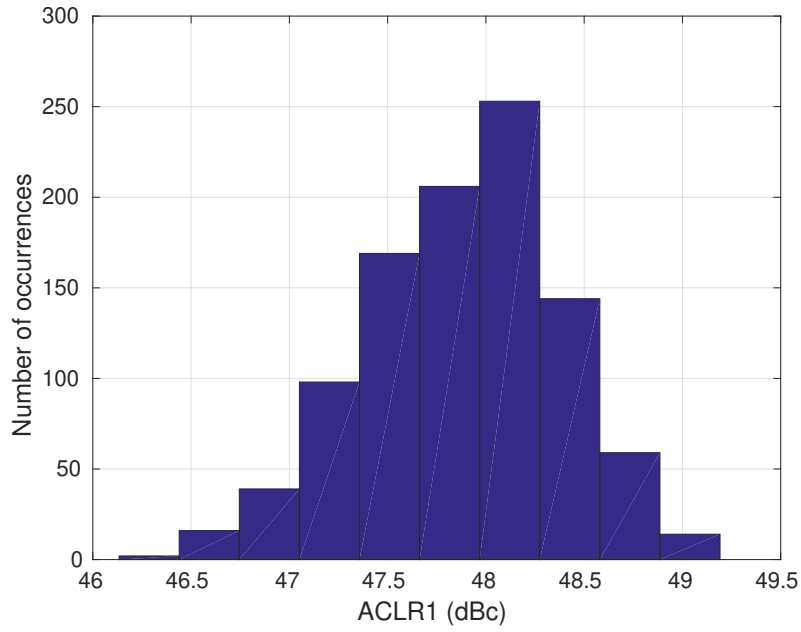
FIGURE 4.8: ACLR1 histogram obtained from 1000 runs of transient Monte Carlo simulations with 5% t_d mismatch

realization of the predistorter system in a cost-efficient form, design in CMOS process is a natural next step. Firstly, regarding the choice of technology, we would like to have the MSPD to be implemented in a modern CMOS process like 28 nm FDSoI, where the whole radio transceiver can be integrated with digital baseband processor to obtain low-cost small-cell BS solution. The salient features of this process is that the thick-oxide devices support a nominal voltage of 1.8V. Thick-oxide devices can be used as the transistor choice in order to maximize the signal swings necessary to obtain high SNR and linearity in the signal path.

This requires selection of the potential architectures which are available in the literature for the design of the subsystems. The three main subsystems that are required are:

- Coefficient DACs,

- Multipliers and

- Time delays.

Apart form the above three subsystems, summation blocks are also needed, which can be implemented in current mode by connecting the current sources and if necessary followed by a transimpedance amplifier as employed in [8]. In the following sections, we present some pointers to the potential candidate circuit architectures of the aforementioned three building blocks of the MSPD.

4.4.1   Coefficient DACs

Table 4.2 shows the predistorter coefficients with 64 bits double float and finite wordlength (8 bits) precision for the predistorter with reduced nonlinearity order and memory depth, i.e., $K = 5$ and $M = 2$. As pointed out in Section. 4.3, though 6 bits of resolution for the coefficient DACs were enough, we chose to use 8 bits for the coefficient DACs. Since, we would like to use the same DAC, 2 additional bits to support variations have been considered, though at this stage we did not yet analyze the variations in predistorter coefficients over PVT and aging of the PA.

TABLE 4.2: Predistorter coefficients of the simplified predistorter, with 64 bits double float and finite wordlength precision

Coeff.	64 bits double float	8-bit precision
a_{10}	1.058026709060108 + 0.022926907839086i	1.0546875 + 0.0234375i
a_{11}	-0.062283568758925 - 0.039968278829881i	-0.0625000 - 0.0390625i
a_{12}	0.029042404407057 + 0.026703667511766i	0.0312500 + 0.0234375i
a_{30}	0.075753170004712 - 0.125072957483959i	0.0781250 - 0.1250000i
a_{31}	-0.055432179665305 - 0.030447239911687i	-0.0546875 - 0.0312500i
a_{32}	-0.168011351079807 - 0.030308560985868i	-0.1718750 - 0.0312500i
a_{50}	0.674331940254317 + 0.252884324071456i	0.6718750 + 0.2500000i
a_{51}	-0.415135291277771 + 0.005189024624980i	-0.4140625 + 0.0078125i
a_{52}	0.286503200406367 - 0.090425278892664i	0.2890625 - 0.0937500i

There can be a large amount of variation between the same parameters when we would like to change the nonlinearity order and memory depth of the predistorter. It has been observed with the help of simulations. For example, when $K = 7$ and $M = 1$, the value of the coefficient a_{51} is 0.607568604930277 + 0.575913585510551i, which when converted to 8 bits becomes 0.6093750 + 0.578125i. The values demand that the coefficient DAC should be able to handle large variations, and hence it is imperative to consider margins for the DACs resolution.

Regarding the speed of the coefficient DACs, since the predistorter developed is not real-time adaptive, but is operated in open-loop, the update speed of the DAC can be lower. That is the coefficients only change when there is a variation in the PA characteristic requiring the predistorter coefficient update.

Numerous different DAC architectures are available in literature [108, 11, 109]. But for the aforementioned specifications one of the suitable candidates would be a charge redistribution DAC.

The charge-redistribution DAC works on the principle of switched-capacitor (SC) circuits. Fig. 4.9 shows a simple illustration of an N-bit binary-encoded charge-redistribution

DAC. With a total of 2^N number of capacitors scaled in a binary fashion converts the digital data to analog according to the N-bit digital data [109].
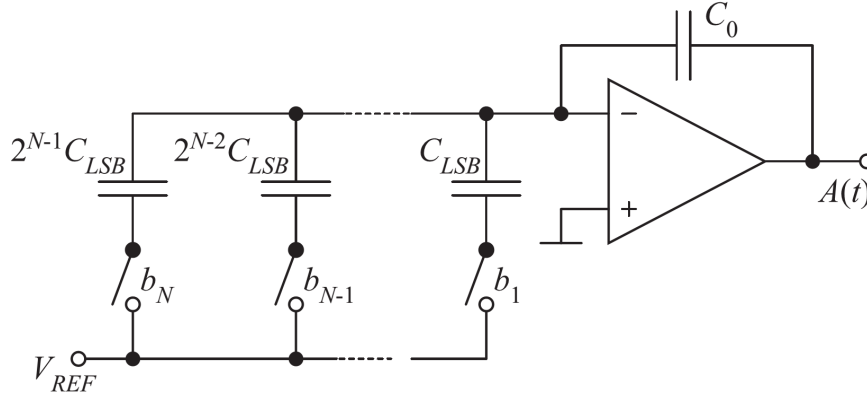
FIGURE 4.9: Illustration of an N-bit binary-encoded charge-redistribution DAC [11]

It can be noted here that the DACs that are used in [8] are current steering DACs, which combines the coefficient DAC and the multiplier. Since the predistorter in [8] is adaptive, it is a right choice albeit with higher power consumption. Also, the multiplier DAC is highly non-linear.

In order to have a faster start-up and have reduced power consumption of the DACs, we can initially have faster sampling clock and later reduce the clock frequency so that the switching activity is reduced and hence the dynamic power consumption.

Thermometric encoding is used to avoid glitches when compared to binary encoding in DACs [11]. But Thermometric comes with an increased cost of routing 2^N signals instead of N signals and hence increase in power and area. A clever combination of both known as segmentation can reduce glitches and with decreased area and power as shown in [110], where 12 bits of DAC resolution is attained by binary-encoded 8 LSBs and thermometer-encoded 4 MSBs. This DAC in [110] is employed in the context of a low-power near-Nyquist rate SAR ADC in 65 nm CMOS process. The LSB unit capacitor is 250 fF, custom-made with M6 and M7 metal layers. They are generally available in all advanced CMOS processes. The ADC achieves an ENOB of 10.1 bits with 97 nW of power with a sampling rate of 40 KSPS. The ADC power consumption drops to 1 nW when the sampling rate is reduced to 250 samples per second. The DAC consumes 25% of the total power. In our context a similar DAC for e.g., binary-encoded 4 LSBs and thermometer-encoded 4 MSBs can be used to attain 8 bit resolution.

4.4.2   Multipliers

The multipliers form the majority of the basic building blocks of the signal processing chain in the analog baseband part of the MSPD. As shown previously, from Fig. 4.4, the multipliers are needed in the following subsystems of the kernels:

- Firstly, to retrieve the envelope component of the complex baseband signal, which is $x_{I,m}^2(t) + x_{Q,m}^2(t)$. hence multiplier acts as a squarer

- Obtaining the even order harmonics of the envelope. Multiplier acts as a squarer, similarly as above

- Multiplying the envelope component or its harmonics with the corresponding coefficients

- Finally, in the complex multiplier as shown in Eq. 4.11 and Eq. 4.12

For the case of complex valued 80 MHz bandwidth signal, the I and Q channel real signals have 40 MHz unilateral signal bandwidth each. Hence the squared envelope (envelope power) is a real signal which has 80 MHz unilateral bandwidth. The square of the envelope power signals hence has 160 MHz unilateral bandwidth. The bandwidth becomes 200 MHz after the final complex multiplication with both the I and Q channel separately. So the bilateral bandwidth of the complex valued signal output from the APD with a nonlinearity order $k = 5$ becomes 400 MHz. We can conclude that each of the multiplier has to accommodate different bandwidths in the signal chain.

The analog signal multiplication can be performed in either voltage-input voltage-output, namely voltage-mode [111] or as current-input current-output, known as current-mode [12] or in the other combinations.

Fig. 4.10 shows one of the state-of-the-art current-mode multiplier/divider circuit presented in [12]. The principle of operation is as explained. The gate-source potentials for the loops with transistors M1, M2, M3 and M4, and M1, M2, M6 and M7, can be expressed as:

$$2V_{\text{GS}}(I_2) = V_{\text{GS}}(I_{D1,2}) + V_{\text{GS}}[I_{D1,2} + 2(I_1 \mp I_0)], \tag{4.15}$$

results in

$$I_{D1,2} = I_2 - (I_1 \mp I_0) + \frac{(I_1 \mp I_0)^2}{4I_2}, \tag{4.16}$$

when the square law characteristic of the CMOS transistors is considered, which is an approximation. Hence, the output current expression is:

$$I_{\text{OUT}} = I_{D2} - I_{D1} + 2I_O, = \frac{I_O I_1}{I_2} \tag{4.17}$$
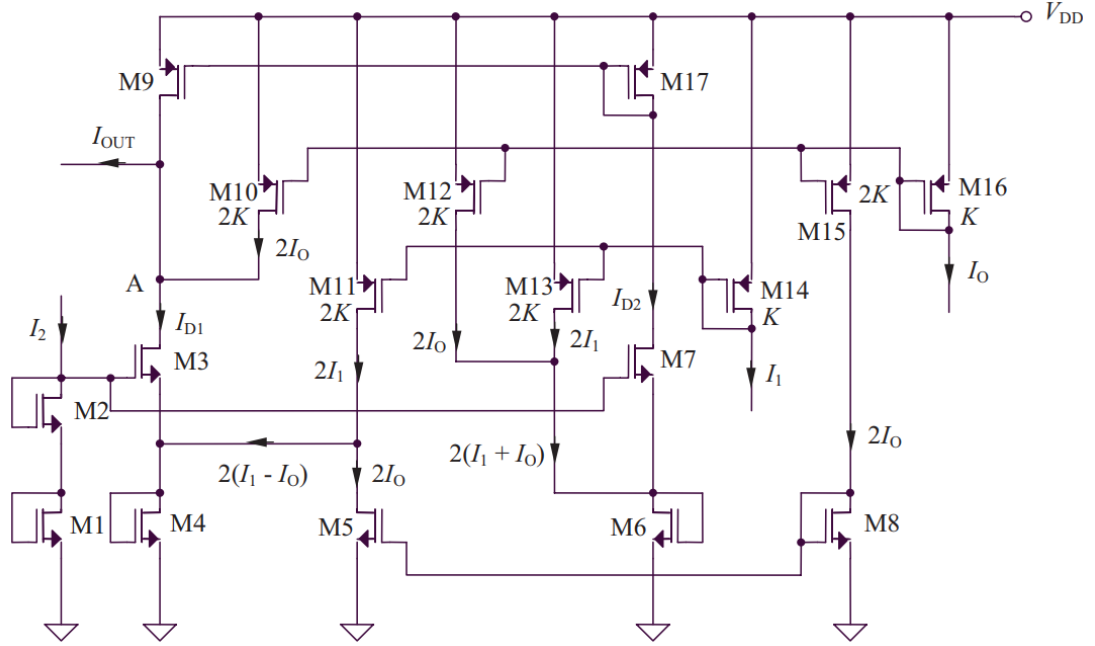
FIGURE 4.10: Current-mode multiplier circuit of [12]

In our case, where the circuit is used as multiplier, the current I_2 can be used to tune the output current range.

The non-idealities of the multiplier arises from the second-order effects of the CMOS transistors, namely, channel-length modulation, body-effect, mobility degradation and transistors mismatch. The influence of second-order effects are weaker when compared to that of the main square-law characteristic of the CMOS transistors. They can be dealt with different circuit and layout techniques to minimize or mitigate their effects as mentioned in [12]. The circuit is implemented in a 1.8 V nominal voltage, 180 nm CMOS process. The supply voltage used is 1.2 V. It achieves very good linearity, with a linearity error of 0.75% and a 3-dB bandwidth of approximately 80 MHz at a very low power consumption of 60 μW. The input-referred noise voltage is 0.6 μV/$\sqrt{\text{Hz}}$. This circuit can be hence utilized by scaling the bandwidth, as needed by the various multipliers of the MP part of the MSPD.

4.4.3 Time delays

As mentioned in Section 4.3, both the I and Q signals, that are the inputs to the MP APD, have to be delayed when being fed to the subsequent memory kernels. Note that the delay on the wideband signals, as in our case can be only performed in voltage domain [31].

The required t_d is about 1.4 ns for the case of input signal with 80 MHz of bandwidth, an amount equal to inverse of nine times the sampling frequency (720 MHz). The transfer function of an ideal delay cell is: $H(s) = e^{-st_d}$. Its gain is 1 and its phase is linear versus frequency as shown in Fig. 4.11.
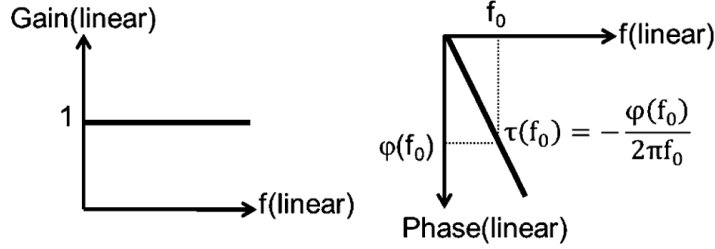
FIGURE 4.11: The gain and phase transfer function of an ideal time delay cell [13]

The delay t_d at frequency f_0 is:

$$t_d(f_0) = -\phi(f_0)/2\pi f_0, \tag{4.18}$$

ideally independent of f_0 (linear phase). Achieving constant true time delay is tougher, as it not only requires constant group delay independent of frequency but also a constant ratio between phase and frequency, which is independent of frequency [112, 13]. There are different IC compatible circuits to approximate a time delay, e.g., transmission lines, LC delay lines, switched capacitor delay circuits and gm-RC or gm-C all-pass filters [13]. However, at MHz and low-GHz frequencies, transmission lines and LC delay lines in CMOS are unpractical due to the low quality factor of coils, loss of the transmission lines and their large sizes. Switched capacitor time-delay circuits on the other hand are not fast enough for low-GHz applications. One of the few remaining options is to exploit an all-pass filter (APF) approximation of a delay. The simplest is a first-order all-pass filter given as [113]:

$$H_{ap1}(s) = 1 - s(t_d/2)\frac{1}{+}s(t_d/2) = 2\frac{1}{+}s(t_d/2) - 1. \tag{4.19}$$

The above equation of APF is a combination of low-pass filter (LPF) and an inverter. It can be implemented as gm-RC or gm-C topology. From the bandwidth performance and power consumption point of view, gm-C topology of [112, 13] outperforms gm-RC realizations, as proved in [112, 13]. The block-level view of the first-order APF is shown in Fig. 4.12 and the transfer function is as shown in Fig. 4.13. Known as true-time delays, these circuits are used in beamsteering applications, which demands delaying GHz bandwidth signal in ps to ns delay range [13, 114]
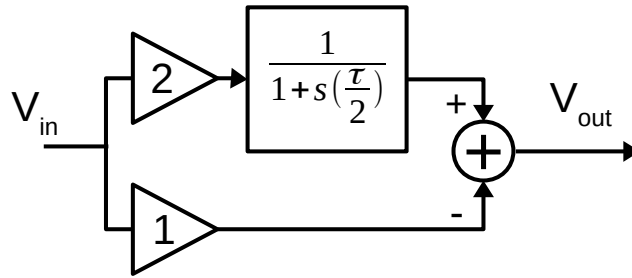
FIGURE 4.12: The block view of the first-order all-pass filter
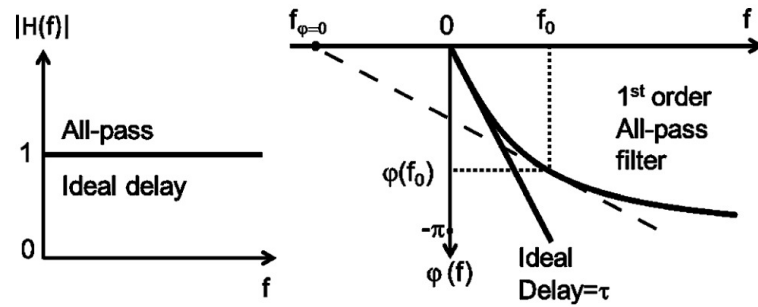
FIGURE 4.13: The gain and phase transfer function of a first order APF [13]
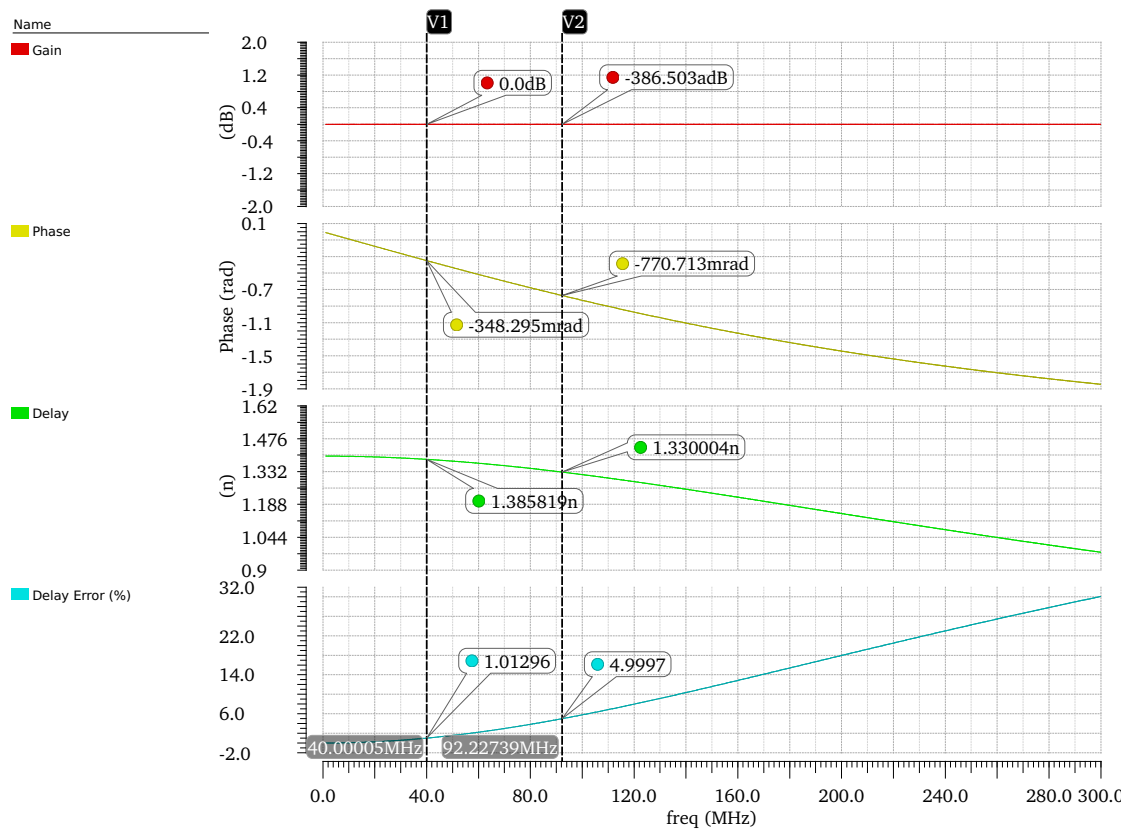
FIGURE 4.14: The gain and phase transfer function of an ideal first order APF with t_d of $1.4\,\mathrm{ns}$

Fig. 4.14 shows the response of the ideal first order APF model shown in Fig. 4.12, when the delay t_d is 1.4 ns. Only the input signal's in-phase and quadrature components, with an unilateral signal bandwidth of 40 MHz are subjected to the time delays. We can observe that the delay at $f_0 = 40$ MHz is about 1.386 ns, giving an error of about 1% with the ideal delay of 1.4 ns. Hence the whole modulated signal in-phase and quadrature components with unilateral bandwidth of 40 MHz can be easily delayed by a first order APF section, without any further enhancements of phase linearization and bandwidth extension, as discussed in [13]. It can be observed that the 5% error in delay occurs after 92 MHz.

The schematic of the gm-C delay cell proposed in [13] is as shown in Fig. 4.15 Delay tuning can be implemented using capacitor banks. Coarse delay tuning techniques can be employed to have the delay tuning by a large factor, for example if we intend to delay the signal by 2.8 ns instead of 1.4. In a similar way fine delay tuning can be used to tune around a coarse time-delay setting. The APF is implemented in 140 nm CMOS and consumes 90 mW power to delay a signal with $1 - 2.5$ GHz bandwidth by a maximum delay of 550 ps, with an error of±2%. Again, for our signal case we just need to delay 40 MHz signal hence the power consumption of the circuit implementation should be able to scale accordingly.
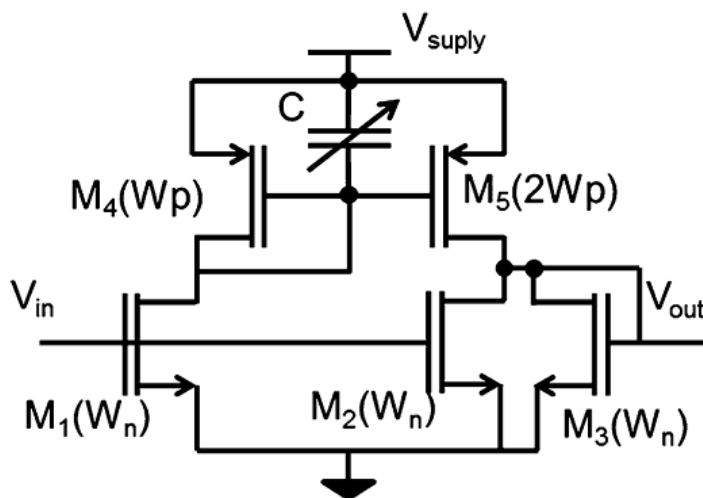
FIGURE 4.15: The first order APF of [13]

The state-of-the-art true-time delay element presented in [114] utilizes a 9th order APF, which is not necessary for our predistorter. It achieves the highest amount of delay and bandwidth product among the existing APF ICs and it can be exploited to see if the power consumption can be scaled down according to the signal bandwidth. It is realized with gm-C topology in 130 nm CMOS and consumes $112 - 364$ mW power to delay a signal with $0.1 - 2$ GHz bandwidth by a maximum delay of 1700 ps, with an error of±8%.

4.5   Conclusions

In this chapter, a mixed-signal predistorter architecture based on the FIR-MP, suitable for wideband RF power amplifier linearization is proposed. The drawbacks of the existing DPDs, which requires at least five times bandwidth for the transmitter chain, starting from the digital section, and that of ARFPDs, which requires power-hungry and challenging to design RF components, can be eliminated using the MSPD.

MP extracted PA model of ADL5606 is linearized over an 80 MHz signal. Results prove that the FIR-MP based MSPD in ideal system-level simulations provides an improvement of 14.3 dB and 13.5 dB, in adjacent and alternate channel leakage ratio, ACLR1 and ACLR2, respectively, in comparison with FIR envelope memory polynomial (FIR-EMP) model, used in ARFPD.

The impact of various non-idealities are assessed using behavioral-level electrical simulations. This has aided in deriving the requirements for the integrated circuit implementation. The major subsystems are namely the coefficient DACs, multipliers and time-delays. The simulations show that a resolution of 8 bits for the coefficient DACs and a signal path SNR of 60 dB is required to achieve ACLR1 above 45 dBc to meet the 3GPP ACLR specifications, with as little as 9 complex coefficients in the analog domain. Time-delay mismatch has been simulated with the help of Monte Carlo simulations and show that 5% mismatch is tolerable to still meet the 3GPP specifications.

We can conclude that the MSPD has potential to achieve low-power and low-cost and hence its suitability in the context of future small-cell base stations.

In the realization of the MSPD as an IC, we have provided a discussion on the potential technology choice and various suitable architectural choices for the subsystems to achieve high-performance at low-power.

Chapter 5

Conclusions and Future Directions

The next-generation 5G wireless communications aiming at increase of the network capacity by 1000 times, has to address various challenges. Cell densification using small-cell base stations is one among various solutions. The power amplifier is still the most power-hungry component in the base stations, whose linearity/efficiency trade-off can be addressed with predistortion. Low-power wideband predistortion methods in the context of 5G small-cell base stations have been developed in the thesis. With a brief introduction to the wireless systems and the power amplifier, Chapter 1 has provided the necessary background for the dissertation. Various memory-unaware and memory-aware digital (DPD) and analog RF predistortion (ARFPD) techniques have been briefly summarized in Chapter 2. Advantages and disadvantages of both the predistortion methods were discussed. We conclude that the box-oriented approaches such as FIR-EMP and Volterra simplifications such as MP tend to show promising capabilities and note the need for a new low-power wideband predistortion models in the context of 5G small-cell base stations.

Chapter 3 presents the proposed FIR-MP DPD model, which improves the memory correction performance of the memory polynomials. PA model extraction methodology along with the parameter identification for the FIR-MP using Small-Signal Assisted Parameter Identification (SSAPI) algorithm has been discussed. Based on the simulations performed on the MP models extracted on a small-cell base station PA, we have shown the linearization performance of the proposed FIR-MP DPD. The digital implementation of the proposed predistorter in 28nm FDSOI CMOS is presented.

A novel mixed-signal predistorter architecture based on the FIR-MP, suitable for wideband linearization of RF PA is presented in Chapter 4. The hybrid solution employing

digital FIR filter and MP analog predistorter (APD) overcomes the drawbacks of the existing DPDs and ARFPDs. The linearization performance of the proposed predistorter in an ideal configuration, along with most important non-idealities has been shown with the help of system-level and electrical-level behavioral simulations. Pointers to the potential architectures of the subsystems of the MSPD IC realization are provided.

Thesis Contributions

The contributions of the thesis are briefly summarized in this section

Proposed FIR-MP Model

The existing DPD models addressing memory effects are simplifications of full Volterra model, such as memory polynomial model [71] and generalized memory polynomial model [43]. Memory polynomial model's complexity is highly reduced when compared to the full Volterra model, thereby reducing its accuracy. Nonetheless, it is still one of the most attractive predistortion models providing significant performance with very few coefficients. A new low-complexity digital baseband predistorter with an FIR filter preceding a memory polynomial known as FIR-MP is proposed to improve the performance of the memory polynomial. The adjacent channel leakage ratio (ACLR) performance comparison between the conventional MP and the proposed FIR-MP is done based on simulations with multi-carrier modulated signals of 20 MHz and 80 MHz bandwidths. The PA models used for the simulations are extracted from the measurements of a commercial 1 W GaAs HBT PA (ADL5606). At the ideal system-level simulations, the improvements in ACLR over the conventional MP are 7.2 dB and 15.6 dB, respectively, for 20 MHz and 80 MHz signals.

Digital CMOS Implementation

Implementation of the high-performance and low-complexity digital baseband predistorter based on the proposed FIR-MP model in 28 nm Fully-Depleted Silicon-on-Insulator (FDSOI) Complementary Metal Oxide Semiconductor (CMOS) technology has been carried out. It is shown that with a fraction of the power and die area of that of the MP, a huge improvement in ACLR is attained. The choice of selection of various parameters of the predistorter along with the subsequent digital-to-analog converter (DAC) is presented. The impact of fixed-point representation is assessed using ACLR metrics, which shows that a wordlength of 14 bits is sufficient to obtain ACLR beyond

45 dBc with a margin of 10 dB. With an overall power consumption of 9.18 mW and 116.2 mW, respectively, for 20 MHz and 80 MHz signals, the FIR-MP DPD proves to be a suitable candidate for small-cell base station PA linearization. For the 20 MHz signal bandwidth case the FIR-MP DPD solution in 28nm FDSOI CMOS consumes 9.18 mW, which is about 4.4X lower in comparison with the state-of-the-art integrated solution of DPD (40 mW) [46] and 22X lower than ARFPD IC (200 mW) [8].

Proposed Mixed-Signal Predistorter

We have explored the possibility of implementing the proposed FIR-MP model in mixed signal domain, known as mixed-signal predistorter (MSPD), as an alternative solution to existing DPD and ARFPD solutions, which can significantly reduce the hardware complexity and challenges involved in their design. The FIR filter and MP sections are partitioned into digital and analog domains respectively. This way the digital FIR filter improves the memory correction performance without any bandwidth expansion and the MP predistorter in analog baseband provides superior linearization. MSPD avoids 5X bandwidth requirement for the transmitter and the power-hungry RF components when compared to DPDs and ARFPDs, respectively. This makes the MSPD solution a very low-power candidate and especially attractive in the context of small-cell base stations. The PA model of ADL5606 is linearized over an 80 MHz signal; results prove that the FIR-MP based MSPD in ideal system-level simulations provides an improvement of 14.3 dB and 13.5 dB in adjacent and alternate channel leakage ratio ACLR1 and ACLR2, respectively, in comparison with FIR envelope memory polynomial (FIR-EMP) model used in ARFPD. The impact of various non-idealities are simulated at electrical-level to derive the requirements for the integrated circuit implementation. The simulations show that a resolution of 8 bits for the coefficients and a signal path SNR of 60 dB are required to achieve ACLR1 above 45 dBc, with as little as 9 complex coefficients in the analog domain. A brief study of major subsystems of the MSPD is carried out to derive the initial specifications. The major subsystems are multipliers, coefficient DACs and analog time delays. With a study on the existing architectures of these subsystems, we have provided a discussion on various possible suitable architectures that can be employed to obtain the silicon realization of the MSPD.

Future Directions

The thesis has resulted in the development of FIR-MP predistorter model, which can be extended in a number of following directions:

- Firstly, the digital FIR filter is clocked at a higher sampling rate, a rate equal to that demanded by the MP section. This is similar to that of FIR-EMP implementations as presented in [9, 10]. Though the filter is linear and operates on just the signal bandwidth we had to do this to facilitate parameter identification of FIR filter and MP section at the same rate, which might not be necessary. Hence, there is a necessity to study multi-rate coefficient identification methodology so that the FIR filter during parameter identification is clocked at high rate in order to facilitate the learning of MP and once learnt can be clocked at a frequency equal to the signal bandwidth, when operated in the forward path. This will be extremely beneficial for the proposed mixed-signal predistorter (MSPD) implementation as it will have digital baseband section without interpolator and the FIR section being clocked at low rate. Note that for FIR-MP DPD implementation we still need interpolator in order to accommodate the IMD correction components.

- Based on the proposed FIR-MP model and the mixed-signal predistorter (MSPD) architecture, an implementation in CMOS would be a logical next step. As explained in Section. 4.4, all the major subsystems show feasibility towards low-power analog CMOS predistorter IC implementation. Ultimately, the performance and power consumption of the MSPD IC should be evaluated with the real PA in the measurements step.

- Performing measurements on various types of PAs and checking the impact of PA variabilities on the predistorter.

- For the digital implementation of FIR-MP, LUT approach similar to [46] can be an interesting proposition. LUTs with memory can reduce the computational complexity of the DPD.

- A short-term direction for the DPD implementation in 28 nm FDSOI CMOS process could be to exploit its excellent body biasing capability. With forward body-biasing the threshold of the transistors can be reduced subsequently reducing the dynamic power consumption. This can be detrimental to the leakage power consumption.

- The scope of the thesis was limited to single-band, single-input single-output (SISO) linearization methods, it will be interesting to extend FIR-MP into dual-band, triple-band and multi-input multi-output (MIMO) predistortion systems

Bibliography

[1] A. Birafane, M. El-Asmar, A. B. Kouki, M. Helaoui, and F. M. Ghannouchi, "Analyzing LINC Systems," *IEEE Microwave Magazine*, vol. 11, no. 5, pp. 59--71, Aug. 2010.

[2] J. K. Cavers, "Amplifier linearization using a digital predistorter with fast adaptation and low memory requirements," *IEEE Transactions on Vehicular Technology*, vol. 39, no. 4, pp. 374--382, Nov. 1990.

[3] Y. Cho, J. Lee, S. Jin, B. Park, J. Moon, J. Kim, and B. Kim, "Fully Integrated CMOS Saturated Power Amplifier With Simple Digital Predistortion," *IEEE Microwave and Wireless Components Letters*, vol. 24, no. 8, pp. 533--535, Aug. 2014.

[4] P. Jardin and G. Baudoin, "Filter Lookup Table Method for Power Amplifier Linearization," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 3, pp. 1076--1087, May 2007.

[5] S. Boumaiza, J. Li, and F. M. Ghannouchi, "Implementation of an adaptive digital/RF predistorter using direct LUT synthesis," in *2004 IEEE MTT-S International Microwave Symposium Digest (IEEE Cat. No.04CH37535)*, vol. 2, Jun. 2004, pp. 681--684 Vol.2.

[6] S. Boumaiza, J. Li, M. Jaidane-Saidane, and F. M. Ghannouchi, "Adaptive digital/RF predistortion using a nonuniform LUT indexing function with built-in dependence on the amplifier nonlinearity," *IEEE Transactions on Microwave Theory and Techniques*, vol. 52, no. 12, pp. 2670--2677, Dec. 2004.

[7] K. Y. Son, B. Koo, and S. Hong, "A CMOS Power Amplifier With a Built-In RF Predistorter for Handset Applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 8, pp. 2571--2580, Aug. 2012.

[8] F. Roger, "A 200mW 100MHz-to-4GHz 11th-order complex analog memory polynomial predistorter for wireless infrastructure RF amplifiers," in *Solid-State Circuits*

Conference Digest of Technical Papers (ISSCC), 2013 IEEE International, Feb. 2013, pp. 94--95.

[9] H. Huang, A. Islam, J. Xia, P. Levine, and S. Boumaiza, ''Linear filter assisted envelope memory polynomial for analog/radio frequency predistortion of power amplifiers,'' in *Microwave Symposium (IMS), 2015 IEEE MTT-S International*, May 2015, pp. 1--3.

[10] H. Huang, J. Xia, A. Islam, E. Ng, P. M. Levine, and S. Boumaiza, ''Digitally Assisted Analog/RF Predistorter With a Small-Signal-Assisted Parameter Identification Algorithm,'' *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 12, pp. 4297--4305, Dec. 2015.

[11] J. J. Wikner, *Studies on CMOS Digital-to-Analog Converters*, ser. Linköping studies in science and technology Dissertations.   Linköping: Univ, 2001, no. 667, oCLC: 248446368.

[12] C. Popa, ''Improved Accuracy Current-Mode Multiplier Circuits With Applications in Analog Signal Processing,'' *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 22, no. 2, pp. 443--447, Feb. 2014.

[13] S. K. Garakoui, E. A. M. Klumperink, B. Nauta, and F. E. van Vliet, ''Compact Cascadable g m -C All-Pass True Time Delay Cell With Reduced Delay Variation Over Frequency,'' *IEEE Journal of Solid-State Circuits*, vol. 50, no. 3, pp. 693--703, Mar. 2015.

[14] I. Humar, X. Ge, L. Xiang, M. Jo, M. Chen, and J. Zhang, ''Rethinking energy efficiency models of cellular networks with embodied energy,'' *IEEE Network*, vol. 25, no. 2, pp. 40--49, Mar. 2011.

[15] M. Pickavet, W. Vereecken, S. Demeyer, P. Audenaert, B. Vermeulen, C. Develder, D. Colle, B. Dhoedt, and P. Demeester, ''Worldwide energy needs for ICT: The rise of power-aware networking,'' in *2008 2nd International Symposium on Advanced Networks and Telecommunication Systems*, Dec. 2008, pp. 1--3.

[16] M. Deruyck, W. Vereecken, E. Tanghe, W. Joseph, M. Pickavet, L. Martens, and P. Demeester, ''Power consumption in wireless access network,'' in *2010 European Wireless Conference (EW)*, Apr. 2010, pp. 924--931.

[17] W. Vereecken, W. V. Heddeghem, M. Deruyck, B. Puype, B. Lannoo, W. Joseph, D. Colle, L. Martens, and P. Demeester, ''Power consumption in telecommunication networks: Overview and reduction strategies,'' *IEEE Communications Magazine*, vol. 49, no. 6, pp. 62--69, Jun. 2011.

[18] J. Lorincz, T. Garma, and G. Petrovic, "Measurements and Modelling of Base Station Power Consumption under Real Traffic Loads," *Sensors*, vol. 12, no. 4, pp. 4281--4310, Mar. 2012.

[19] A. Chatzipapas, S. Alouf, and V. Mancuso, "On the minimization of power consumption in base stations using on/off power amplifiers," in *2011 IEEE Online Conference on Green Communications*, Sep. 2011, pp. 18--23.

[20] X. Ge, J. Yang, H. Gharavi, and Y. Sun, "Energy Efficiency Challenges of 5G Small Cell Networks," *IEEE Communications Magazine*, vol. 55, no. 5, pp. 184--191, May 2017.

[21] C. Desset, B. Debaillie, V. Giannini, A. Fehske, G. Auer, H. Holtkamp, W. Wajda, D. Sabella, F. Richter, M. J. Gonzalez, H. Klessig, I. Gódor, M. Olsson, M. A. Imran, A. Ambrosy, and O. Blume, "Flexible power modeling of LTE base stations," in *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2012, pp. 2858--2862.

[22] L. Guan and A. Zhu, "Green Communications: Digital Predistortion for Wideband RF Power Amplifiers," *IEEE Microwave Magazine*, vol. 15, no. 7, pp. 84--99, Nov. 2014.

[23] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 122--130, Feb. 2014.

[24] B. Bangerter, S. Talwar, R. Arefi, and K. Stewart, "Networks and devices for the 5G era," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 90--96, Feb. 2014.

[25] L. Dartois, "LTE and Pre-5G Physical Layer Standards and Implementations," Paris, May 2017.

[26] X. Dong, "Designing remote radio heads (RRHs) on high-performance FPGAs," https://www.eetimes.com/document.asp?doc_id=1278555, Jul. 2011.

[27] M. Alzenad, M. Z. Shakir, H. Yanikomeroglu, and M. Alouini, "FSO-Based Vertical Backhaul/Fronthaul Framework for 5G+ Wireless Networks," *IEEE Communications Magazine*, vol. 56, no. 1, pp. 218--224, Jan. 2018.

[28] C. Han, T. Harrold, S. Armour, I. Krikidis, S. Videv, P. M. Grant, H. Haas, J. S. Thompson, I. Ku, C. X. Wang, T. A. Le, M. R. Nakhai, J. Zhang, and L. Hanzo, "Green radio: Radio techniques to enable energy-efficient wireless networks," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 46--54, Jun. 2011.

[29] 3GPP, "3GPP specification: LTE Evolved Universal Terrestrial Radio Access (E-UTRA) ; Base Station(BS) radio transmission and reception (3GPP TS 36.104 version 15.0.0 Release 15)," http://www.3gpp.org/dynareport/36104.htm, 2017.

[30] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3GPP heterogeneous networks," *IEEE Wireless Communications*, vol. 18, no. 3, pp. 10--21, Jun. 2011.

[31] B. Razavi, *RF Microelectronics.* Prentice Hall, 2012.

[32] P. Reynaert, "Polar Modulation," *IEEE Microwave Magazine*, vol. 12, no. 1, pp. 46--51, Feb. 2011.

[33] E. McCune, *Practical Digital Wireless Signals.* Cambridge University Press, Feb. 2010.

[34] P. Banelli, S. Buzzi, G. Colavolpe, A. Modenini, F. Rusek, and A. Ugolini, "Modulation Formats and Waveforms for 5G Networks: Who Will Be the Heir of OFDM?: An overview of alternative modulation schemes for improved spectral efficiency," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 80--93, Nov. 2014.

[35] X. Zhang, L. Chen, J. Qiu, and J. Abdoli, "On the Waveform for 5G," *IEEE Communications Magazine*, vol. 54, no. 11, pp. 74--80, Nov. 2016.

[36] Y. Cai, Z. Qin, F. Cui, G. Y. Li, and J. A. McCann, "Modulation and Multiple Access for 5G Networks," *IEEE Communications Surveys Tutorials*, vol. 20, no. 1, pp. 629--646, Firstquarter 2018.

[37] F. M. Ghannouchi, O. Hammi, and M. Helaoui, *Behavioral Modeling and Predistortion of Wideband Wireless Transmitters.* John Wiley & Sons, Jul. 2015.

[38] M. K. Kazimierczuk, *RF Power Amplifier.* John Wiley & Sons, Dec. 2014.

[39] S. C. Cripps, "RF Power Amplifiers for Wireless Communications, (Artech House Microwave Library)," *Artech House*, 2006.

[40] E. Westesson and L. Sundstrom, "A complex polynomial predistorter chip in CMOS for baseband or IF linearization of RF power amplifiers," in *Proceedings of the 1999 IEEE International Symposium on Circuits and Systems, 1999. ISCAS '99*, vol. 1, Jul. 1999, pp. 206--209 vol.1.

[41] ------, "Low-power complex polynomial predistorter circuit in CMOS for RF power amplifier linearization," in *Solid-State Circuits Conference, 2001. ESSCIRC 2001. Proceedings of the 27th European*, Sep. 2001, pp. 486--489.

[42] L. Ding, G. Zhou, D. Morgan, Z. Ma, J. Kenney, J. Kim, and C. Giardina, ''A robust digital baseband predistorter constructed using memory polynomials,'' *IEEE Transactions on Communications*, vol. 52, no. 1, pp. 159--165, Jan. 2004.

[43] D. Morgan, Z. Ma, J. Kim, M. Zierdt, and J. Pastalan, ''A Generalized Memory Polynomial Model for Digital Predistortion of RF Power Amplifiers,'' *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3852--3860, Oct. 2006.

[44] P. Desgreys, V. N. Manyam, K. Tchambake, D. K. G. Pham, and C. Jabbour, ''Wideband power amplifier predistortion: Trends, challenges and solutions,'' in *2017 IEEE 12th International Conference on ASIC (ASICON)*, Oct. 2017, pp. 100--103.

[45] F. Ghannouchi and O. Hammi, ''Behavioral modeling and predistortion,'' *IEEE Microwave Magazine*, vol. 10, no. 7, pp. 52--64, Dec. 2009.

[46] C. Mayer, D. McLaurin, J. Fan, S. Bal, C. Angell, O. Gysel, M. McCormick, M. Manglani, R. Schubert, B. Reggiannini, J. Kornblum, L. Wu, L. Leonard, S. Bhal, A. Kagan, and T. Montalvo, ''A direct-conversion transmitter for small-cell cellular base stations with integrated digital predistortion in 65nm CMOS,'' in *2016 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, May 2016, pp. 63--66.

[47] M. Hussein, O. Venard, B. Feuvrie, and Y. Wang, ''Digital predistortion for RF power amplifiers: State of the art and advanced approaches,'' in *New Circuits and Systems Conference (NEWCAS), 2013 IEEE 11th International*, Jun. 2013, pp. 1--4.

[48] P. L. Gilabert, G. Montoro, and E. Bertran, ''FPGA Implementation of a Real-Time NARMA-Based Digital Adaptive Predistorter,'' *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 58, no. 7, pp. 402--406, Jul. 2011.

[49] H. Qian, H. Huang, and S. Yao, ''A General Adaptive Digital Predistortion Architecture for Stand-Alone RF Power Amplifiers,'' *IEEE Transactions on Broadcasting*, vol. 59, no. 3, pp. 528--538, Sep. 2013.

[50] Y. Ma, Y. Yamao, Y. Akaiwa, and C. Yu, ''FPGA Implementation of Adaptive Digital Predistorter With Fast Convergence Rate and Low Complexity for Multi-Channel Transmitters,'' *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 11, pp. 3961--3973, Nov. 2013.

[51] S. A. Bassam, M. Helaoui, and F. M. Ghannouchi, ''2-D Digital Predistortion (2-D-DPD) Architecture for Concurrent Dual-Band Transmitters,'' *IEEE Transactions on Microwave Theory and Techniques*, vol. 59, no. 10, pp. 2547--2553, Oct. 2011.

[52] C. Quindroit, N. Naraharisetti, P. Roblin, S. Gheitanchi, V. Mauer, and M. Fitton, "FPGA implementation of orthogonal 2D digital predistortion system for concurrent dual-band power amplifiers based on time-division multiplexing," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 61, no. 12, pp. 4591--4599, 2013.

[53] Y. Liu, P. Roblin, Y. Hai, S. Shao, and Y. Tang, "Novel multiband linearization technique for closely-spaced dual-band signals of wide bandwidth," in *Microwave Symposium (IMS), 2015 IEEE MTT-S International*, May 2015, pp. 1--4.

[54] B. Fehri and S. Boumaiza, "Dual-Band Digital Predistortion Using a Single Transmitter Observation Receiver and Single Training Engine," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 1, pp. 315--321, Jan. 2017.

[55] S. Amin, P. Händel, and D. Rönnow, "Digital Predistortion of Single and Concurrent Dual-Band Radio Frequency GaN Amplifiers With Strong Nonlinear Memory Effects," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 7, pp. 2453--2464, Jul. 2017.

[56] M. Younes, A. Kwan, M. Rawat, and F. Ghannouchi, "Three-Dimensional digital predistorter for concurrent tri-band power amplifier linearization," in *Microwave Symposium Digest (IMS), 2013 IEEE MTT-S International*, Jun. 2013, pp. 1--4.

[57] P. M. Suryasarman and A. Springer, "A Comparative Analysis of Adaptive Digital Predistortion Algorithms for Multiple Antenna Transmitters," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 62, no. 5, pp. 1412--1420, May 2015.

[58] Z. A. Khan, E. Zenteno, P. Händel, and M. Isaksson, "Digital Predistortion for Joint Mitigation of I/Q Imbalance and MIMO Power Amplifier Distortion," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 1, pp. 322--333, Jan. 2017.

[59] C. Yu, L. Guan, E. Zhu, and A. Zhu, "Band-Limited Volterra Series-Based Digital Predistortion for Wideband RF Power Amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 12, pp. 4198--4208, Dec. 2012.

[60] Y. Liu, J. Yan, H.-T. Dabag, and P. Asbeck, "Novel Technique for Wideband Digital Predistortion of Power Amplifiers With an Under-Sampling ADC," *IEEE Transactions on Microwave Theory and Techniques*, vol. 62, no. 11, pp. 2604--2617, Nov. 2014.

[61] H. Ku, M. D. McKinley, and J. S. Kenney, ''Quantifying memory effects in RF power amplifiers,'' *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 12, pp. 2843--2849, Dec. 2002.

[62] P. Asbeck, L. Larson, D. Kimball, and J. Buckwalter, ''CMOS handset power amplifiers: Directions for the future,'' in *2012 IEEE Custom Integrated Circuits Conference (CICC)*, Sep. 2012, pp. 1--6.

[63] C. Presti, D. Kimball, and P. Asbeck, ''Closed-Loop Digital Predistortion System With Fast Real-Time Adaptation Applied to a Handset WCDMA PA Module,'' *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 3, pp. 604--618, Mar. 2012.

[64] J. Wood, ''What's New in Digital Predistortion,'' https://ieeetv.ieee.org/video/what-s-new-in-digital-predistortion, Jul. 2015.

[65] C. D. Presti, A. G. Metzger, H. M. Banbrook, P. J. Zampardi, and P. M. Asbeck, ''Efficiency improvement of a handset WCDMA PA module using Adaptive Digital Predistortion,'' in *Microwave Symposium Digest (MTT), 2010 IEEE MTT-S International*, May 2010, pp. 804--807.

[66] A. A. M. Saleh and J. Salz, ''Adaptive linearization of power amplifiers in digital radio systems,'' *The Bell System Technical Journal*, vol. 62, no. 4, pp. 1019--1033, Apr. 1983.

[67] J. Wood, ''System-Level Design Considerations for Digital Pre-Distortion of Wireless Base Station Transmitters,'' *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 5, pp. 1880--1890, May 2017.

[68] V. Volterra, *Sopra le funzioni che dipendono da altre funzioni.* Tip. della R. Accademia dei Lincei, 1887.

[69] ------, *Theory of Functionals and of Integral and Integro-Differential Equations.* Courier Corporation, Jan. 2005.

[70] J. Staudinger, J. C. Nanan, and J. Wood, ''Memory fading Volterra series model for high power infrastructure amplifiers,'' in *2010 IEEE Radio and Wireless Symposium (RWS)*, Jan. 2010, pp. 184--187.

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

[72] T. Jiang, ''Behavioral modeling and FPGA implementation of digital predistortion for RF and microwave power amplifiers,'' Politecnico di Torino, Tech. Rep., 2016.

[73] W. Gao, ''Linearization of Wideband Wi-Fi Power Amplifiers Using RF Analog Memory Predistortion,'' in *2018 IEEE International Conference on Communications (ICC)*, May 2018, pp. 1--6.

[74] A. Zhu, J. C. Pedro, and T. J. Brazil, ''Dynamic Deviation Reduction-Based Volterra Behavioral Modeling of RF Power Amplifiers,'' *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 12, pp. 4323--4332, Dec. 2006.

[75] H. W. Kang, Y. S. Cho, and D. H. Youn, ''Adaptive precompensation of Wiener systems,'' *IEEE Transactions on Signal Processing*, vol. 46, no. 10, pp. 2825--2829, Oct. 1998.

[76] P. Gilabert, G. Montoro, and E. Bertran, ''On the Wiener and Hammerstein models for power amplifier predistortion,'' in *Microwave Conference Proceedings, 2005. APMC 2005. Asia-Pacific Conference Proceedings*, vol. 2, Dec. 2005, pp. 4 pp.--.

[77] K. Narendra and P. Gallman, ''An iterative method for the identification of nonlinear systems using a Hammerstein model,'' *IEEE Transactions on Automatic Control*, vol. 11, no. 3, pp. 546--550, Jul. 1966.

[78] T. Liu, S. Boumaiza, and F. M. Ghannouchi, ''Deembedding static nonlinearities and accurately identifying and modeling memory effects in wide-band RF transmitters,'' *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 11, pp. 3578--3587, Nov. 2005.

[79] ------, ''Augmented hammerstein predistorter for linearization of broad-band wireless transmitters,'' *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 4, pp. 1340--1349, Jun. 2006.

[80] O. Hammi and F. M. Ghannouchi, ''Twin Nonlinear Two-Box Models for Power Amplifiers and Transmitters Exhibiting Memory Effects With Application to Digital Predistortion,'' *IEEE Microwave and Wireless Components Letters*, vol. 19, no. 8, pp. 530--532, Aug. 2009.

[81] F. Mkadem and S. Boumaiza, ''Physically Inspired Neural Network Model for RF Power Amplifier Behavioral Modeling and Digital Predistortion,'' *IEEE Transactions on Microwave Theory and Techniques*, vol. 59, no. 4, pp. 913--923, Apr. 2011.

[82] R. N. Braithwaite, ''Wide bandwidth adaptive digital predistortion of power amplifiers using reduced order memory correction,'' in *Microwave Symposium Digest, 2008 IEEE MTT-S International*, Jun. 2008, pp. 1517--1520.

[83] J. Namiki, "An Automatically Controlled Predistorter for Multilevel Quadrature Amplitude Modulation," *IEEE Transactions on Communications*, vol. 31, no. 5, pp. 707--712, May 1983.

[84] T. Nojima, Y. Okamoto, and S. Ohyama, "Predistortion nonlinear compensator for microwave SSB-AM system," *Electronics and Communications in Japan (Part I: Communications)*, vol. 67, no. 5, pp. 57--66, 1984.

[85] W. Woo, M. Miller, and J. Kenney, "A hybrid digital/RF envelope predistortion linearization system for power amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 1, pp. 229--237, Jan. 2005.

[86] S. P. Stapleton and F. C. Costescu, "An adaptive predistortion system," in *[1992 Proceedings] Vehicular Technology Society 42nd VTS Conference - Frontiers of Technology*, May 1992, pp. 690--693 vol.2.

[87] ------, "An adaptive predistorter for a power amplifier based on adjacent channel emissions [mobile communications]," *IEEE Transactions on Vehicular Technology*, vol. 41, no. 1, pp. 49--56, Feb. 1992.

[88] T. Rahkonen, T. Kankaala, and M. Neitola, "A programmable analog polynomial predistortion circuit for linearising radio transmitters," in *Solid-State Circuits Conference, 1998. ESSCIRC '98. Proceedings of the 24th European*, Sep. 1998, pp. 276--279.

[89] K. Gumber, P. Jaraut, M. Rawat, and K. Rawat, "Digitally assisted analog predistortion technique for power amplifier," in *2016 88th ARFTG Microwave Measurement Conference (ARFTG)*, Dec. 2016, pp. 1--4.

[90] K. Gumber and M. Rawat, "A Modified Hybrid RF Predistorter Linearizer for Ultra Wideband 5G Systems," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 7, no. 4, pp. 547--557, Dec. 2017.

[91] T. Rahkonen, T. Kankaala, M. Neitola, and A. Heiskanen, "Using Analog Predistortion for Linearizing Class A–C Power Amplifiers," *Analog Integrated Circuits and Signal Processing*, vol. 22, no. 1, pp. 31--39, Jan. 2000.

[92] O. Hammi, F. Ghannouchi, and B. Vassilakis, "A Compact Envelope-Memory Polynomial for RF Transmitters Modeling With Application to Baseband and RF-Digital Predistortion," *IEEE Microwave and Wireless Components Letters*, vol. 18, no. 5, pp. 359--361, May 2008.

[93] J. Xia, E. Ng, and S. Boumaiza, "Wideband Compensation of RF Vector Multiplier for RF Predistortion Systems," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 63, no. 11, pp. 1084--1088, Nov. 2016.

[94] L. Ding, Z. Ma, D. R. Morgan, M. Zierdt, and J. Pastalan, "A least-squares/Newton method for digital predistortion of wideband signals," *IEEE Transactions on Communications*, vol. 54, no. 5, pp. 833--840, May 2006.

[95] J. Wood, *Behavioral Modeling and Linearization of RF Power Amplifiers*. Artech House, Jun. 2014.

[96] H. Ku and J. S. Kenney, "Behavioral modeling of nonlinear RF power amplifiers considering memory effects," *IEEE Transactions on Microwave Theory and Techniques*, vol. 51, no. 12, pp. 2495--2504, Dec. 2003.

[97] N. Messaoudi, M. C. Fares, S. Boumaiza, and J. Wood, "Complexity reduced odd-order memory polynomial pre-distorter for 400-watt multi-carrier Doherty amplifier linearization," in *2008 IEEE MTT-S International Microwave Symposium Digest*, Jun. 2008, pp. 419--422.

[98] "ADL5606 Datasheet and Product Info — Analog Devices," http://www.analog.com/en/products/amplifiers/rf-amplifiers/driver-amplifiers/adl5606.html#product-overview.

[99] D. Na and K. Choi, "Low PAPR FBMC," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 182--193, Jan. 2018.

[100] "Mini-Circuits ZFRSC-42," https://www.minicircuits.com/WebStore/dashboard.html?model=ZFRSC-42-S%2B.

[101] T. Gotthans, G. Baudoin, and A. Mbaye, "Optimal order estimation for modeling and predistortion of power amplifiers," in *2013 IEEE International Conference on Microwaves, Communications, Antennas and Electronic Systems (COMCAS 2013)*, Oct. 2013, pp. 1--4.

[102] W. T. Lin, H. Y. Huang, and T. H. Kuo, "A 12-bit 40 nm DAC Achieving SFDR > 70 dB at 1.6 GS/s and IMD < –61dB at 2.8 GS/s With DEMDRZ Technique," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 3, pp. 708--717, Mar. 2014.

[103] L. Guan and A. Zhu, "Low-Cost FPGA Implementation of Volterra Series-Based Digital Predistorter for RF Power Amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 58, no. 4, pp. 866--872, Apr. 2010.

[104] L. Guan, R. Kearney, C. Yu, and A. Zhu, "High-performance digital predistortion test platform development for wideband RF power amplifiers," *International Journal of Microwave and Wireless Technologies*, vol. 5, no. 02, pp. 149--162, Apr. 2013.

[105] T. Jiang, R. Quaglia, V. Camarchia, and M. Pirola, ''FPGA-based digital predistortion of A 3.5 GHz GaN Doherty power amplifier,'' in *10th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM 2014)*, Sep. 2014, pp. 20--24.

[106] ''HDL Coder,'' https://www.mathworks.com/products/hdl-coder.html.

[107] ''Fixed-Point Designer,'' https://www.mathworks.com/products/fixed-point-designer.html.

[108] M. Gustavsson, J. J. Wikner, and N. Tan, *CMOS Data Converters for Communications*, ser. The Springer International Series in Engineering and Computer Science. Springer US, 2000.

[109] T. C. Carusone, D. Johns, and K. Martin, *Analog Integrated Circuit Design*, 2nd ed. Hoboken, NJ: Wiley, Dec. 2011.

[110] P. Harpe, E. Cantatore, and A. van Roermund, ''A 10b/12b 40 kS/s SAR ADC With Data-Driven Noise Reduction Achieving up to 10.1b ENOB at 2.2 fJ/Conversion-Step,'' *IEEE Journal of Solid-State Circuits*, vol. 48, no. 12, pp. 3011--3018, Dec. 2013.

[111] J. N. Babanezhad and G. C. Temes, ''A 20-V four-quadrant CMOS analog multiplier,'' *IEEE Journal of Solid-State Circuits*, vol. 20, no. 6, pp. 1158--1168, Dec. 1985.

[112] S. K. Garakoui, E. A. M. Klumperink, B. Nauta, and F. F. E. V. Vliet, ''A 1-to-2.5GHz phased-array IC based on gm-RC all-pass time-delay cells,'' in *2012 IEEE International Solid-State Circuits Conference*, Feb. 2012, pp. 80--82.

[113] K. Bult and H. Wallinga, ''A CMOS analog continuous-time delay line with adaptive delay-time control,'' *IEEE Journal of Solid-State Circuits*, vol. 23, no. 3, pp. 759--766, Jun. 1988.

[114] I. Mondal and N. Krishnapura, ''A 2-GHz Bandwidth, 0.25-1.7 ns True-Time-Delay Element Using a Variable-Order All-Pass Filter Architecture in 0.13 mu m CMOS,'' *IEEE Journal of Solid-State Circuits*, vol. 52, no. 8, pp. 2180--2193, Aug. 2017.

Appendix A

Principle of IMD cancelation in RF

The basic principle of operation for an ARFPD can be understood by taking cube of a unit amplitude two tone RF input signal, $X(t) = \cos(\omega_1 t) + \cos(\omega_2 t)$, where ω_1 and ω_2 are frequencies inside the signal BW. This is similar to the nonlinear distortion that the signals get subjected to in a PA, which results in the following equations:

$$
\begin{aligned}
X^3 = & \frac{1}{4}[\cos(3\omega_1) + \cos(3\omega_2)] + 3[\cos(2\omega_1 + \omega_2) + \cos(\omega_1 + 2\omega_2)] \\
& + 9[\cos(\omega_1) + \cos(\omega_2)] \\
& + \mathbf{3[\cos(2\omega_1 - \omega_2) + \cos(2\omega_2 - \omega_1)]}
\end{aligned}
\tag{A.1}
$$

In the above and forthcoming equations t is removed for simplicity. In (A.1) except for the last two terms, which are IMD3 terms, all the other terms either fall outside the band of interest or on the top of the original signal itself.

In a similar way if we square the signal we can get

$$
\begin{aligned}
X^2 = & \frac{1}{2}[\cos(2\omega_1) + \cos(2\omega_2)] + \cos(\omega_1 + \omega_2) + 1 \\
& + \mathbf{\cos(\omega_1 - \omega_2)}
\end{aligned}
\tag{A.2}
$$

The squared signal consists of high frequency terms and DC terms, which are not of interest but also contains a baseband signal term $cos(\omega_1 - \omega_2)$

Now to cancel out the IMD terms, we can multiply the RF signal X by the baseband correction signal $Y = k_3 \cos(\omega_1 - \omega_2)$, which gives

$$X.Y = [\cos(\omega_1) + \cos(\omega_2)]k_3\cos(\omega_1 - \omega_2)$$

$$= \frac{k_3}{2}\underbrace{[\cos(\omega_1) + \cos(\omega_2)]}_{\text{Original input signal}} + \frac{k_3}{2}\underbrace{[\cos(2\omega_1 - \omega_2) + \cos(2\omega_2 - \omega_1)]}_{\text{Predistortion signal}} \quad \text{(A.3)}$$

Hence by controlling the coefficient $k3$ we can cancel out the IMD3 components $((2\omega_1 - \omega_2)$ and $(2\omega_2 - \omega_1))$ of the PA when the input to the PA is $X.Y$, which is the sum of original signal and the predistortion signal, instead of just the original signal X. In a similar way other higher order IMD components can be canceled out.

Fig. A.1 illustrates the method of generation of the predistortion signal, as discussed formerly.
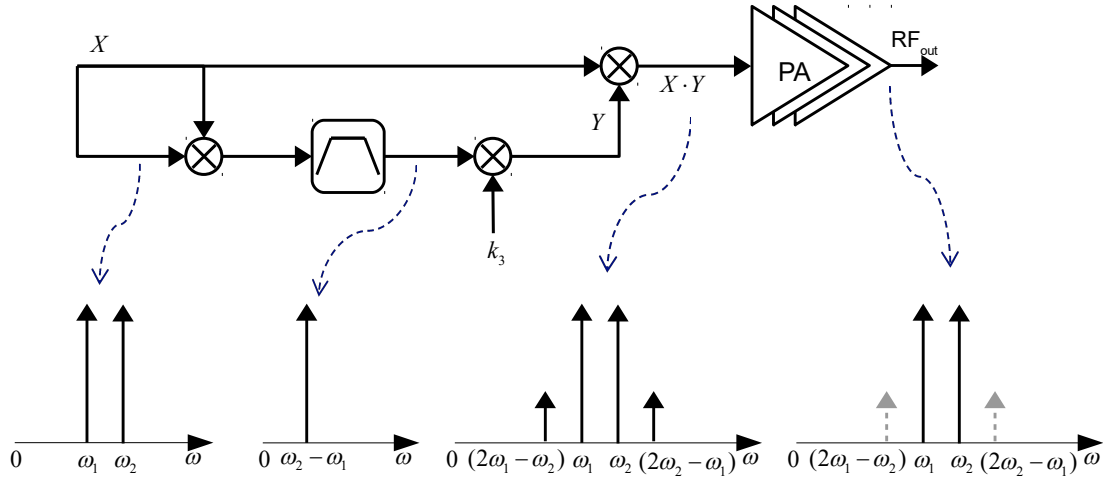
FIGURE A.1: Illustration of predistortion signal generation for the case of two-tone signal.

Appendix B

Digital ASIC Design Methodology

We have used MATLAB driven ASIC design flow, as explained in Section 3.6. The flowchart shown in Fig. B.1 summarizes the steps that were used in obtaining the final digital IC implementation in 28 nm FDSOI CMOS process. Various metrics of the IC such as the power estimation, gate count and area were also obtained.
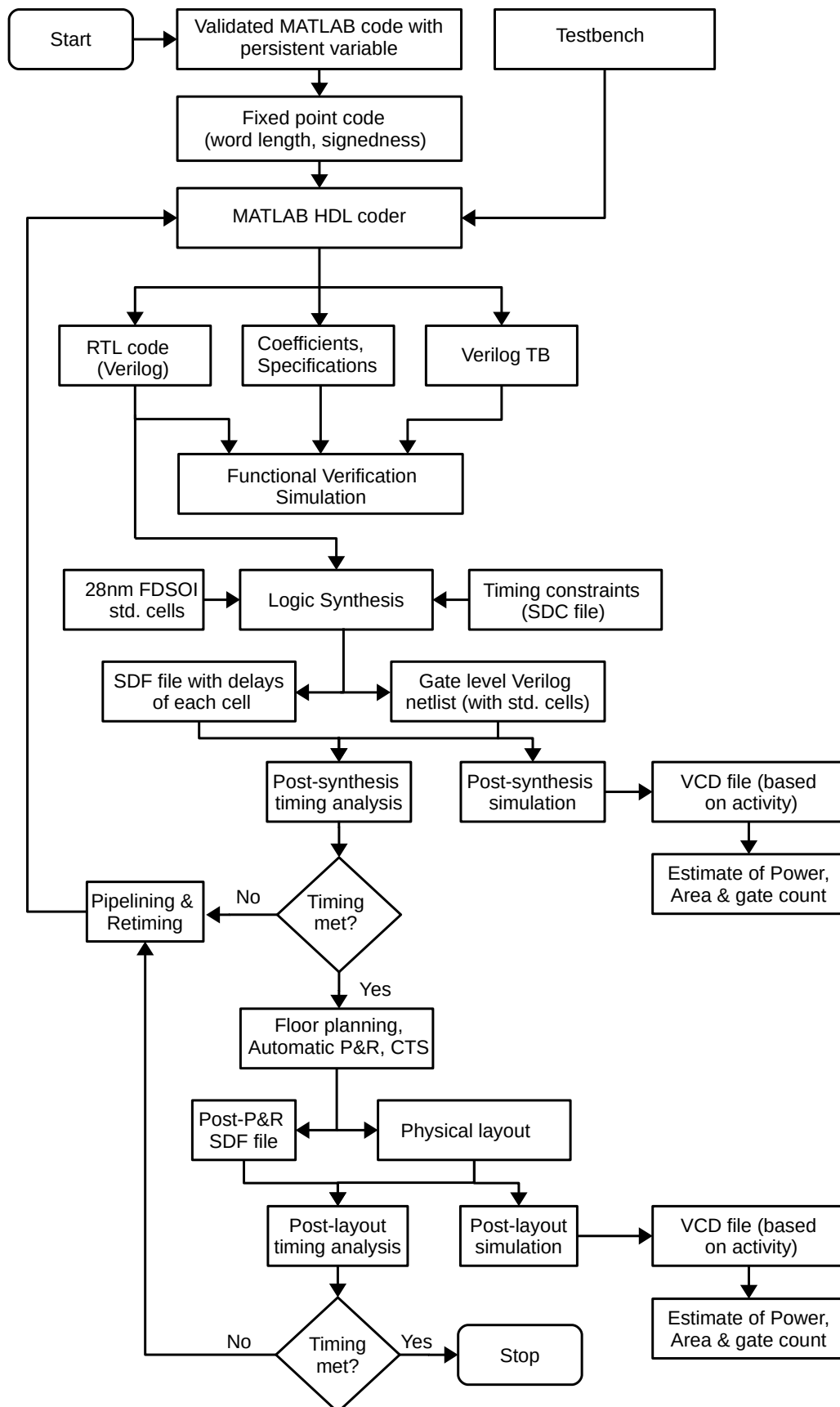
FIGURE B.1: Flowchart describing the digital ASIC implementation Methodology

Appendix C

MATLAB codes

C.1 Example MATLAB code implemented with persistent variables

The following MATLAB code snippet listed below is used implement the FIR filter for the 20 MHz signal case. Persistent variables are used to mimic the registers that are used to implement sample delays in an FIR filter.

```matlab
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function [delayed_xI,delayed_xQ,yI,yQ] = FIRfilter_20MHz(xI,xQ,h0_r,h0_i,h1_r,h1_i)

% declare and initialize the delay registers
persistent ud_xI1 ud_xQ1;
if isempty(ud_xI1) && isempty(ud_xQ1)
    ud_xI1 = 0; ud_xQ1 = 0;
end

% Complex multiplier chain
% yI0 = h0_r*xI - h0_i*xQ;
% yQ0 = h0_i*xI + h0_r*xQ;
%
% yI1 = h1_r*ud_xI1 - h1_i*ud_xQ1;
% yQ1 = h1_i*ud_xI1 + h1_r*ud_xQ1;

[yI0, yQ0] = compMult(xI,xQ,h0_r,h0_i);
[yI1, yQ1] = compMult(ud_xI1,ud_xQ1,h1_r,h1_i);

% delayout input signal
delayed_xI = ud_xI1;
delayed_xQ = ud_xQ1;

% output signal
yI = yI0+yI1;
yQ = yQ0+yQ1;
```

```
% update the delay line
ud_xI1 = xI;
ud_xQ1 = xQ;


end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

C.2   Example MATLAB code generated by the fixed-point designer

The following is the code generated by MATLAB fixed-point designer for the above MATLAB function, which is the 20 MHz FIR filter function written with persistent variable. Signed notation of 14-bit wordLength was employed in the entire datapath.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%                                                                       %
%            Generated by MATLAB 9.3 and Fixed-Point Designer 6.0       %
%                                                                       %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%#codegen
function [delayed_xI,delayed_xQ,yI,yQ] = FIRfilter_20MHz_v3_fixpt(xI,xQ,h0_r,h0_i,h1_r,h1_i)

% declare and initialize the delay registers
fm = get_fimath();

persistent ud_xI1 ud_xQ1;
if isempty(ud_xI1) && isempty(ud_xQ1)
    ud_xI1 = fi(0, 1, 14, 12, fm); ud_xQ1 = fi(0, 1, 14, 13, fm);
end

% Complex multiplier chain
% yI0 = h0_r*xI - h0_i*xQ;
% yQ0 = h0_i*xI + h0_r*xQ;
%
% yI1 = h1_r*ud_xI1 - h1_i*ud_xQ1;
% yQ1 = h1_i*ud_xI1 + h1_r*ud_xQ1;

[fmo_1, fmo_2] = compMult(xI,xQ,h0_r,h0_i);
yI0 = fi(fmo_1, 1, 14, 13, fm);
yQ0 = fi(fmo_2, 1, 14, 13, fm);
[fmo_3, fmo_4] = compMult(ud_xI1,ud_xQ1,h1_r,h1_i);
yI1 = fi(fmo_3, 1, 14, 15, fm);
yQ1 = fi(fmo_4, 1, 14, 15, fm);

% delayout input signal
delayed_xI = fi(ud_xI1, 1, 14, 12, fm);
delayed_xQ = fi(ud_xQ1, 1, 14, 13, fm);

% output signal
```

```matlab
yI = fi(yI0+yI1, 1, 14, 12, fm);
yQ = fi(yQ0+yQ1, 1, 14, 13, fm);

% update the delay line

ud_xI1(:) = xI;
ud_xQ1(:) = xQ;


end

function [mr,mi] = compMult(xI,xQ,hr,hi)

fm = get_fimath();

mr = fi(fi_signed(hr*xI) - hi*xQ, 1, 14, 13, fm);
mi = fi(hi*xI + hr*xQ, 1, 14, 13, fm);
end




function y = fi_signed(a)
    coder.inline( 'always' );
    if isfi( a ) && ~(issigned( a ))
        nt = numerictype( a );
        new_nt = numerictype( 1, nt.WordLength + 1, nt.FractionLength );
        y = fi( a, new_nt, fimath( a ) );
    else
        y = a;
    end
end

function fm = get_fimath()
        fm = fimath('RoundingMethod', 'Floor',...
            'OverflowAction', 'Wrap',...
            'ProductMode','FullPrecision',...
            'MaxProductWordLength', 128,...
            'SumMode','FullPrecision',...
            'MaxSumWordLength', 128);
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

Titre : Prédistorsion Mixte pour des Micro-Cellules 5G

Mots clés : Prédistorsion signaux mixtes, petite cellule, Communications sans fil 5G, Polynôme de mémoire

Résumé : Les stations de base à petite échelle (pico-cellules et femtocellules) seront un des leviers principaux qui permettront d'atteindre l'objectif 1000X, objectif fixé par les grands acteurs du domaine des télécommunications visant à augmenter la capacité des réseaux mobiles sans fil 5G d'un facteur 1000 par rapport aux réseaux 4G. Dans ce type de réseau, l'amplificateur de puissance (PA) est responsable de la majorité de la consommation de puissance de la station de base. Pour minimiser sa consommation de puissance, le PA est polarisé proche de sont point de compression mais avec l'augmentation des largeurs de bande, ce dernier subit des effets de mémoire accrus qui viennent s'ajouter aux problèmes classiques de non-linéarités. Les systèmes de prédistorsion numérique (DPD), et analogique/RF(ARFPD) peuvent être utilisés pour améliorer le compromis linéarité / efficacité des PAs. Cependant pour les pico-cellules et femto-cellules utilisées dans le standard 5G, les prédistorseurs conventionnels ne sont adaptés pour des raisons de complexité et de consommation de puissance.

Le modèle "Memory Polynomilal" (MP) est l'un des modèles de prédistorsion les plus attractifs pour modéliser les PAs, fournissant des performances intéressantes avec peu de coefficients. Cependant, la précision de ce modèle se dégrade pour les signaux large bance. Pour palier ce problème, nous proposons un nouveau modèle, le FIR-MP qui combine un filtre FIR au modèle MP classique. Pour valider et quantifier la précision du modèle proposé, nous avons effectué des simulations avec un modèle extrait par mesure de l'amplificateur sur étagère ADL5606 (GaAs 1W HBT PA). Les résultats de ces simulations présentent des améliorations du taux de fuite des canaux adjacents (ACLR) de 7,2 dB et 15,6 dB, respectivement, pour des signaux à 20 MHz et 80 MHz par rapport au modèle MP classique. Le FIR-MP a été également synthétisé en technologie CMOS FDSOI 28 nm. Les résultats de la synthèse ont donné une puissance globale de $9,18$ mW and $116,2$ mW, respectivement, pour les signaux de 20 MHz and 80 MHz. Basé sur le modèle proposé de FIR-MP, une nouvelle approche à signaux mixtes pour linéariser les PAs a été aussi étudiée. En fait, le filtre numérique FIR améliore la performance de correction de la mémoire sans aucune expansion de la bande passante et la linéarisation en bande de base permet d'éviter l'utilisation de composants RF dans le lineariseur. Ainsi, les contraintes en bande passante requises pour le DAC, les filtres de reconstruction et les blocs RF de l'émetteur sont relâchées comparés aux techniques conventionnelles de linéarisation numériques et RF. Nous avons ainsi étudié l'impact des diverses non-idéalités en utilisant un signal modulé à 80 MHz afin de dériver les exigences pour la mise en uvre du circuit. Les simulations ont montré qu'une résolution de 8 bits pour les coefficients et un SNR de 60 dB sont nécessaires pour atteindre un ACLR1 supérieur à 45 dBc. Ces résultats constituent un premier signe favorable dans l'optique d'une implémentation matérielle de la solution proposée, étape indispensable pour évaluer précisément sa consommation de puissance et sa complexité pour pouvoir la comparer à l'état de l'art des linéariseurs.

Title : Mixed-Signal Predistortion for Small-Cell 5G Wireless Nodes

Keywords : Mixed-signal Predistortion, Small-cell, 5G wireless communications, Memory polynomial

Abstract : Small-cell base stations (picocells and femtocells) handling high bandwidths (> 100 MHz) will play a vital role in realizing the 1000X network capacity objective of the future 5G wireless networks. Power Amplifier (PA) consumes the majority of the base station power, whose linearity comes at the cost of efficiency. With the increase in bandwidths, PA also suffers from increased memory effects. Digital predistortion (DPD) and analog RF predistortion (ARFPD) tries to solve the linearity/efficiency trade-off. In the context of 5G small-cell base stations, the use of conventional predistorters becomes prohibitively power-hungry. Memory polynomial (MP) model is one of the most attractive predistortion models, providing significant performance with very few coefficients. We propose a novel FIR memory polynomial (FIR-MP) model which significantly augments the performance of the conventional memory polynomial predistorter. Simulations with models extracted on ADL5606 which is a 1 W GaAs HBT PA show improvements in adjacent channel leakage ratio (ACLR) of 7.2 dB and 15.6 dB, respectively, for 20 MHz and 80 MHz signals, in comparison with MP predistorter. Digital implementation of the proposed FIR-MP model has been carried out in 28 nm FDSOI CMOS technology. With a fraction of the power and die area of that of the MP a huge improvement in ACLR is attained. An overall estimated power consumption of 9.18 mW and 116.2 mW, respectively, for 20 MHz and 80 MHz signals is obtained.

Based on the proposed FIR-MP model a novel low-power mixed-signal approach to linearize RF power amplifiers (PAs) is presented. The digital FIR filter improves the memory correction performance without any bandwidth expansion and the MP predistorter in analog baseband provides superior linearization. MSPD avoids 5X bandwidth requirement for the DAC and reconstruction filters of the transmitter and the power-hungry RF components when compared to DPD and ARFPD, respectively. The impact of various non-idealities is simulated with ADL5606 (1 W GaAs HBT PA) MP PA model using 80 MHz modulated signal to derive the requirements for the integrated circuit implementation. A resolution of 8 bits for the coefficients and a signal path SNR of 60 dB is required to achieve ACLR1 above 45 dBc, with as little as 9 coefficients in the analog domain. Discussion on the potential circuit architectures of subsystems is provided. It results that an analog implementation is feasible. It will be worth in the future to continue the design of this architecture up to a silicon prototype to evaluate its performance and power consumption.