Behavioral Modeling and Digital Predistortion of Wide- and Multi-Band Transmitter Systems

by

Farouk Mkadem

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2014

© Farouk Mkadem 2014

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Farouk Mkadem

Abstract

The demands for high data rates and ubiquitous/broadband wireless access necessitate the development of radio systems that deploy wide- and multi-band signals. These signals lead to high spectral utilization with negative trade-offs, such as rapidly varying envelopes and high peak-to-average power ratios (PAPRs). To deploy these types of signals, radio frequency (RF) transmitters face several challenges in power consumption, i.e., efficiency, and sources of distortions, i.e., nonlinearity. These challenges are most apparent in power amplifier (PAs) and degrade the overall performances of RF transmitters.

PA efficiency and linearity are characteristics that cannot be satisfied simultaneously. At high input power, PAs exhibits high efficiency; however, a PA in that region is inherently nonlinear. Conversely, at low input power, PAs are fairly linear at the expense of efficiency. With the deployment of wide- and multi-band signals, PAs exhibit strong static nonlinearity and memory effects at high input power. These effects lead to distortions, resulting in degradation of the error vector magnitude (EVM) and spectral regrowth. This regrowth creates adjacent channel interference and violates the emission requirements mandated by regulatory bodies. Digital predistortion (DPD) has been devised to mitigate the PA nonlinearity at high input power. Subsequently, DPD improves the achievable PA linearity versus power efficiency trade-off.

DPD incorporates an extra nonlinear function before the PA in order to preprocess the input signal. As a result, the cascaded system (DPD+PA) behaves linearly. DPD+PA linearity requires the DPD function to produce nonlinearities that have equal magnitude and are out phased compared to those generated by the PA. Thus, accurate PA behavioral modeling is essential for the development of DPD and is usually explored before DPD development.

In the context of wide- and multi-band signals, PA behavioral models/DPD schemes face several problems that have not been addressed appropriately in the literature. One particular problem is the exponential growth of the number of coefficients with nonlinearity order and memory depth. This growth leads to models' high complexity identification and implementation. Therefore, it restricts DPD performances in trading off linearity and efficiency. Moreover, for multi-band signals, the carrier frequencies' separation can be very large (in the order of hundreds of MHz). Consequently, the conventional single-input single-output DPD is not viable due to the unrealistic sampling rate required to cover a large range of frequencies. Finally, because a very high sampling rate is needed in the PA output observation path, the practicability of the DPD is rendered more complex and sometimes unfeasible when it is deploying wideband signals.

This thesis explores new schemes suitable for modeling and linearizing PA outputs when driven with wide- and multi-band signals. First, a new pruning technique is introduced to alleviate the complexity of DPD implementation. This pruning identifies the minimum set of dominant kernels needed in the Volterra series modeling scheme based on Wiener G-Functionals. The pruned Volterra series allows significant improvements of the numerical computation and stability of the DPD scheme. Second, new generalized memory polynomial (GMP) models for multi-band DPD are proposed. These multi-input multi-output GMP models show excellent potential in linearizing different nonlinear multi-band PAs. In addition, they involve cross terms and a reduced number of coefficients, which allows robustness against time delay misalignment between the multi-band signals. Finally, a new method to reduce the conventional high output sampling rate in the PA output observation path is proposed. Detailed analysis assesses the extent to which the output-sampling rate can be reduced through careful selection of the set of kernels used to represent the Volterra series DPD. This analysis results in a $^2/_5$ reduction when the G-Functionals pruning technique is deployed. Extensive analysis and measurement validation are discussed and carefully analyzed to prove the validity of the different proposed solutions and to guarantee their generalizability.

Acknowledgements

I want to express my gratitude to the following people for their help over the period of my PhD and MASc degrees. Their continuous support and concern have made the achievement of this work possible.

First and foremost, I would like to acknowledge my supervisor Dr. Slim Boumaiza whose guidance and patience guided my work and assisted me in writing this thesis. He ensured my academic development while emphasizing the importance of collaborative work with fellow students.

I would also like to thank all emerging radio systems group for their continued involvement and help. They were my friend and my family in Canada, and I enjoyed collaborating with them.

I would also like to express my gratitude to all my professors at the University of Waterloo and the Tunisia Polytechnic School for every bit of inspiration and insight they caused.

Last but not least, I don't find suitable words to express my love, gratitude and respect for those to whom I am offering this work: my dear beloved family. They were, are and will be forever my inspiration and my strength for any accomplishment and success.

To Amor
My father, my friend and my hero
You will be always remembered

"A happy family is but an earlier heaven"
George Bernard Shaw

Contents

List of Tables

List of Figures

Nomenclature

f_s	sampling rate
f_c	carrier frequency
t_s	time step
1D-GMP	single-band generalized memory polynomial
2D-CR-GMP	dual-band complexity reduced generalized memory polynomial
2D-DPD	dual-input digital predistortion
2D-GMP	dual-band generalized memory polynomial
3D-CR-GMP	tri-band complexity reduced generalized memory polynomial
3D-DPD	tri-band digital predistortion
3D-GMP	tri-band generalized memory polynomial
3G	third generation wireless network technology
4G	fourth generation wireless network technology
ACLR	adjacent channel leakage ratio
ACPR	adjacent channel power ratio
ADC	analog-to-digital converter
ADS	advanced design system
AM/AM	amplitude-to-amplitude distortion
AM/PM	amplitude-to-phase distortion
ANN	artificial neural network

BER	bit error rate
BS	base station
CAD	computer aided design
CR-GMP	complexity-reduced generalized memory polynomial
CW	continuous wave
DAC	digital-to-analog converter
dB	decibels
dBc	decibels relative to the carrier
DC	direct current
DDR	dynamic deviation reduction
DE	drain efficiency
DPD	digital predistortion
DSP	digital signal processor
DUT	device under test
EPA	error power amplifier
EVM	error vector magnitude
FET	field-effect transistor
FIR	finite impulse response
FPGA	field programmable gate array
GaN	gallium nitride
GHz	giga Hertz
GMP	generalized memory polynomial
GST	generalized sampling theorem
HSPA	high-speed packet access

IF	intermediate frequency
IMD	intermodulation distortion
IMD3	third-order intermodulation distortion
IMD5	fifth order intermodulation distortion
IMD7	seven order intermodulation distortion
INR	input Nyquist rate
km/h	kilometres per hour
LDMOS	laterally diffused metal oxide semiconductor
LNA	low noise amplifier
LO	local oscillator
LSE	least squares error
LTE	long term evolution
Mbit/s	megabit per second
MHz	megahertz
MIMO	multi-input multi-output
NMSE	normalized mean square error
ns	nanoseconds
ONR	output Nyquist rate
PA	power amplifier
PAE	power added efficiency
PAPR	peak-to-average-power ratio
PSD	power spectrum density
PSDE	power spectrum density error
RF	radio frequency

RLS	recursive least squares
SISO	single-input single-output
SNR	signal-to-noise ratio
SSG	small signal gain
TOR	transmitter observation receiver
WiMAX	worldwide interoperability for microwave access

Chapter 1

Introduction

The information and communications technology sector has been continuously growing to fulfill increasingly evolving applications. An exponential growth of mobile cellular telephones and broadband services subscriptions has been recorded over the past ten years, according to the international telecommunication union [33]. This rapid growth is being driven by wireless network usage evolution from voice centric to broadband Internet access and video transmission.

Third generation wireless-network technology (3G)'s ongoing deployment has provided a partial solution for mobile broadband Internet access. Certainly, 3G has enabled the deployment of high-speed mobile networks, such as worldwide interoperability for microwave access (WiMAX) and high-speed packet access (HSPA). These networks allow people to connect to the Internet, anytime, anywhere. Nevertheless, 3G do not offer a complete solution for tomorrow's wireless networks as was expected by international telecommunication union and other standard-setting agencies around the world. Thus, fourth-generation wireless network technology (4G) has been developed and should eventually provide:

- **Similar data throughput as wired broadband technology:** 4G should offer speed comparable to that of wired broadband technologies along with high mobility features [33].
- **Significantly lowered network latencies:** Broadband networks suffer from a long time delay between when the source sends a packet and when the destination receives it. Accordingly, 4G is promising to shorten this delay.
- **Improved spectral efficiencies:** 4G is promising new wide-/multi-band modulation schemes to increase the spectrum usage efficiency. Thus, 4G will enable more data throughput within the same 3G and anterior standards bands [3].

Table 1.1: Fourth-generation wireless network technology features

Features	**Values**
Signal bandwidth	Up to 100 MHz
C-plane capacity	Up to 300 active data clients in every 5 MHz cell
Peak to average power ratio	Up to 12 dB
Peak downlink speed	Up to 299.6 Mbit/s
Peak uplink speed	Up to 75.4 Mbit/s
Low data transfer latencies	$<$ 50 ns from idle
Mobility	Up to 500 km/h depending on the frequency band

Table 1.1 shows the main features that 4G was conceived to provide. Among others, 4G networks should allow an instantaneous bandwidth of up to 100 megahertz (MHz), thereby enabling a downlink speed of about 300 megabits per second (Mbit/s). This reduces the latency time to less than 50 nanoseconds (ns) and offers a high mobility operation for up to 500 kilometres per hour (km/h) (network accessible in operating cars, trains, etc.). To accommodate such features, the wireless network is expected to involve:

- **Wideband signals:** According to Cooper's law [22], the spectrum efficiency doubles every 30 months. However, the user demand for data capacity increases at a much higher speed [2]. Therefore, wider bandwidth signals are required to meet the demand for broadband data access that cannot be satisfied solely through spectrum efficiency improvement.
- **Multi-band signals:** The wireless spectrum usage between 100 MHz and 5 GHz shows no unallocated continuous 100 MHz band [95]. Alternatively multiple carriers' aggregation is envisaged. Moreover, a wireless access provider is expected to deploy different 4G and/or anterior wireless standards simultaneously to meet various users' needs. Conventionally, the stacking of different radios was conceived to enable various standards. However, radios are currently expected to handle these various standards using the same hardware by enabling multi-band operation [52].

1.1 Wireless Communication Challenges

Emerging wireless networks were conceived to provide ubiquitous, fast and reliable wireless access. Thus, they require wide-/multi-band signals, adaptive modulation and multi-input

multi-output (MIMO) transmitters [13]. These requirements make radio frequency (RF) transmitters implementation very complex and create the major challenges discussed in this section.

The non-availability of a continuous wideband spectrum is the main challenge for enabling 4G networks. Unfortunately, most of the wireless spectrum opportunities are currently either bought or reserved [33]. Moreover, there is no continuous wideband signal allocation within the spectrum dedicated to mobile communication. Alternatively, multiple carrier aggregation enables the existence of wide frequency bands [60]. This technique is motivated by discrete spectrum bands' availability, because they can be combined to support high bandwidth requirements. Moreover, RF transmitters should deploy simultaneously different modulation schemes for different carriers and different signal bandwidths due to back compatibility requirement. In this way, the network results in deploying wide-/multi-band signals that increase the signal peak-to-average-power ratio (PAPR). For example, 4G networks PAPR can reach up to 12 decibels (dB), as shown in Table 1.1. Unfortunately, this PAPR growth reduces the signal quality, i.e., linearity. Therefore, RF transmitters are operated in the back-off region far from their peak power capabilities to avoid strong nonlinearity. This nonlinearity generates high error vector magnitude (EVM) and spectrum regrowth that interferes with adjacent channels. Unfortunately, at the back-off region, the RF transmitters experience poor efficiency.

The design of wide-/multi-band ultra linear RF transmitters becomes increasingly complex when combined with the requirement for high power efficiency. Cascades of different hardware optimized for each standard can be deployed in order to enable multi-band standards. However, this solution suffers from two drawbacks. First, the cost of the network increases due to the deployment of different hardware. Second, the size of the resulting cascaded RF systems becomes large. This is contradictory to wireless communication evolution toward remote radio heads [57], where the size and the cost of the RF system should be small enough to be mobile and easily installed.

Current state-of-the-art RF transmitters are only capable of coping with single-band scenarios with optimized linearity and efficiency, or wide-/multi-band scenarios with lower requirements for linearity and efficiency. In the literature, the efficiency and linearity of RF transmitters have only been optimized for narrow bandwidth signals, typically in the range of $1.4 - 20\ MHz$ [20, 48]. These types of RF transmitters are not suitable for 4G networks, i.e., when using wide-/multi-band signals. Furthermore, their performance drops dramatically when they are operated at regions different from where they were optimized. Alternatively, linearity and efficiency are compromised to satisfy wide-/multi-band operation. These RF transmitters are also not suitable for a 4G network, as these transmitter dramatically increase its cost and complicate its implementation and mobility.

Among the different constituents of RF transmitters, the power amplifier (PA) inherits a large portion of the previously enumerated challenges. Indeed, the RF transmitters' power consumption (efficiency) and the sources of distortions (nonlinearity) are apparent in the PA. Unfortunately, PA efficiency and linearity are characteristics that cannot be satisfied simultaneously [24]. At high input power, the PA exhibits high efficiency; however, the PA in that region is inherently nonlinear. Conversely, at low input power, the PA is fairly linear, whereas its efficiency is very poor. Furthermore, PA nonlinearity is more pronounced in the case of wide-/multi-band signals, where linearity and efficiency are compromised to ensure wide-/multi-band operation.

One trivial solution for tackling the PA nonlinearity is the use of brute force to operate the PA deeply in the back-off region far from its power capability. This approach guarantees that the PA output signal will pass the mandatory spectrum mask and the EVM requirements [1]. Despite its simplicity, this solution is increasingly discarded. In fact, it leads to costly and power inefficient RF transmitters, because of the poor PA efficiency in low input power regions. Alternatively, several linearization techniques are being devised to mitigate the PA nonlinearity at high input power, consequently improving the achievable PA linearity versus power efficiency trade-off. Table 1.2 illustrates typical performances of the major linearization techniques, namely feedback, feed-forward [29, 62], and digital predistortion (DPD) [36, 45, 87]. Among these techniques, DPD incorporates an extra nonlinear function before the PA to preprocess the input signal. This results in a cascaded system (DPD+PA) that behaves linearly. DPD+PA overall linearity requires primarily the DPD function to produce nonlinearities that have equal magnitude and are out of phase with those generated by the PA. DPD is the focus of this thesis since it offers higher efficiency and greater flexibility at a lower cost.

Table 1.2: Comparison of different linearization techniques

Technique	Correction	Bandwidth	Efficiency	Flexibility	Cost
Feedback	High	Narrow	High	Moderate	Low
Feed-Forward	High	High	Low	Moderate	High
Digital Predistortion	Moderate	Moderate	High	High	Low

The main DPD attribute is its principle simplicity. Actually, the predistortion function can be generated by modeling the inverse behavior of the PA, as shown in Figure 1.1 and explained in the following [84]. First, the output signal is attenuated by small signal gain (SSG); the DPD is a nonlinear function that does not change the average power of the signal being transmitted. Then, the PA input and output signal samples are recorded.

Finally, the DPD is constructed as a dynamic nonlinear mathematical function that models some PA input signal samples as a function of the output one; i.e., it is an inverse model. Thus, a key point of DPD is the construction of an accurate PA inverse behavioral model. Consequently, building a good DPD linearizer depends on the ability to develop an accurate and simple PA behavioral model.

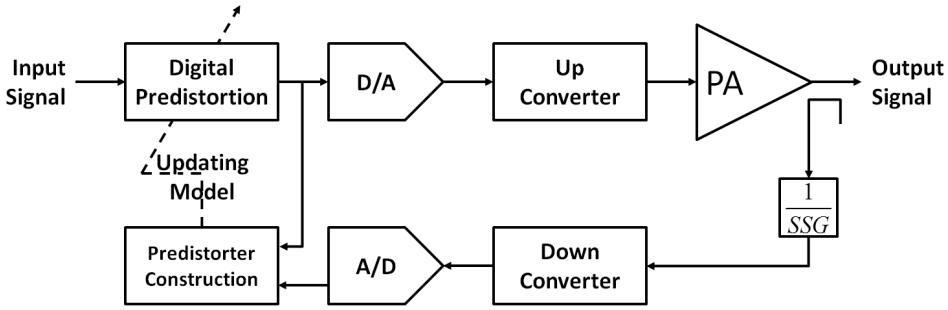

Figure 1.1: Digital predistortion model generation principle

DPD practical implementation and design face several challenges in the context of wide-/multi-band signals. These problems extend from the three different components of the DPD solution, as shown in Figure 1.2, namely:

1. the DPD engine,
2. the DPD coefficients identification, and
3. the transmitter's observation receiver path.

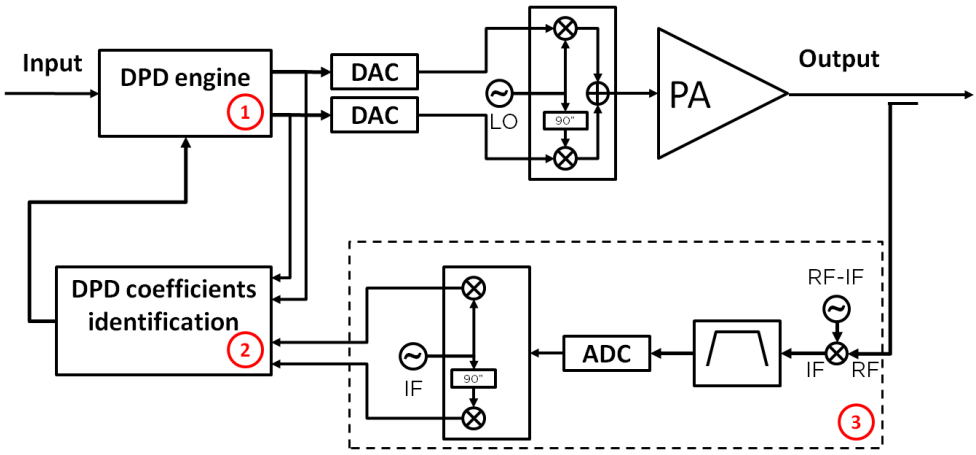

Figure 1.2: Digital predistortion block diagram

Several DPD engines were proposed in the literature. They vary from simple look-up table [18, 23, 56] to more complex structures such as Volterra series [31, 40, 59, 64, 118, 120–122]. The Volterra series is a universal and excellent framework for modeling nonlinear

dynamic systems [93]. However, one particular problem of the Volterra series is the exponential growth of the number of coefficients with high nonlinearity and memory depth orders. Although it has been proven that the Volterra series formulation is convergent for fading memory systems [97], the number of coefficients needed for practical PA degree of nonlinearity and memory depth is still large. For example, a Volterra series approximation of a polynomial degree of 7 with 5 memory taps contains some 8,364 coefficients that need to be determined; if only odd orders were involved. Historically, this was an intractable problem and Volterra series approximations were limited to low degrees and only a couple of memory taps. In addition, due to the large degree of freedom in the Volterra series formulation, a large number of parameters should be determined/estimated before determining the model coefficients; namely, the nonlinearity order and the memory span for each order. Traditionally, these parameters are estimated using a trial and evaluation technique leading to not necessarily the optimum selection of these parameters.

In the context of multi-band scenarios, the conventional implementation of the DPD faces several issues. Mainly, the carrier frequencies separation can be very large (in the order of hundreds of MHz), and a single-input single-output (SISO) DPD is not viable due to the unrealistic sampling rate required to cover a large range of frequencies. Moreover, a SISO Volterra model will require unrealistic number of coefficients due to high memory taps needed if the multi-band signal is combined and treated as one signal. If fact, the memory span of the PA is an intrinsic property, i.e., independent of the signal being used to drive the PA. Therefore, combining a multi-band signal to one single signal will lead to an increase of the required memory taps to model the memory span. Since the Volterra series number of coefficients increases exponentially with the memory depth, using a SISO model for a multi-band DPD will results in a huge number of coefficients that can not be identified neither in real time systems of in computer aided design (CAD) tools. Thus, MIMO behavioral models/DPDs are envisaged.

In the literature, multi-band DPD models were developed using two strategies; generalization of SISO model structure to accommodate multi-band signals [8, 88, 107, 112, 113] or pruning of a general Volterra series formulation [30]. These approaches lead to different challenges. Either, large number of coefficients results as the construction of the DPD model is not specific to the multi-band PA operation or extra complexity due to pruning needed to accommodate the general Volterra series formulation to the multi-band case.

Deployment of wide-/multi-band signals require a wideband transmitter observation receiver (TOR) with high-speed analog-to-digital converters (ADCs) and flat frequency response over a wide range of frequencies. This is particularly important for capturing accurate measurement data essential for the identification of the DPD coefficients. The significant power and cost overheads of high-speed ADCs brings down the overall efficiency

of the DPD+PA cascade [49]. Thus, it limits the usefulness of the DPD approach to enhancing the trade-off between efficiency and linearity. Indeed, it is widely accepted that the application of DPD to PA linearization requires the capture of the PA output signal using a TOR that has a bandwidth equal to at least five times the input signal modulation bandwidth. As this latter widens, this requirement becomes less manageable. For example, for a long term evolution (LTE)-advanced signal where the instantaneous bandwidth reaches 100 MHz, the DPD needs to deploy an ADC with a speed equal at least twice 500 MHz to follow the Nyquist-Shannon sampling theorem. This bandwidth is even more challenging in the case of multi-band signals, where large separation of the frequency carriers requires an unrealistic sampling speed for ADC to sample the entire multi-band signal at once. This requires an RF measurement systems to include high-speed ADCs, e.g., at least 1 giga Hertz (GHz). Such systems present major challenges, as their signal-to-noise ratio (SNR) is very high. In addition, for a linear and accurate measurement of the output signal, accurate calibration of the transmitter, i.e., in-phase/quadrature-phase (I/Q) imbalance correction, and accurate calibration of the receiver, i.e., amplitude and frequency correction, should be performed across a large range of frequency bands [7, 17, 102]. The cost of DPD systems using high TOR bandwidth when handling wide-/multi-band signals, due to the complex models and high-speed ADCs, limit the efficacy of the DPD solution as a good linearizer in efficiency and linearity trade-off. In fact, the power consumption of the high-speed ADCs required for wide-/multi-band DPD systems is no longer negligible [49] and significantly affects the overall efficiency of the DPD+PA cascade. Furthermore, the large size memory and processor needed for the computation of the complex behavioral models require a large cooling system, thereby increasing the size and cost of the RF system. This compromises the evolution toward remote radio heads [58], where the aim is a small and simple architecture for wireless networks.

Several algorithms and approaches were developed in the literature for DPD coefficients identification. The direct and the indirect learning were both used for the identification of the DPD coefficients [84]. While the direct learning identifies first a PA model and then a DPD model is estimated [117], the indirect learning approximates the DPD function to the inverse model of the PA [91]. These learning identification methods were used solely and combined for better enhancement [39]. The indirect learning showed broader application and success for the identification of the DPD coefficients and will be adopted in this thesis. Different algorithms were developed and used for the indirect learning, namely least squares error (LSE), recursive least squares (RLS), and intelligent algorithms [9, 19, 27, 47, 71, 108, 116].

1.2 Thesis Objectives, Statements and Organization

From the discussion in Section 1.1, development of new DPD engines and pruning technique with reduced complexity are required for PA driven with wide-/multi-band signals. In addition, reduced TOR bandwidth is seek without compromising the modeling/linearization performances. The objective of this thesis is to tackle these limitations as summarized in the following statements:

- **Systematic pruning of Volterra series:** a new pruning approach based on Wiener G-Functionals is sought for the selection of reduced number of dominant kernels in the Volterra series scheme.
- **DPD with reduced TOR bandwidth:** a reduced TOR bandwidth solution is proposed and validate for wideband PA behavioral/DPD models.
- **Multi-band complexity reduced DPD:** a new strategy to construct SISO and MIMO PA behavioral/DPD models is proposed that takes into account PA behavior driven with single- and multi-input excitations.

This thesis is organized as follows.

Chapter 2 introduces PA nonlinearity and efficiency dilemma. Then, literature review of different linearization techniques is proposed with a special focus on DPD.

Chapter 3 expounds on the pruning of single-band DPD model using G-Functionals [73, 75]. This pruning approach starts with dominant kernels identification in the Volterra series scheme. Then, the pruned Volterra series is applied to synthesize a DPD function. The resulting model reduces the kernels' number and its range of values, and produces better numerical conditioning when compared to the full Volterra series.

Chapter 4 investigates the reduction of the TOR's bandwidth in DPD system [67]. It starts by demonstrating that the general sampling theorem (GST) [123] is not applicable for synthesizing a DPD function. Alternatively a novel technique for identifying a DPD is proposed using a reduced TOR bandwidth. The theoretical basis of the proposed technique is presented. Then, a detailed analysis assesses the bandwidth reduction sensitivity extent to the kernels set selected to represent the Volterra series DPD. This analysis reveals a robust DPD synthesis when the TOR bandwidth is reduced from the conventional five times to twice the input signal bandwidth. The G-Functionals pruning technique, proposed in Chapter 3, is used to determine a low pass equivalent Volterra model that identifies the most suitable structure needed for a given PA under test and signal characteristics. Finally,

the linearized PA and the GST are used for PA behavioral modeling using a reduced TOR bandwidth.

Chapters 5 and 6 expound a generalized memory polynomial (GMP) models for single-/multi-band DPD [68, 74]. First, PA block diagrams characterizing the behavior of a PA under single-/multi-band stimulus are proposed. Second, GMP forward models are derived from the feedback PA block diagrams for the case of single-band, dual-band, and tri-band PAs. A general formulation for the multi-band case is then deduced. Finally, it will be shown that the proposed GMP models are robust against time delay misalignment between the different multi-band signals.

Lastly, Chapter 7 concludes the thesis by summarizing the contributions and the achieved results. Future works and possible improvement are discussed.

Chapter 2

Power Amplifier Behavioral Modeling and Linearization Techniques

This chapter gives a literature review of wide-/multi-band PA behavioral modeling/DPD. It starts with an introduction to PA efficiency, nonlinearity, memory effects, and linearization techniques. Next, wide-/multi-band PA behavioral models/DPD schemes latest development are presented. Throughout this chapter, the proposed solutions are introduced and discussed to distinguish their novelties and the problems that they tackled.

2.1 Power Amplifier Overview

In this section, wireless radio system architecture is described. Then, PA nonlinearity and efficiency characterization is discussed. Finally, the different linearization techniques are introduced with a special focus on DPD.

Figure 2.1 shows a typical wireless radio system. It is mainly composed of three sub-blocks: the transmitter, the air, and the receiver.

Receiver: the receiver is responsible for capturing an RF signal and down-converting it to digital. The receiver is composed by:

- An antenna that captures the analog signal.
- A band-pass filter to reject all the unwanted signals, i.e., adjacent and harmonic signals.
- Low-noise amplifiers (LNAs) are used in lieu of PAs to amplify the filtered signal to a suitable power level for an optimal demodulator operation.

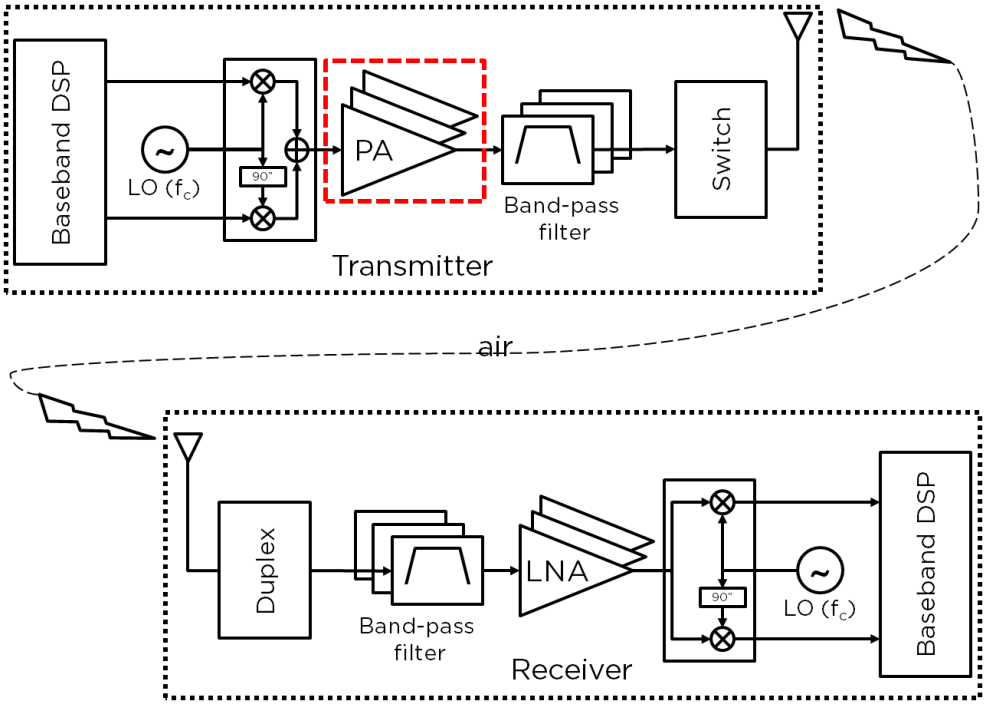

Figure 2.1: Wireless radio system architecture

- A demodulator that down-converts and transforms the RF signal into in-phase (I) and quadrature-phase (Q) baseband components.
- Digital signal processor (DSP) that extracts the modulation scheme from the I and Q components.

Transmitter: the transmitter is responsible for generating a digital signal and converts it to RF. The transmitter is identical to the receiver with two differences:

- A modulator that combines the I and Q baseband components with a local oscillator (LO) into a specific RF signal called the carrier frequency (f_c).
- PAs: the signal at the output of the modulator is at low power. Thus it cannot propagate in the air and reach the receiver before dying due to air attenuation. Thus, a power amplification stage is considered to amplify the RF signal up to a certain high output power given by the regulatory bodies. For traditional wireless radio systems, the transmitter contains as many PAs as there are standards deployed. Thus, in the transmitter, different PAs are stacked in parallel after the modulator. In a base station (BS), the output power of the signal should be very high. Thus, the amplification stage is performed in two steps, where the first amplifier is referred to the driver and the second one is the main PA. For modern wireless system, wide-/multi-band PA are used instead of multiple PA

to reduce the size and cost of the system and enhance its performances.

The transmission from the baseband DSP to its receiver counterpart is characterized by two main factors:

- **Linearity** which determines the quality of the signal transmission. The linearity is measured using EVM or the bit error rate (BER) and adjacent channel power ratio (ACPR).
- **Efficiency** which characterizes the amount of power needed to transfer a digital signal to an RF one.

In this thesis, only the efficiency and linearity of the transmitter are investigated. The PA is the major contributor to nonlinearity and power consumption in the transmitter. In the next subsections, PAs nonlinearity and efficiency are discussed.

2.1.1 Power Amplifier Efficiency

PA efficiency characterizes the amount of input power needed to generate the output RF signal. It can be evaluated using two figures of merit:

- Drain efficiency (DE), η, gives the ratio of the RF output signal power to the direct current (DC) supplied power, as expressed in (2.1):

$$\eta = \frac{P_{out}}{P_{DC}} \tag{2.1}$$

  where P_{out} and P_{DC} designate the RF PA output power and the supplied DC power, respectively.

- Power added efficiency (PAE), expressed in (2.2), involves the PA input signal power. Consequently, the efficiency is characterized in converting all the input supplied powers to the RF output one.

$$PAE = \frac{P_{out} - P_{in}}{P_{DC}} = \eta \left(1 - \frac{1}{G}\right) \tag{2.2}$$

  where P_{in} is the RF PA input power and G is its gain.

Figure 2.2 illustrates typical DE and gain power curves versus the PA input power. It is clear from Figure 2.2 that the efficiency decreases rapidly as the input power back-off increases. However, at the high input power region, the PA output gain drifts from the ideal linear amplification as a consequence of PA nonlinearity. Therefore, improving PA efficiency comes at the expense of its linearity and vice versa [24]. In the next subsection, PA nonlinearity is discussed.

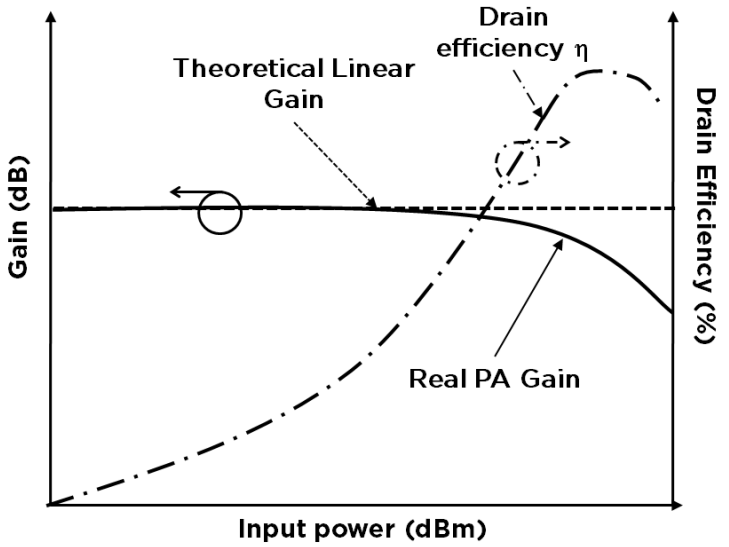

Figure 2.2: Power amplifier drain efficiency and linearity

2.1.2 Power Amplifier Nonlinearity

PA nonlinearity can be characterized when using different input signals, namely continuous wave (CW), two-tone, and modulated signals.

2.1.2.1 Power Amplifier Nonlinearity Characterization With Continuous Wave

A PA under CW stimulus causes two types of distortions, namely amplitude and phase distortions:

- Amplitude distortion is usually expressed as an output amplitude power level change as a function of the input amplitude power level. This is known as amplitude-to-amplitude distortion (AM/AM) and is given in (2.3):

$$r_{out}(t) = g(r_{in}(t)) \tag{2.3}$$

  where r_{in} and r_{out} are the PA input and output amplitude signals, respectively, and $g(.)$ gives a nonlinear function. Sometimes, the AM/AM characteristic is plotted as a gain compression curve when using (2.4).

$$G(t) = \frac{r_{out}(t)}{r_{in}(t)} = \frac{g(r_{in}(t))}{r_{in}(t)} \tag{2.4}$$

- Output amplitude distortion in a nonlinear PA, i.e., AM/AM, is usually accompanied with a phase distortion, i.e., amplitude-to-phase distortion (AM/PM). The AM/PM corresponds to a non-constant PA output signal phase shift when the PA input envelope signal changes. AM/PM is expressed in (2.5):

$$\phi_{out}(t) - \phi_{in}(t) = f(r(t)) \tag{2.5}$$

  where ϕ_{in}, ϕ_{out} and $r(t)$ are the PA input and output phases and the PA input amplitude signal, respectively; and, $f(.)$ gives a nonlinear function.

Figures 2.3 and 2.4 illustrate a typical Gain and AM/PM characteristics, respectively. The CW characterization was a first attempt to study the PA nonlinearity. However, very limited information is extracted from this test, i.e., AM/AM and AM/PM. This was not enough to characterize signal regrowth and memory effects. Thus, a two-tone test was introduced to provide a more detailed PA nonlinearity characterization, as discussed in the next subsection.

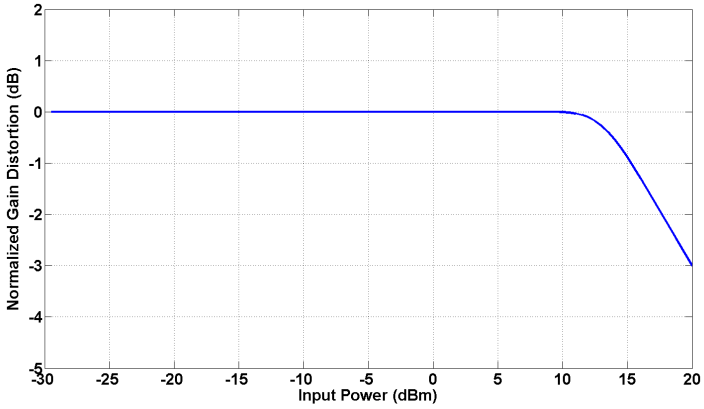

Figure 2.3: Gain compression of a typical nonlinear power amplifier driven with continuous wave

2.1.2.2 Power Amplifier Nonlinearity Characterization With a Two-Tone Test

The two-tone test provides a good illustration of the amplitude and phase distortions discussed previously. The test consists of stimulating a nonlinear PA driven with two sinusoids at different carrier frequencies, f_1 and f_2. The frequency separation $|f_2 - f_1|$ should be negligible compared to carrier $f_c = (f_1 + f_2)/2$; i.e., $|f_2 - f_1| \ll f_c$. The two-tone test representation allows a good understanding of the in-band, out-of-band, intermodulation, and amplitude and phase distortions. An example of a two-tone test is

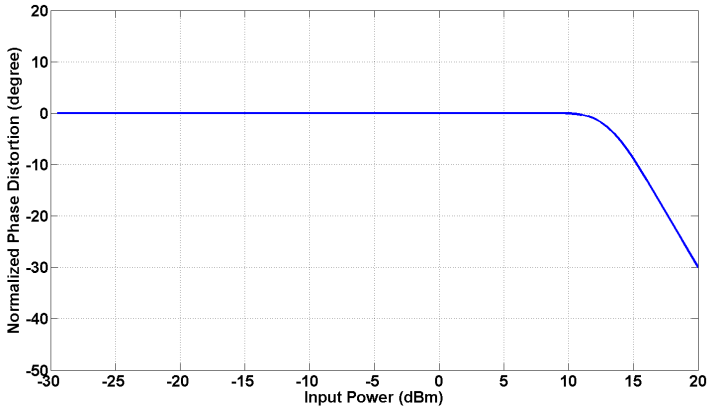

Figure 2.4: Phase distortion of a typical nonlinear power amplifier driven with continuous wave

when using a PA represented by a third-degree memoryless polynomial, as shown in (2.6), where $x(t)$ and $y(t)$ represent the PA input and output signals, respectively; a_1 describes the linear SSG; and a_2 and a_3 are the quadratic and cubic nonlinearity coefficients, respectively.

$$y\left(t\right) = a_1 x\left(t\right) + a_2 x^2\left(t\right) + a_3 x^3\left(t\right) \tag{2.6}$$

Considering the following two tones $v_1(t) = Asin(w_1 t)$ and $v_2(t) = Asin(w_2 t)$, where A is the amplitude of each tone. The PA output driven with $v_1(t)$ and $v_2(t)$, is as follows:

$$\begin{aligned} y(t) =& a_2 A^2 \left[1 + \cos\Big(\left(w_1 - w_2\right) t\Big)\right] \\ &+ a_1 A \left[\sin\left(w_1 t\right) + \sin\left(w_2 t\right)\right] + \frac{9}{4} a_3 A^3 \left[\sin\left(w_1 t\right) + \sin\left(w_2 t\right)\right] \\ &+ \frac{3}{4} a_3 A^3 \left[\sin\Big(\left(2w_1 - w_2\right) t\Big) + \sin\Big(\left(2w_2 - w_1\right) t\Big)\right] \\ &- \frac{1}{2} a_2 A^2 \left[\cos\left(2w_1 t\right) - \cos\left(2w_2 t\right)\right] - a_2 A^2 \cos\Big(\left(w_1 + w_2\right) t\Big) \\ &- \frac{1}{4} a_3 A^3 \left[\sin\left(3w_1\right) - \sin\left(3w_2\right)\right] \\ &- \frac{3}{4} a_3 A^3 \left[\sin\Big(\left(2w_1 + w_2\right) t\Big) - \sin\Big(\left(2w_2 + w_1\right) t\Big)\right] \end{aligned} \tag{2.7}$$

From (2.7), and as illustrated in the frequency domain in Figure 2.5, one can conclude that the first linear term in (2.6), i.e., $a_1 x$, amplifies the fundamental tones. The quadratic term in (2.6), i.e., $a_2 x^2$, down-converts the RF signal to the DC band and generates the second harmonic band. The cubic term in (2.6), i.e., $a_3 x^3$, generates intermodulation

distortion (IMD) and third harmonic band. The IMD is this case are known as third-order IMD (IMD3) and are located around the fundamental frequency. Note that, the second and third harmonic bands can be easily filtered out by a band-pass filter around the fundamental frequency. Therefore, their significance is minor.

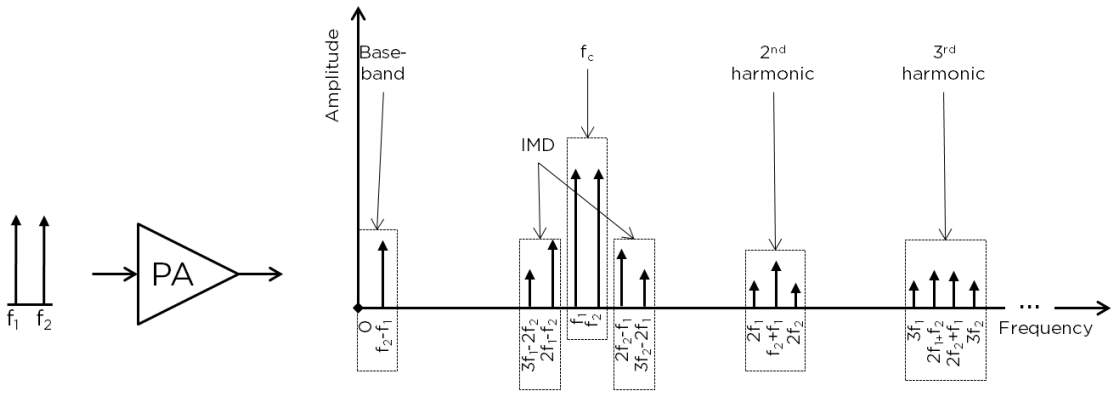

Figure 2.5: Frequency representation of a two-tone test

From (2.7) and Figure 2.5, it is clear that only the cubic term generates frequency components around the carrier frequency, and the quadratic term does not. This fact can be generalized for all odd and even terms [63]. In fact, the PA is assumed to have a narrow fractional bandwidth compared to the carrier frequency; therefore, all even-order terms are harmonics and baseband. Consequently, the even-order do not interfere with the fundamental frequency.

When the PA exhibits memory effects, Figure 2.6 shows that the two-tone test allows their identification. In fact, new IMD that occur at the same IMDs caused by the static nonlinearity of (2.7) appear. These IMD are distortions that alter the already existing IMDs as a phase shift. In this way, the discrepancy of IMDs around the two tones demonstrates a remixing of these as a consequence of memory effects. Memory effects characterization can be demonstrated by varying the two-tone spacing and measuring the discrepancy of the left and right IMDs around the two tones.

The two-tone test can be used to characterize a nonlinear PA driven with a modulated signal. This is done by setting the frequency separation of the two tones, i.e., $|f_2 - f_1|$, equal to the bandwidth of the modulated signal [10, 42, 55, 86]. Yet, measurements showed that although this characterization gives an estimation of the PA response, the two-tone test fails to accurately model the PA response and its behavior. Therefore, PA characterization when using modulated signal is required.

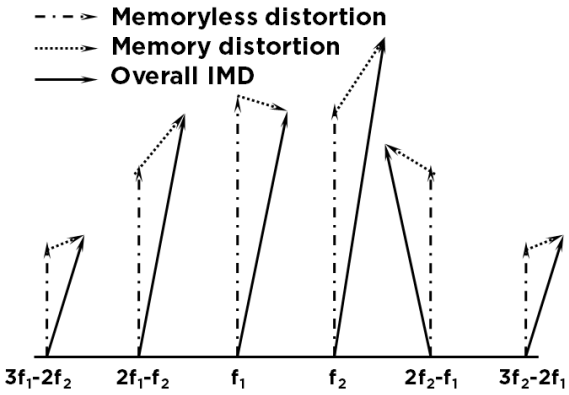

Figure 2.6: Intermodulation distortion of a nonlinear power amplifier exhibiting memory effects

2.1.2.3 Power Amplifier Nonlinearity Characterization With Modulated Signal

PAs driven with modulated signals exhibit stronger nonlinearities and memory effects when compared to the case of two-tone stimulus. Consequently, different characterization methods are put in practice to determine the unwanted signals engendered by the nonlinear PA. Two established measures exist for this purpose, namely ACPR and EVM characteristics.

- The ACPR, also expressed as the adjacent channel leakage ratio (ACLR), is the ratio between the total neighbouring channels' powers to the main channel signal power. It describes the degree of the signal regrowth into adjacent channels, as shown in Figure 2.7. The ACPR is an important criterion for wireless radio standards because it characterizes the maximum power allowed to be radiated outside the allocated band. This measure is used to prevent interference with other adjacent radios [1].

- The EVM is a measure that quantifies the imperfection to the output signal when compared to the input one. The EVM is illustrated in Figure 2.8, where v is the ideal output signal, w is the measured PA output signal, e is the magnitude error, θ is the phase error, and $w - v$ is the EVM. For a signal sequence of length N, the EVM is defined in (2.8).

$$EVM = \sqrt{\frac{\sum_{j=1}^{N}\left[\left(I_j - \widetilde{I_j}\right)^2 + \left(Q_j - \widetilde{Q_j}\right)^2\right]}{\sum_{j=1}^{N}\left[I_j^2 + Q_j^2\right]}} \tag{2.8}$$

  where I_j and Q_j are the ideal output signal in-phase and quadrature components and $\widetilde{I_j}$ and $\widetilde{Q_j}$ are their output measured counterparts, respectively.

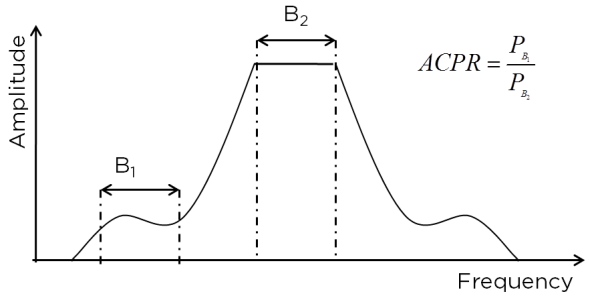

Figure 2.7: Adjacent channel power ratio

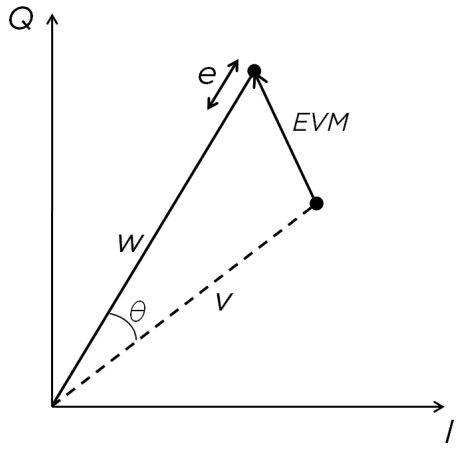

Figure 2.8: Error vector magnitude

In the previous, only distortions caused by change of the current input amplitude were discussed, i.e., static nonlinearities. However, as the bandwidth of the input signal gets wider, new sources of distortion happen, which are known as memory effects. These memory effects occur when the distorted output signal amplitude and phase vary as functions of not only the current input one but also of the previous input and output signals. Unlike the static nonlinearity, memory effects do not generate new spectral components. Instead, the memory effects change the shapes of the already existing signal components, as it was illustrated in Figure 2.6.

PAs exhibit fading memory effects [97]. Thus, PA output at time t depends on a finite spanning history of the input and output signals. To characterize PA memory effects, PA circuit characterization is essential to determine their sources. Consequently, accurate yet simple modeling is possible as it will be shown in Chapters 5 and 6. Figure 2.9 shows the principle sub-blocks of which a PA is composed; namely the transistor, the input and output matching networks, and the bias network. The transistor amplifies the signal and is the main contributor to the static nonlinearity. In fact, depending on the PA biasing point, the transistor transforms some input current to a voltage shape that may cause nonlinearity. The three networks, i.e., biasing, input and output matching networks, are responsible of the memory effects, due to their non-constant frequency responses. Two different types of memory effects exist:

- Short-term memory effects take place on a time scale that is close to the carrier frequency f_c. The input and output matching networks generate short-term memory effects because of the reactive components present in these networks (chip capacitors, inductors, transmission lines).

- Long-term memory effects take place on a time scale that is close or smaller than the signal bandwidth. Some of these long-term memory effects are manifested by the spread of the AM/AM and AM/PM around the compression. This is caused by thermal effects, charge trapping effects and DC-self biasing. These long-term memory effects consist of even distortions of the DC bias network. This causes reflection of the input signal, thereby resulting in memory effects.

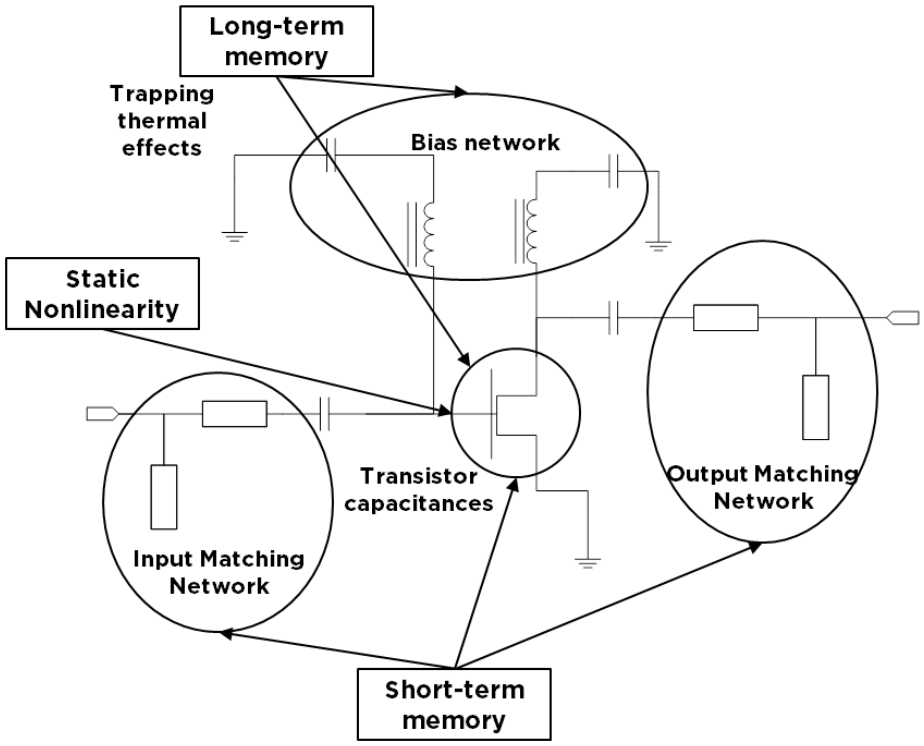

Figure 2.9: Nonlinearity, short-term and long-term memory effects sources in power amplifiers

2.1.3 Power Amplifier Linearization Techniques

As discussed in the last subsection, linearity and efficiency are two critical design factors that cannot be satisfied simultaneously. Different linearization techniques were introduced in the literature to enhance the linearity-efficiency trade-off. PA are driven to high input power for high efficiency and linearization is applied to enhance the linearity in this efficient region. Three major techniques for PA linearization have been reported in the literature, namely feedback [29, 83], feed-forward [53] and DPD [36, 45, 87] linearizers. In the following, an overview about these linearization techniques is given with a special focus on DPD.

2.1.3.1 Feedback Linearization

The feedback approach is a method based on algebraic transformation of nonlinear systems into linear ones [83]. The feedback linearization principle consists of compensation for the nonlinear system output signal by a feedback loop. In this way, the current output signal is used to adjust the future input ones. This is achieved by comparing the output signal to a desired output. Then, an error signal is deduced and substituted from the next input signal. Figure 2.10 shows a typical feedback linearizer block diagram. Its transfer function is given in (2.9), where $x(t)$ and $y(t)$ represent the feedback system input and the output, respectively, and G and A are PA and feedback loop transfer functions, respectively.

$$y(t) = \frac{G}{1 + GA} x(t) \tag{2.9}$$

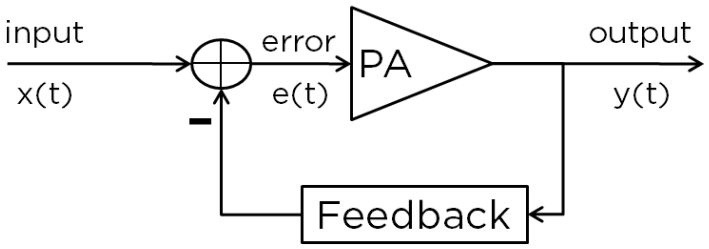

Figure 2.10: Feedback linearization block diagram

Feedback linearization is reasonably simple to implement. However, it suffers from many drawbacks when applied to PA linearization. Its major problem gains importance when it is applied to wideband signals. Indeed, the feedback linearizer reduces the PA gain and suffers from serious potential instability. Better linearity comes at the expenses of $1/(1 + AG)$ gain reduction, as shown in (2.9). In addition, it is generally difficult to preserve stability in the case of large frequency range, i.e., when feedback linearization is used with wideband signals. In fact, the feedback system stability is determined by the feedback loop gain AG variation with frequency. This is difficult to control over a wide bandwidth signal.

2.1.3.2 Feed-Forward Linearization

Figure 2.11 illustrates a typical feed-forward linearizer. Feed-forward linearization principle is based on the generation of an error signal, $e(t)$. This error signal is intended to represent the additive distortion products that are present at an attenuated version of PA output signal, $y_{PA}(t)/SSG$.

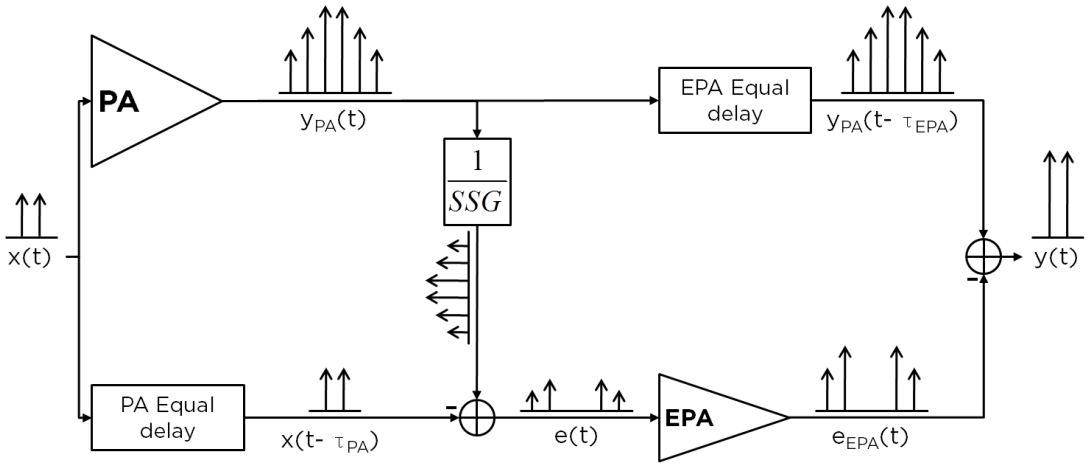

Figure 2.11: Feed-Forward linearization block diagram

The feed-forward linearization works as follows. The error signal is scaled up to its original magnitude using an error power amplifier (EPA). Then, it is subtracted from the PA output, $y_{PA}(t - \tau_{EPA})$, where τ_{EPA} is the propagation delay caused by the EPA. These steps result in a linear PA. It is important to note that, there is no bandwidth limitation attributed to the linearization effort. Since phase shifts is introduced in the carrier of the PA and EPA, PA and EPA equal delays are introduced to compensate for these propagation delays. These two delay blocks must be very tightly tuned, in order to realize good linearization performance. In addition, the gain of the EPA must be very well matched to the gain of the PA so that good cancellation of the distortion can be realized. Moreover, the EPA must be linear to amplify the error signal, $e(t)$, with no additional distortions. Finally, the analog delays and the EPA must be extremely low-loss, in order for the whole linearization effort to be worthwhile in terms of efficiency. Feed-forward linearization is very effective, however, it is complex to implement and introduces non negligible addition power consumption.

2.1.3.3 Digital Predistortion

The DPD principle is based on intentionally introducing a nonlinear function in the base-band part of the transmitter, which is commonly called the digital pre-distorter. This function generates additional distortions that are out of phase to those introduced by the PA. Thereby, DPD leads to a linear cascade, as illustrated in Figure 2.12. The DPD corrects the power and phase level of a primary input signal. In this way, the corrected signal nonlinear amplification results in a linear amplification of the primary input signal. The cascaded system (DPD+PA) linearity relies primarily on the DPD ability to produce non-linearities that are complementary in magnitude and out of phase to the PA ones. Hence,

digital pre-distorter successful deployment requires proper choice of the predistortion function and accurate PA behavior modeling.

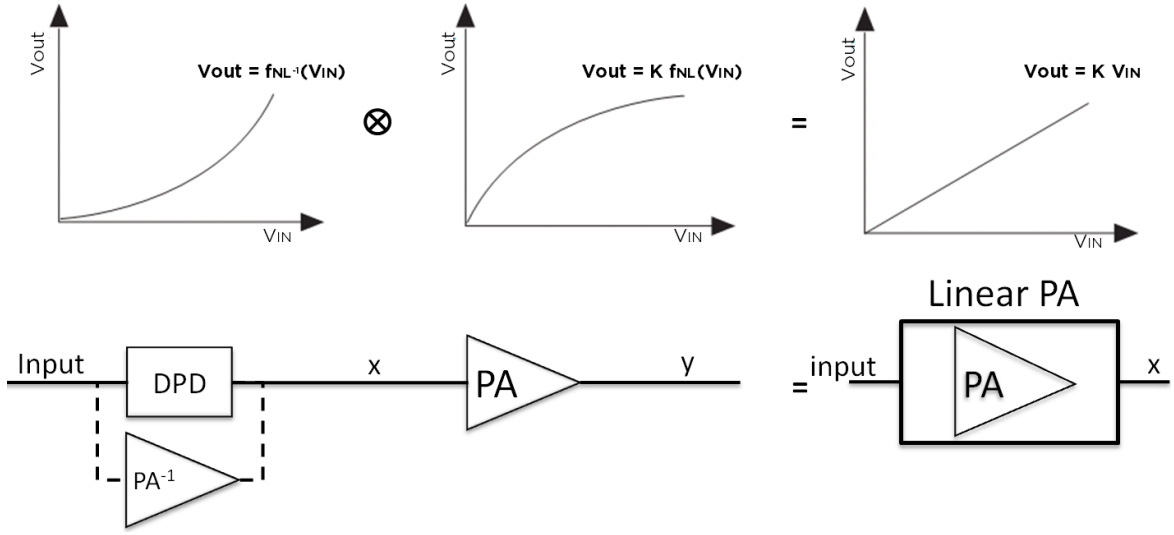

Figure 2.12: Digital predistortion principle

It is possible to introduce a post-distorter function instead of pre-distorter one. In this case, the complementary nonlinear function is placed after the PA. However, post-distortion is undesirable, because it leads to processing the signal at a high power level. This results on a high power consumption and decreases the overall efficiency. Nevertheless, it is demonstrated that post-distortion and predistortion exit simultaneously and are equal [46, 91].

In the following, the DPD principle is illustrated for a memoryless PA, as shown in Figure 2.13. To achieve the PA output level, V_{out_A}, an adjustment of the input signal amplitude from V_{in_A} to V_{in_B} is required. This correction can be deduced by projecting the input V_{in_A} in the DPD transfer function, which provides the required level V_{in_B}. The DPD is, therefore, responsible for the transformation from level V_{in_A} to level V_{in_B}. This DPD transformation is found to be equal to the inverse of the PA transfer function. The DPD transformation is achieved by finding a function that maps the PA output signal to the input one. This is known as the indirect learning.

Predistortion main attribute is the simplicity of its principle. In addition, the DPD is unconditionally stable and does not suffer from a bandwidth limitation thanks to the open loop-scheme. Moreover, due to the DPD low power consumption relative to the PA, there is no significant degradation in the transmitters power efficiency. However, DPD is optimized for specific PA operating conditions (average output power, modulation scheme, bandwidth, PAPR, etc.). Unfortunately, the DPD loses its complementarity with the PA

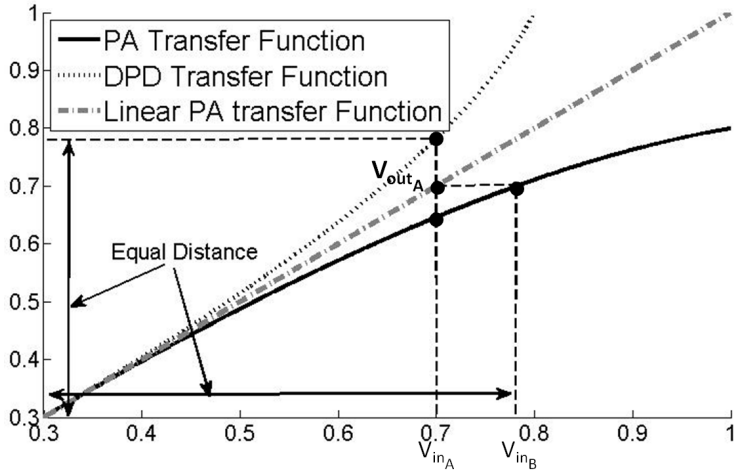

Figure 2.13: Digital predistortion principle illustration

nonlinearities as these conditions change. Thus, constructing a good DPD is conditional on the determination of a simple, yet accurate PA behavioral model that is insensitive to a signal or environmental change. It is shown in Chapter 6 that, if the DPD function is not set correctly for the case of multi-band signal, the linearization performance can degrades due to change in the time delay misalignment.

2.2 Wide- and Multi-Band Digital Predistortion Literature Review

The DPD principle, as explained in the previous section, relies on producing PA inverse function. To synthesize this function, a good understanding and accurate modeling of PA distortion is essential. Actually, the DPD construction is achieved usually through two steps:

1. Model the PA behavior to identify PA nonlinearity characteristics.
2. Construct the DPD function by modeling the PA inverse function.

PA behavioral models, also called black-box models, consider the PA device under test (DUT) as an unknown device. A behavioral model finds the best dynamic nonlinear function that relates some measured output data to the measured input data, as shown in Figure 2.14. Behavioral models are used at high levels of abstraction in the overall circuit or system design [4], due to their simplicity and lack of necessity for an in-depth knowledge of the system. Behavioral models have the advantage of fast convergence suitable for

simulation. They can be applicable for processes that are not understood and require little expertise in the domain of the system application. However, they cannot guarantee reliable extrapolation; it is widely known that behavioral models are only valid where they were characterized and constructed [4]. Indeed, they are non-scalable, very dependent on the measurement data, and very sensitive to noise. They give little understanding about the system being modeled and they can only be applicable to an already existing process since they require measurement. In this thesis, behavioral models will be adopted as modeling schemes, due to their principle simplicity, and, because they take advantage of the advanced development of DSP techniques. However, it is shown throughout Chapters 5 and 6 that the modeling procedure will not be limited only to black-box modeling. Instead, physical inspiration and PA behavior will be used to derive reduced complexity DPD schemes without compromising its simplicity and linearization capability.

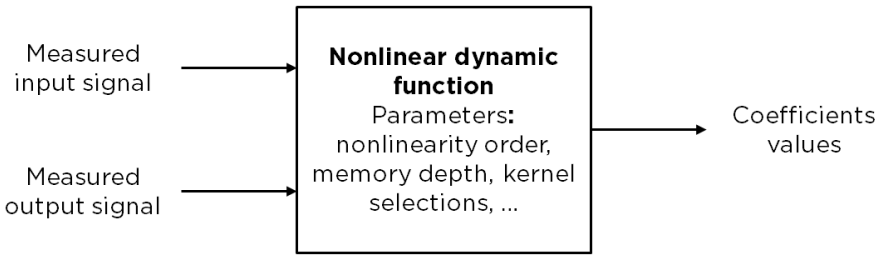

Figure 2.14: Black box modeling procedure

PA behavioral modeling is not a trivial problem, and different PA model schemes have been suggested in the past. In the literature, these models can be classified in two categories: memoryless models and memory capable models. With a memoryless model, the current output signal depends only on the current input ones through a nonlinear mechanism. This instantaneous nonlinearity is referred to static nonlinearity. The Saleh [90], memoryless polynomial [99], the LUT [18, 23, 56] models and diode based linearizers [79, 94] were the first techniques that aimed to fit the PA static nonlinearity. As the input signal bandwidth gets wider, the PA begins to exhibit memory effects. Accordingly, the PA output signal no longer depends only on the current input one but also on past input and output ones. In other words, the PA becomes a dynamic nonlinear system. For such a PA, memoryless predistortion can achieve only limited linearization performance. Therefore, the DPD model needs to capture these memory effects by introducing structures capable of modeling them. Since the focus of this thesis is on wide-/multi-band signals, only memory capable models are tackled.

In the following, detailed descriptions of the main behavioral models/DPD are presented. First, important and recent SISO models are presented and their limitation are discussed. Second, new MIMO models are presented and their limitation are discussed. In

both cases, the proposed solutions in this thesis are put into context. The Volterra series is considered as a reference for modeling dynamic nonlinear systems due to its richness [101]. Nevertheless, the Volterra series has always been criticized for its prohibitive complexity and restricted applicability to high nonlinear PAs. Alternatively, the excellent capability of artificial neural networks (ANNs) to accurately approximate continuous functions [26, 34, 43] has been successfully exploited to model RF and microwave devices/circuits [5, 12, 21, 28, 44, 65, 66, 69–71, 82, 96, 98, 104, 109]. In this thesis the Volterra series scheme is adopted. Chapter 3 introduces an efficient approach to prune the Volterra series while allowing high nonlinearity order and large memory span with reduced complexity.

2.2.1 Single-Input Single-Output Behavioral Models

2.2.1.1 Volterra series base model

A Volterra series can be thought of as a Taylor series with memory [93]. In other words, it is a polynomial approximation to the target function at the present instant convolved with the target function over some history. The mathematical formulation of the Volterra series is given in (2.10), where $x(t)$ and $y(t)$ are the input and output signals, respectively; and h_n are the Volterra series kernels.

$$\begin{aligned} y(t) &= \sum_{n=0}^{\infty} H_n[x(t)] \\ &= \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_n(\tau_1, \cdots, \tau_n)\, x(t-\tau_1) \cdots x(t-\tau_n)\, d\tau_1 \cdots \tau_n \end{aligned} \tag{2.10}$$

In the discrete time domain, the Volterra series formulation is given in (2.11), where $x(n)$ and $y(n)$ are the discrete input and output signals, p is the nonlinearity order, and $M_0, \cdots, M_p$ are the respective memory depths for the order $0, \cdots, p$. Chua et al. [14] demonstrated that this formulation is applicable to fading memory systems, e.g., PA and consequently DPD.

$$y(n) = \sum_{p=1}^{\infty} \left[\sum_{m_1=0}^{M_0} \cdots \sum_{m_p=0}^{M_p} h_p(m_1, \cdots, m_p)\, x(n-m_1) \cdots x(n-m_p) \right] \tag{2.11}$$

One major drawback of the Volterra series is its large number of coefficients; the number of coefficients increases exponentially with the nonlinearity degree and the memory

depth. For example, a Volterra series with a nonlinearity degree of 7 and 5 memory taps contains some 8,364 coefficients that need to be determined, assuming only odd orders are involved. Historically, this was an intractable problem, and Volterra series were limited to low nonlinearity degrees and only a reduced number of memory taps. However, unlike other behavioral models, e.g., ANN, Gaussian functions, etc., the advantage of the Volterra series resides on the linearity of its coefficients with respect to the input signal. Indeed, Volterra series identification is determined using simple and powerful linear algorithms that are capable of fast convergence into the global minima [81]. Therefore, number of techniques to reduce the number of coefficients in the Volterra series have been implemented in the literature. In the following, an overview of the major Volterra series approximation is given.

Two Box Models

A very simple reduction of the Volterra series is two-box models, known also as Wiener or Hammerstein models [37, 54, 61, 69, 70, 110]. These models have been employing a cascade of nonlinear function and a linear filter. In the case of the PA, the first box of the Hammerstein scheme captures its static nonlinear behavior, while the second one is intended to take into account memory effects, as shown in Figure 2.15. For the Wiener model, the two boxes are arranged differently than in the Hammerstein case as shown in Figure 2.16.

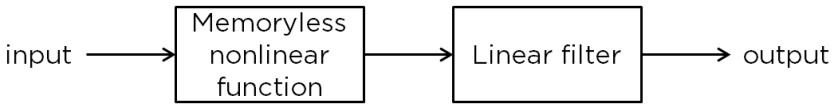

Figure 2.15: Block diagram of the Hammerstein model

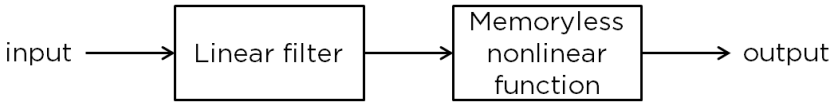

Figure 2.16: Block diagram of the Wiener model

In the literature, polynomial functions, LUT and ANN system has been used to construct the static nonlinear function of the Hammerstein/Wiener models. Authors in [61] showed the limitation of conventional Hammerstein/Wiener scheme in mimicking the wideband PA behavior. Alternatively, they suggested an augmented version of these models with weakly nonlinear one with multiple filters to solve the modeling inaccuracy. Parallel Hammerstein/Wiener models [54] have been also suggested to address the traditional

schemes limited capability in accounting for the memory effects. This was achieved by stacking extra branches in parallel, however, their parameters identification become too tedious. In [69, 70] authors used ANN for an accurate construction of the static nonlinear part, then, the error signal between the output and the input signal of the memoryless sub-model was filtered and post-injected at the model output through two linear filters.

The main problem of two box models resides in their identification; nonlinear algorithms are needed to extract accurately the two boxes coefficients. A popular technique to divert this issue is to use two steps identification. First, the memoryless block is identified and then its response is used for the identification of the memory block [69]. It was demonstrated in the literature that the two box models achieve inferior linearization capabilities when compared to Volterra series and MP models even for narrow band signals [36]. Thus, it is expected that two-box model performance will be inferior to other models in the case of wideband signals.

Memory polynomial

The MP model [51, 78] is the most popular simplification of the Volterra model that excludes cross terms to alleviate the complexity of the Volterra series. In other words, MP exploits the diagonal kernels of the Volterra series corresponding to the pure powers of the input signals samples. MP model is expressed in (2.12), where $x(n)$ and $y(n)$ are the discretized input and output signals, respectively; M is the memory depth; N is the highest polynomial order; and, $a_{i,j}$ are complex coefficients.

$$y(n) = \sum_{j=0}^{M-1} \sum_{i=1}^{N} a_{i,j} |x(n-j)|^{i-1} x(n-j) \tag{2.12}$$

A typical MP structure is shown in Figure 2.17, where M is equal to the number of branches and N is the highest order of the different polynomials in each branch. Figure 2.17 suggests that the order of nonlinearity in the different branches does not have to be the same. However, for ease of representation, orders of the different polynomials of the MP model are usually considered to be equal. Despite the simplicity of the MP structure, attempts to reduce the number of coefficients of the MP model are proposed. In [63], elimination of the even-order coefficients from the MP formulation is proposed. This was motivated by the fact that even-order only affect the harmonics, i.e., very far from the PA carrier frequency. In [11], the number of MP coefficients was reduced by assigning individual nonlinearity orders in the different MP branches. The author in [11] succeed in decreasing the number of coefficients to almost one third and the conditioning num-

ber by three orders of magnitude while maintaining the same modeling and linearization capabilities.

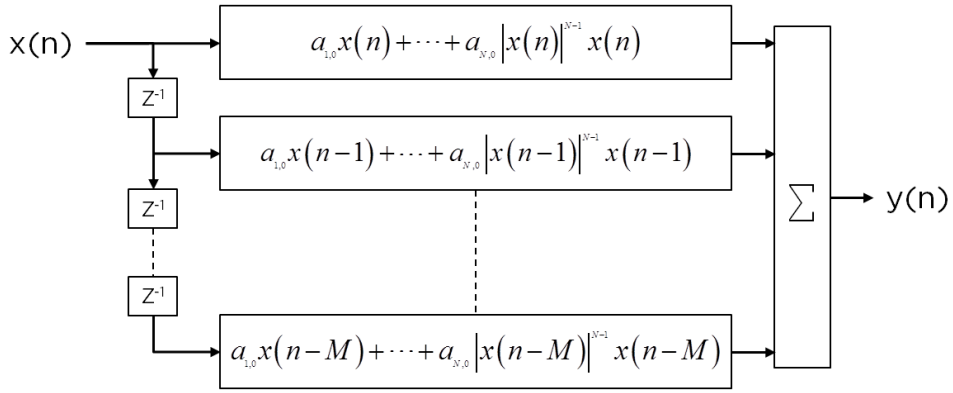

Figure 2.17: Memory polynomial structure

Due to the large nonlinearity order required to model/linearize a given PA, the MP models high conditioning number and range of the coefficients render it not suitable for real time implementation. In [77, 78], authors proposed a new MP formulation that is more suitable for hardware implementation. It uses the Horner's rule and a new arrangement in coefficients that allows their range to be smaller. Consequently, the new MP formulation requires less numbers of bits to preserve the precision of the implementation of the DPD. This formulation is given in (2.13).

$$y(n) = a_1 x(n) \left(1 + \left(\frac{a_2}{a_1} |x(n)| \left(\cdots 1 + \left(\frac{a_{n-1}}{a_{n-2}} |x(n)| \left(1 + \frac{a_n}{a_{n-1}} |x(n)|\right)\right)\right)\right)\right) \quad (2.13)$$

Compared to the Volterra series formulation, the MP reduces considerably the number of coefficients. However, due to the lack of cross terms, its application to linearize highly nonlinear PA has limited capabilities, e.g., for Doherty PA [71]. In Chapter 3, a new Volterra series pruning technique based on G-Fucntionals is proposed. This pruning allows a selection of a PA/DPD model with reduced number of coefficients while preserving cross terms. In addition, it also results in a small range of coefficients suitable for hardware implementation.

Dynamic Deviation Reduction

In dynamic deviation reduction (DDR) techniques [40, 120–122], the Volterra series is expanded around the deviation of the signal from its static value, which may be DC or, as more commonly used in PA modeling, the large-signal RF tone. In this way, the Volterra

series is expressed in terms of a static polynomial plus dynamic polynomial contributions of up to several orders. The model reduction is achieved by limiting the number of dynamic orders used in the approximation. Some success with using only first-order dynamics has been claimed [121]. However, in more complex PA structures used in wireless communications, higher dynamic orders are needed [40, 120]. Although the DDR approach reduces considerably the number of kernels of the Volterra series, its identification is still computationally extensive. In fact, when using wideband signals, strong memory effects happen and require high order of dynamics resulting in a large number of coefficients. A description of the Volterra series DDR and its derivation is provided in the following paragraphs.

For a nonlinear, causal, and time-invariant system with fading memory, a truncated discrete time domain Volterra series, expressed in (2.14), is used to model its output response. $x(n)$ and $y(n)$ are the nonlinear system input and output discrete signals, respectively; $h_p(i_1, \cdots, i_p)$ is the p^{th}-order Volterra kernel; and, P and M denote the nonlinearity degree and memory depth, respectively.

$$y(n) = \sum_{p=1}^{P} \sum_{i_1=0}^{M} \cdots \sum_{i_p=0}^{M} h_p(i_1, \cdots, i_p) \prod_{j=1}^{p} x(n - i_j) \tag{2.14}$$

In the following, a description of the DDR technique is given. A dynamic deviation function given in (2.15) is introduced to reduce the complexity of the Volterra series formulation, where $x(n-i)$ is the delayed input signal.

$$e(n, i) = x(n - i) - x(n) \tag{2.15}$$

Substituting (2.15) in (2.14) results on the following expression:

$$\begin{aligned} y(n) &= \sum_{p=1}^{P} \sum_{i_1}^{M} \cdots \sum_{i_p}^{M} h_p(i_1, \cdots, i_p) \prod_{j=1}^{p} [x(n) + e(n, i_j)] \\ &= \sum_{p=1}^{P} x^p(n) \sum_{i_1=0}^{M} \cdots \sum_{i_p=0}^{M} h_p(i_1, \cdots, i_p) \\ &\quad + \sum_{p=1}^{P} \sum_{r=1}^{p} x^{p-r}(n) \binom{p}{r} \sum_{i_1=0}^{M} \cdots \sum_{i_p}^{M} h_p(i_1, \cdots, i_p) \prod_{j=1}^{r} e(n, i_j) \end{aligned} \tag{2.16}$$

The output signal, $y(n)$, can be expressed as a summation of static $y_s(n)$ and dynamic $y_d(n)$ characteristics, where a_p are the memoryless coefficient of the response, and $w_{p,r}(.)$ represents the r^{th}-order dynamic kernel.

$$y_s(n) = \sum_{p=1}^{P} a_p x^p(n) \tag{2.17}$$

$$y_d(n) = \sum_{p=1}^{P} \sum_{r=1}^{p} x^{p-r}(n) \sum_{i_1=0}^{M} \cdots \sum_{i_r=0}^{M} w_{p,r}\left(i_1, \cdots, i_r\right) \prod_{j=1}^{r} e\left(n, i_j\right) \tag{2.18}$$

The previous formulation in (2.18) introduces a new variable, r. It represents the order of the dynamics that allows the truncation of the Volterra series to be a simpler version. For instance, if $r = 1$, the Volterra series formulation is truncated as given in (2.19), where only the static and the first-order dynamic behavior are taken into account.

$$y(n) = \sum_{p=1}^{P} a_p x^p(n) + \sum_{p=1}^{P} x^{p-1}(n) \sum_{i=0}^{M} w_{p,1}\left(i\right) e\left(n, i\right) \tag{2.19}$$

Another rearrangement of the Volterra series coefficients is performed to result in the Volterra series DDR shown in (2.20).

$$\begin{aligned} y(n) = & \sum_{p=1}^{P} h_{p,0}\left(0, \cdots, 0\right) x^p(n) \\ & + \sum_{p=1}^{P} \left\{ \sum_{r=1}^{P} \left[x^{p-r}(n) \sum_{i_1=1}^{M} \cdots \sum_{i_r=i_{r-1}}^{M} h_{p,r}\left(0, \cdots, 0, i_1, \cdots, i_r\right) \prod_{j=1}^{r} x\left(n - i_j\right) \right] \right\} \end{aligned} \tag{2.20}$$

In the Volterra series DDR, the variable r is used to limit the dynamic response order. This method reduces considerably the number of coefficients when compared to the full Volterra series, as shown in Table 2.1.

The DDR-based Volterra series is still relatively limited to low nonlinearity orders, since it preserves the important dependency of the number of kernels on the nonlinearity order. For wideband signals, e.g., LTE-advanced, where the PA/DPD models run at 500 MHz, the DDR method is still computationally complex. For example, if the Volterra series model accounts for dynamic up to 100 ns (low memory values) when running at 500 MHz, the DPD requires the use of the 50 past values of the input signal. Combined with the nonlinearity degree equal to 5 (relative low nonlinear order), and when using the Volterra series DDR model (where the dynamics are set equal to 1, the simplest reduced Volterra series model), the resulting number of coefficients of the DPD is around 250. This optimistic, yet still large number of coefficients, increases the complexity of the identification of the DPD model even when using a lab computer equipped with MATLAB. More difficult implementation will be faced when dealing with real-time systems, in terms of memory

Table 2.1: Number of coefficient when modeling DPD with a 100 MHz bandwidth input signal

DPD	**Memory**	**Nonlinearity**	**Volterra series DDR**			
Speed	**Depth**	**Order**	**r=0 (MP)**	**r=1**	**r=2**	**r=∞ (Full)**
300 MHz	50 ns	5	45	78	888	$> 113e3$
300 MHz	50 ns	7	75	109	$> 1e3$	$> 3e6$
300 MHz	100 ns	5	90	153	$> 3e3$	$> 2e6$
300 MHz	100 ns	7	150	214	$> 5e3$	$> 255e6$
500 MHz	50 ns	5	75	128	$> 2e3$	$> 1e6$
500 MHz	50 ns	7	125	179	$> 3e3$	$> 78e6$
500 MHz	100 ns	5	150	253	$> 9e3$	$> 31e6$
500 MHz	100 ns	7	250	354	$> 14e3$	$> 7e9$

and computation. Examples of the resulting number of coefficients in the case of wideband DPD are summarized in Table 2.1. This table shows that the conventional DPD solution for wideband signals, even when using the Volterra series DDR, is not practical. In fact, it leads to very complex model with unrealistic number of coefficients to be identified.

The Volterra series DDR has only been applied to signals with relatively low bandwidth, up to 20 MHz, due to the uniform selection of the kernels. This pruning technique leads to a large number of coefficients. Therefore, new methods are needed to truncate the number of kernels in the Volterra series scheme when dealing with wideband signals. In Chapters 3 and 5, two novel pruning approaches for Volterra series are introduced. These approaches were used for PA behavioral modeling/DPD schemes that were driven with signals having a bandwidth up to 60 MHz. The resulting reduced complexity model show similar modeling performances and linearization capabilities with fewer numbers of coefficients to the DDR approach. This was achieved by using non-uniform selection of the kernels. Thus, high selection of nonlinearity order and memory depth was possible without augmenting extensively the number of coefficients. Consequently, it will be concluded that these new pruning approaches are more suitable to be applied to wideband signals.

2.2.1.2 Transmitter Observation Receiver Bandwidth for Wideband Signal

In the previous subsection, an overview of different DPD engines was given. In Chapters 3, 5 and 6, new DPD engines will be introduced that allow the DPD model construction

with high nonlinearity and memory depth without exponentially increasing the number of coefficients. However, when deploying wideband signals, another problem rises. To capture the PA output signal driven with a wideband signal, a wideband TOR with high-speed ADCs and flat response over a wide range of frequencies is required. This is particularly important for capturing accurate measurement data essential for the DPD coefficients identification. However, the significant power and cost overheads of high-speed ADCs [49] bring down the overall efficiency of the DPD+PA cascade. This, it eventually limits the usefulness of the DPD approach to enhancing the trade-off between efficiency and linearity. Consequently, reducing of the TOR bandwidth is a must when deploying wideband signals.

Several attempts to relax the TOR bandwidth have been reported in the literature. The relaxed TOR bandwidth consequently reduces the required ADC sampling rate, and alleviates the complexity of the DPD implementation as the signal bandwidth broadens. In [123], Zhu introduced the GST to reconstruct a nonlinear system output signal using a reduced TOR bandwidth. This theorem calls for an additional function which is used to reduce the output signal bandwidth. According to [123], the application of GST in PA behavioral modeling problems requires a priori knowledge of its corresponding inverse function, called $g(.)$ function in [123]. In this way, the PA output signal can be sampled using the input Nyquist rate (INR); i.e., twice the input signal bandwidth, rather than the output Nyquist rate (ONR); i.e., twice the output signal bandwidth. Typically, the ONR is equal to five times the INR. The GST principle is explained in the following subsection. Authors in [6, 103, 119] proclaimed GST successful application to PA behavioral modeling and linearization. However, a close examination of the reported approaches reveals non-conformance with the GST as stated in [123] and introduction of under-sampling technique [50]. As an example, the PA behavioral modeling procedures reported in [103] did not involve the inverse function $g(.)$ as stipulated in [123]. In [119], the authors proposed the identification of the DPD function using the GST and a sampling rate equal to the INR. However, simulation and measurement results were conducted using a sampling rate higher than the INR. Authors attributed this choice to the sensitivity to the noise. In Chapter 4, it is demonstrated that the GST direct application for DPD identification with reduced TOR bandwidth is not possible. Alternatively, a new approach is proposed.

In [32, 100], the Volterra series kernels were constructed using input and output signals sampled at the INR instead of the ONR. This was achieved using an under-sampling technique that exploits the alias of the output signal harmonic and/or inter-modulation components [50]. However, the focus of the current thesis is on lowering the TOR bandwidth used to capture the PA band limited output signal driven with band-limited single harmonic modulated signal. Under-sampling technique cannot be applied in this case as the signal alias will fold on the modulated signal. A new approach is proposed in Chapter

4 to enable DPD construction using only a portion of PA output signal bandwidth. Consequently, the linearization with the reduced TOR bandwidth will enable the application of the GST.

In [15], the author proposed a method based on multiple narrow-band observations at the PA output signal. The proposed method consisted of capturing a frequency portion of the output signal and used a modified DPD algorithm to update the DPD coefficients accordingly. In fact, by capturing only a portion of the output signal, a linear transformation, noted $L\{.\}$, can be applied to both the input and output signals. In this way, the PA/DPD model coefficients can be approximated without acquiring the full input/output signal spectrums.

In [41], the authors proposed linear sub-filtering to adjust the PA/DPD model in sub-blocks. Thereby, the output sampling rate is lowered by any desired order. The proposed approach works as follows. First, a memoryless nonlinear function was applied to the input signal, allowing the generations of IMDs. Next, linear sub-band filters were used to adjust these IMDs, so they align with the suitable DPD model. Note that, there was no restriction on how wide the linear filters bandwidth could be. In [41], the authors split the signal into M adjacent frequency sub-bands by a predefined linear filter bank. Decimation after each linear filter separation was then applied to reduce the sampling rate value. In this manner, the signals were divided into M sub-sampled input signals and M sub-sampled output signals. These signals were then used to adjust the pre-distorted signal so that it was aligned with a suitable DPD function through linear filters. In the final stage, the pre-distorted signals were recombined using up-sampling and filtering to rematch them to the input signal full bandwidth. Using this technique, the authors showed the possible linearization of a 3G BS PA driven with a 4 MHz signal when 4 and 8 sub-bands were used. The technique proposed in [41] provided a good alternative to reduce the sampling rate. However, the authors did not extend their technique to wider bandwidth signals, i.e., 20 MHz or higher. In addition, the proposed technique used linear filters to adjust the IMDs. Thus, only limited results could be achieved when dealing with highly nonlinear PA, especially when the sub-bands interact with each other.

In [38, 114], the authors proposed a band-limited Volterra series model to restrict the linearization application to the available TOR bandwidth. It was shown that linearization was successful within the set observation band. The residual spectrum regrowth introduced by the PA outside the observation bandwidth, and not tackled by the band-limited Volterra series DPD, was mitigated using a sharp band-pass filter. The band-pass filter high-order would imply non negligible insertion loss, which should be included in order to carefully assess the potential performance degradation. While the band-limited Volterra DPD does not allow for complete linearization of PA output IMDs, it appears to be the most promising

practical solution in the literature.

In Chapter 4, it is shown that to compensate for wideband PA nonlinearity, the correct pre-distorted signal can be accurately synthesized using a fraction of the output signal bandwidth. Since the GST cannot be used, a new approach is proposed to reduce the required TOR bandwidth without compromising the linearization capability. This approach allows the identification of the DPD coefficients using a reduced TOR bandwidth by pre-processing the PA input signal using a nonlinear function. It is identified that the sensitivity of the linearization capacity to the agreement between the PA inverse model and the structure of the DPD function (nonlinearity order, memory depth, and dominant distortion products needed to model the PA inverse) used to pre-process the input signal. Theoretical and experimental analysis revealed a robust and accurate PA output linearization when TOR bandwidth is set to only twice the input signal modulation bandwidth instead of the usual five times.

In the following, overview of GST, under-sampling and band-limited Volterra series are given.

Generalized Sampling Theorem

The GST was introduced as a solution to capture a nonlinear system output signal, e.g., a nonlinear PA, at a reduced sampling rate, i.e., less than the ONR. The GST stipulates that for an arbitrary function $f(.)$, with no requirement limit on is output signal bandwidth, if a one-to-one mapping function $g(.)$ can be found such that $g(f(t))$ is band limited, i.e., $\exists w_0 \in \mathbb{R}$ such that the Fourier transformation $G_f(w) = 0 \ \forall |w| \geq w_0 = \frac{\pi}{T_s} = \pi f_s$, then $f(.)$ can be fully recovered by sampling at the points $t_n = nT_s$. This is possible through sampling $g(f(t))$ at reduced rate. Then, $g^{-1}(.)$ is applied as given in (2.21) to recover $f(t)$, where g^{-1} is the inversion of g.

$$f(t) = g^{-1}\left\{\sum_{-\infty}^{\infty} g\left[f(t_n)\right] \frac{sin\left[w_0\left(t - nT_s\right)\right]}{w_0\left(t - nT_s\right)}\right\} \tag{2.21}$$

From this theorem, an arbitrary function $f(.)$ can be sampled at a low rate, whether it is band limited or not, if the function $g(.)$ can be found and applied with $f(.)$. This can be seen as a frequency compression of $f(.)$ and a frequency decompression, in order to recover. Utilization of this technique allows for sampling rate reduction to capture the PA output. However, in Chapter 4, it will be shown that the GST cannot be used for the DPD coefficients identification solely and an alternative technique is proposed. It is also shown that after identifying the DPD, the GST is used to estimate the PA wideband output response using a reduced TOR bandwidth.

Under-Sampling

The conversion of analog signal to a digital one depends on two important criteria that should be carefully set, namely time step and amplitude quantization as shown in Figure 2.18. To ensure unambiguous data, i.e., non-aliased signal, the sampling frequency f_s, i.e., the inverse of the time step $t_s = 1/f_s$, should be equal or greater than twice the analog signal maximum frequency, according to the Nyquist-Shannon theorem [50, 92]. In the case of an envelop signal, i.e., signal having only components around a carrier frequency f_c, sampling according to the previous yields to sampling unoccupied frequencies, from DC to the envelop signal around f_c. To solve the issue, the signal is down-converted to an intermediate frequency (IF) so that the signal occupies all frequency component from DC to $f_s/2$ as shown in Figure 2.19. In this case, the sampling rate f_s is equal or greater than the envelop signal bandwidth rather than its maximum frequency component. Alternatively, under-sampling can be used to sample directly the envelop signal with a new condition on the sampling rate as explained hereafter.

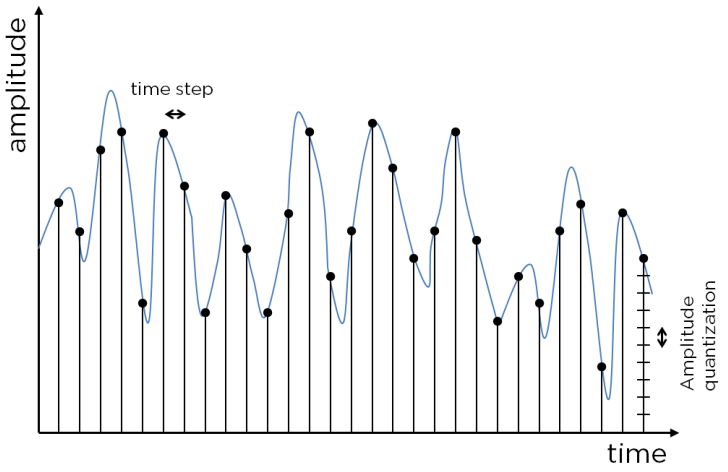

Figure 2.18: Sampling of analog signal [50]

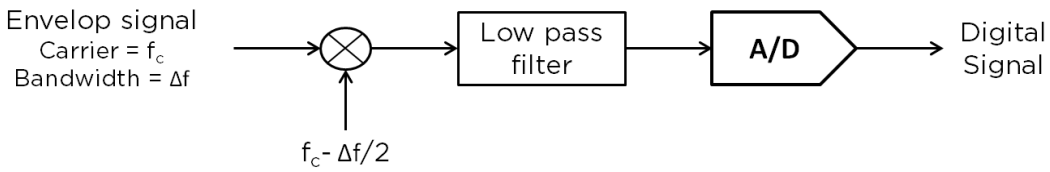

Figure 2.19: Sampling of an envelop signal using down-converting

The under-sampling uses aliased signals to capture the envelop signal as long as one of these aliased signals lies in the first half of the Nyquist range, DC to $f_s/2$ [50]. Any signals which fall outside the bandwidth of the interest should be filtered using a band-pass filter.

In this case, a new sampling rate is required and is given in the following. If an analog signal occupies a bandwidth $\Delta f = f_h - f_l$, where f_l and f_h are the signal lowest and highest frequency components, then the sampling rate has to be within $\frac{2f_h}{n} < f_s < \frac{2f_l}{n-1}$, where n is positive integer such that $n < \frac{f_h}{f_h - f_l}$. It can be deduced that the sampling rate should be at least about twice the bandwidth of the signal $f_s > 2\Delta f$ but much less than the highest frequency component $f_s \ll f_h$. For example, in the case where an envelop analog signal has a bandwidth of 100 MHz and modulated at $f_c = 2.14\,GHz$, without down-converting the signal to IF, a sample rate higher than twice 2.14 GHz should be deployed. However, using the under-sampling, knowing that $f_h = 2.14\ GHz + {}^{100}\!/_{2}\ MHz = 2.19\ GHz$ and $f_l = 2.14\ GHz - {}^{100}\!/_{2}\ MHz = 2.09\ GHz$, the sampling rate should be within $\frac{2\times 2.19\ GHz}{n} < f_s < \frac{2\times 2.09\ GHz}{n-1}$ with $n < \frac{2.19\ GHz}{100\ Mhz} = 21.9 \Rightarrow n_{max} = 21$. Thus, $208\ MHz < f_s < 209\ MHz$ which is about twice the bandwidth of the envelop signal.

It is important to distinguish between the GST and the under-sampling. The GST goal is to reduce the sampling rate less than twice the bandwidth of the envelop signal. Whereas, the under-sampling reduces the sampling rate from twice the signal maximum frequency component to twice its envelop bandwidth.

Band-Limited Volterra series

The band-limited Volterra series [114, 115] is applied in the case where only the distortions within a limited bandwidth is captured, as shown in Figure 2.20. The band limited Volterra series is explained in the following. Let's consider a PA modeled using a Volterra series as given in (2.22) where $x(n)$ and $y(n)$ represent the input and output signals, respectively, $h_p\left(i_1, \cdots, i_p\right)$ is the p^{th}-order Volterra kernel, and $D_p\left[x\left(n\right)\right] = \prod_{j=1}^{p} x\left(x - i_j\right)$ is the p^{th}-order Volterra operator. One problem of Volterra series is that the bandwidth increases with the p^{th}-order Volterra operator as shown in Figure 2.21. The identification of a Volterra series model using reduced bandwidth input and output signal can lead to modeling error. In fact, the Volterra series model output will have non-zero frequency components outside the limited band, while measurement capture only the limited band.

$$y\left(n\right) = \sum_{p=1}^{\infty} \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} h_p\left(i_q, \cdots, i_p\right) D_p\left[x\left(n\right)\right] \tag{2.22}$$

To solve the previous mentioned problem, the band-limited Volterra series incorporated a band-pass filter after every p^{th}-order Volterra operator. This restricts the bandwidth expansion to only the reduced captured band as shown in Figure 2.22. The new p^{th}-order band-limited Volterra operators are given in (2.23), where $*$ is the convolution operator

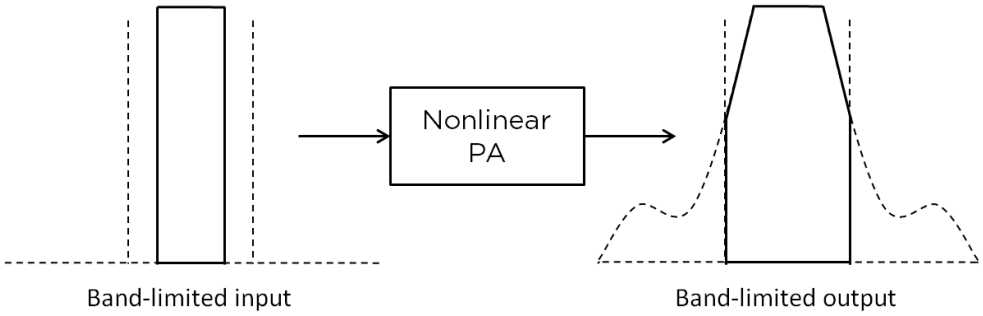

Figure 2.20: Band-limited input and output signals

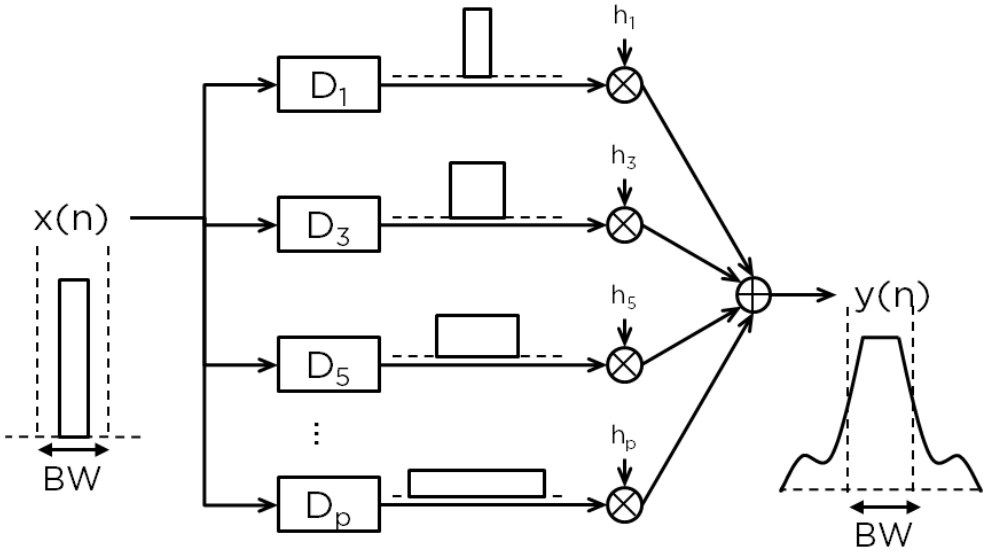

Figure 2.21: General Volterra series model

and $w(n)$ is a band-pass filter. Consequently, the band-limited Volterra series is given (2.24), where K is the length of the filter.

$$T_p[x(n)] = D_p[x(n)] * w(n) \tag{2.23}$$

$$y(n) = \sum_{p=1}^{\infty} \sum_{i_1=1}^{\infty} \cdots \sum_{i_1=p}^{\infty} h_{p,BL}(i_1, \cdots, i_p) \left\{ \sum_{k=0}^{K} \left[\prod_{j=1}^{p} x(n - i_j - k) w(k) \right] \right\} \tag{2.24}$$

Similar to the conventional Volterra series, the band-limited Volterra series kernels are linear with respect to the input signal. Hence linear identification algorithms can be used for the model coefficients identification. In Chapter 4, it will be shown that using only a band-limited input signal portion, wideband nonlinear PA output signal accurate

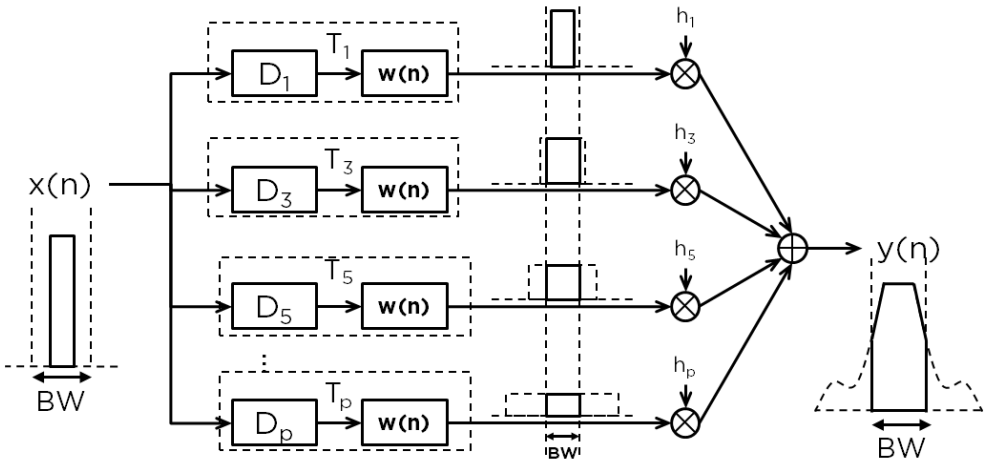

Figure 2.22: General band-limited Volterra series model

linearization and modeling is possible under certain conditions. Comparison with the band-limited Volterra series will be discussed.

2.2.2 Multi-Input Multi-Output Behavioral Models

In the previous, it was shown that wideband signals involve very complex DPD and behavioral models. Moreover, for accurate data acquisition, linear measurement system across a wide bandwidth are required leading to expensive resources for computation and memory storage. These factors become more complicated when dealing with multi-band signals. Indeed, the multi-band signal carrier frequencies separation can be very large i.e., $\gg 100\ MHz$. Therefore, a multi-band PA model using the conventional behavioral model/DPD schemes, i.e., sampling all the frequencies at once, is no longer practical. Actually, using conventional behavioral PA modeling/DPD leads to the capture of many unoccupied frequencies between the carrier frequencies, in order to cover the multi-band signal at once. Thereby, this leads to the loss of hardware resources in sampling noise, i.e., unrealistic ADCs and digital-to-analog converters (DACs) sampling rate requirements to capture unnecessary data. Moreover, even if such ADCs and DACs can be conceived, the model scheme and identification process complexities will be unrealizable in real-time systems. This is due to the large amount of memory and computation resources needed. Accordingly, new multi-band modeling schemes should be envisaged to reduce the requirements of the model complexity and the hardware resources needed to implement these models.

Different attempts to model/linearize multi-band PAs have been presented in the literature. These attempts were limited to only dual-band and few tri-band cases. In the

following, introduction to multi-band nonlinearity is given. Second, major contributions in multi-band PA output modeling and linearization is given.

2.2.2.1 Distortion in case of Multi-Band Power Amplifier

A typical nonlinear PA frequency representation driven by K multi-tone multi-band stimulus is given in Figure 2.23. It is clear that in addition to the inter-modulation generated by each band around its respective carrier frequency, the multi-band signal generates cross modulation. These cross modulation are a result of the mix between the multi-band signals. Consequently, the PA/DPD model can no longer be a SISO model, since it fails for accounting for these cross modulation. Alternatively, a MIMO model shown in Figure 2.24 is envisioned, where $x_i(n)$ and $y_i(n)$ are the PA input and output signals around each carrier frequency f_{c_i}, respectively, and $G_i(.)$ are MIMO nonlinear functions. These functions describe the PA nonlinear outputs around f_{c_i} as given in (2.25), where $j^2 = -1$.

$$y_i\,(n) = G_i\,[x_1(n), \cdots, x_K(n)]\,e^{j2\pi f_i} \qquad i \in \{1 \cdots K\} \tag{2.25}$$

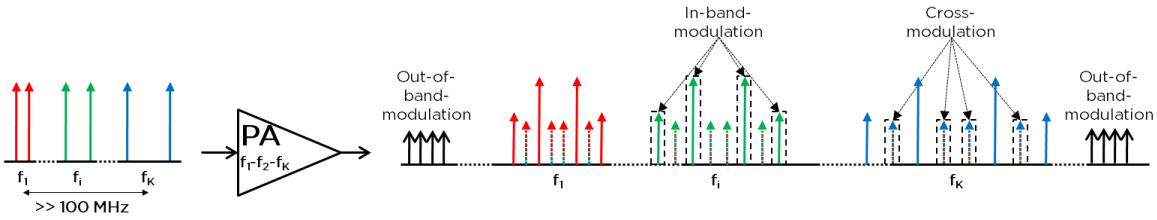

Figure 2.23: Input and output spectrums representation of a nonlinear power amplifier driven by multi-tone multi-band signal

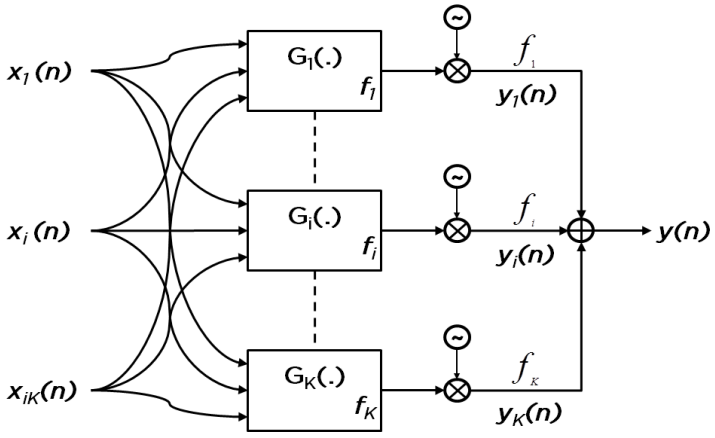

Figure 2.24: Multi-input multi-output multi-band power amplifier model

Ideally, the $G_i(.)$ functions should be expressed in MIMO Volterra series form. In the case of dual-band signal, the Kronecker product can be used as follows:

$$G_i\left[x_1(n), x_2(n)\right] = \sum_{m_1=0}^{M} H_{m_1}^T X_{m_1} + \sum_{m_1=0}^{M}\sum_{m_2=0}^{M} H_{m_1 m_2}^T (X_{m_1} \otimes X_{m_2}) + \cdots + \sum_{m_1=0}^{M} \cdots \sum_{m_p=0}^{M} H_{m_1 \cdots m_p}^T \left(X_{m_1} \otimes \cdots \otimes X_{m_p}\right) \;\; i \in \{1,2\} \tag{2.26}$$

where $\otimes$ denotes the Kronecker product, which is defined as:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \otimes \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 y_1 \\ x_1 y_2 \\ x_2 y_1 \\ x_2 y_2 \end{bmatrix} \tag{2.27}$$

The vectors $H_{m_1 \cdots m_p}$ and X_{m_p} are defined as:

$$H_{m_1 \cdots m_p} = \left[h_{m_1 \cdots m_p,1}, h_{m_1 \cdots m_p,2}, \cdots, h_{m_1 \cdots m_p,2p}\right]^T \tag{2.28}$$

$$X_{m_p} = \begin{bmatrix} x_1\left(n - m_p\right) \\ x_2\left(n - m_p\right) \end{bmatrix} \tag{2.29}$$

where $h_{m_1 \cdots m_p}$ is the p^{th}-order kernel of the Volterra series.

The previous formulation guarantees the optimal functions $G_i(.)$ that describe an accurate dual-band PA. Unfortunately, the number of coefficients in this previous formulation explodes exponentially with the nonlinearity order and memory depth. In Chapters 5 and 6, a complexity-reduced generalized memory polynomial is derived from a simple PA block diagram to represent the functions $G_i(.)$. This proposed formulation allows cross terms between different delayed input signals to account for the cross modulation distortions. In addition, the proposed formulation results in a manageable number of coefficients.

2.2.2.2 Dual-Band Power Amplifier Behavioral Models

In the literature, dual-band PA behavioral models were developed using two strategies; generalization of SISO model structure to accommodate dual-band signals [8, 88, 107] and pruning of general Volterra series formulation [30].

In [8], authors extended the SISO MP formulation to a dual-input model (2D-DPD). The 2D-DPD allows the construction of a predistorted signal using two input signals to

model/linearize cross distortion generated by the two-band interaction on each others. However, it has been demonstrated that the 2D-DPD suffers from numerical instability as reported in [88, 107]. Moreover, it is shown in Chapter 6 that the 2D-DPD is very sensitive to time delay misalignment between the two transmitted signals. This results in modeling/DPD performances degradation.

One particular problem with SISO model generalization to multi-input cases is the increase of the number of coefficients. Authors in [88] proposed two box model structures (memoryless block+MP) to limit the increase of number of coefficients to only the memoryless block. The derived model achieved smaller number of coefficients when compared with the 2D-DPD. Yet, nonlinear identification algorithms are needed to extract accurately the model coefficients. It is demonstrated in the literature that the two box models achieve inferior linearization capabilities when compared to Volterra series and MP SISO models [36]. Thus, the same is expected in the case of dual-band and multi-band signals. In [107], similarly to the 2D-DPD case, the authors generalized the Volterra series DDR model to the case of dual-band. However, the resulting number of coefficients was larger comparing to the 2D-DPD.

Alternatively to the SISO model generalization to dual-band one, general Volterra series schemes pruning was proposed in [30]. A systematic pruning and derivations were applied to yield a simple baseband equivalent low-complexity model. This model does not require empirical pruning and uses only few number of coefficients. However, the model includes kernels that need pre-computation before the model coefficient identification, thus leading to model implementation complexity increase. The baseband equivalent Volterra model showed comparable results to the 2D-DPD when applied to behavioral modeling and DPD. In the following subsection, an overview about the 2D-DPD model is given. The 2D-DPD model is used Chapter 6 as a model reference for the proposed dual-band model.

One particular problem in dual-band DPDs identification is the necessity of dual-output signal sampling at the same time. This is required to ensure that all the PA output signal are time aligned and synchronized with the PA input ones. A common method to solve this issue is using repeated signals. Tn this case, only one TOR is used to capture the different PA output signals. However, in real time system, this solution is not practical as the signal changes constantly. Alternatively, authors in [35] proposed PA model identification first, since it requires all the PA input signals (which are known) but only one output signal. Then estimation of the PA output signals is performed. In this way, all the estimated output signals are time aligned. Consequently, indirect learning can be used for DPD coefficients identification. This method allows the usage of only one TOR, however, it added the complexity of identifying a PA model. Additionally, the DPD model will suffer from the PA model imperfections. Authors in [80] proposed the usage of direct learning

instead of the in-direct learning as it requires only one output signal at a time. For that, an explicit formula is identified for the DPD coefficients based on the assumption that the PA is modeled using 2D-DPD. This allows the identification of the DPD coefficients without the need of inverting large matrices of capturing synchronized output signals. Authors in [16] proposed also an explicit formula for dual-band DPD that doesn't requires multiple receiver to capture a synchronized dual-output signals. However, the performances of these techniques were inferior to the 2D-DPD one.

Dual-band signals are usually located at carrier frequencies that are not necessarily close, thus they involve a large range of frequencies. To characterize these signals, the measurement tools deployed suffer from I/Q imbalance. Authors in [111] proposed a joint predistorter and I/Q imbalance correction for the transmitter imperfection across wide range of frequencies. The proposed technique involved the cascade of a SISO Volterra filter for the compensation of the I/Q imbalance and dual-band LUT table for the correction of the dual-band PA distortions.

In this thesis, the 2D-DPD model is used as a reference model. In Chapter 6, it is shown that a new proposed model is less complex and achieves better linearization/modeling performances when compared to 2D-DPD. In addition, it is shown that the proposed model is more robust to time delay misalignment where the 2D-DPD is not. In the following, an overview of the 2D-DPD is given.

Dual-Band Digital Predistortion (2D-DPD)

Figure 2.25 shows the block diagram of the 2D-DPD model. In the following, the processing steps that lead to 2D-DPD are explained.

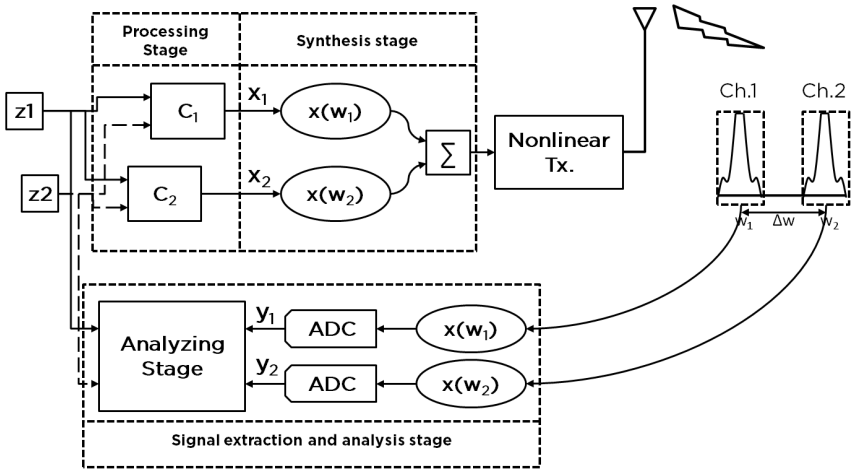

Figure 2.25: Block diagram of the dual-cell processing digital predistortion [8]

Consider two modulated signals, namely, $x_1 e^{jw_1 t}$ and $x_2 e^{jw_2 t}$, around the carrier frequencies, w_1 and w_2, respectively. For simplicity, the 2D-DPD model is first demonstrated for a simple memoryless model, and then by analogy, the case with memory is deduced. For that, consider a memoryless fifth-order nonlinearity, $y = x + x\,|x|^2 + x\,|x|^4$, and a discrete-time baseband equivalent input signal, $x = s_1 + s_2 = x_1 e^{-jwnT} + x_2 e^{-jwnT}$, where $t = nT$. By simple mathematical derivation, the PA output is given in (2.30).

$$\begin{aligned} y = &\left(x_1 + x_1|x_1|^2 + x_1|x_2|^2 + x_1|x_1|^4 + 4x_1|x_1|^2|x_2|^2 + x_1|x_2|^4\right) e^{-jwnT} \\ &+ \left(x_2 + x_2|x_1|^2 + x_2|x_2|^2 + x_2|x_1|^4 + 4x_2|x_1|^2|x_2|^2 + x_2|x_2|^4\right) e^{jwnT} \\ &+ \left(x_1^2 x_2^* + 2x_1^2|x_1|^2 x_2^* + 3x_1^2|x_2|^2 x_2^*\right) e^{-j3wnT} \\ &+ \left(x_1^* x_2^2 + 3x_1^*|x_1|^2 x_2^2 + 2x_1^*|x_2|^2 x_2^2\right) e^{j3wnT} \\ &+ \left(x_1^3 x_2^{*2}\right) e^{-j5wnT} + \left(x_1^{*2} x_2^3\right) e^{j5wnT} \end{aligned} \tag{2.30}$$

Equation (2.30) revels that the output signal contains terms at the fundamental and intermodulation frequencies. In [8], it was assumed that the two carrier frequencies are far enough to allow filtering the intermodulation and not taken them into account by the PA/DPD model. Consequently, the relevant terms of (2.30) leads to (2.31), where $y(-w)$ and $y(w)$ are the signals around the frequencies $-w$ and w.

$$\begin{aligned} y(-w) &= \left(x_1 + x_1|x_1|^2 + x_1|x_2|^2 + x_1|x_1|^4 + 4x_1|x_1|^2|x_2|^2 + x_1|x_2|^4\right) e^{-jwnT} \\ y(w) &= \left(x_2 + x_2|x_1|^2 + x_2|x_2|^2 + x_2|x_1|^4 + 4x_2|x_1|^2|x_2|^2 + x_2|x_2|^4\right) e^{jwnT} \end{aligned} \tag{2.31}$$

Equation (2.31) shows that the output signal around each carrier frequency contains both the dual-input signals. Therefore, a dual-band PA/DPD model should contain both dual-input signals in the same manner than (2.31). For the case of a model with nonlinearity order equal to N and a memory depth equal to $M-1$, (2.31) was generalized to (2.32). In (2.32), i is an index that represent the band 1 or 2 and $c_{k,j,m}^{(i)}$ are the coefficient of the model. The coefficients of the 2D-DPD model are linear with respect to the input signals $x_i(n-m) \times |x_1(n-m)|^{k-j} \times |x_2(n-m)|^j$, therefore, linear identification, e.g., LSE, can be used. It is worth mentioning that, even though the first derivation of the 2D-DPD model, i.e., (2.31), involved only odd-terms, the proposed model in (2.32) contains both even and odd terms. In fact, it was shown in [30] that 2D-DPD odd terms are not enough for linearizing dual-band PA, and can results in numerical instability. The resulting number of coefficients of the 2D-DPD model is small comparing to the dual-band Volterra based

models. However, this number is still large especially when involving even and odd terms as shown in Table 2.2.

$$y_i(n) = \sum_{m=0}^{M-1} \sum_{k=0}^{N} \sum_{j=0}^{k} c_{k,j,m}^{(i)} x_i(n-m) \times |x_1(n-m)|^{k-j} \times |x_2(n-m)|^{j} \tag{2.32}$$

Table 2.2: Number of coefficient for the 2D-DPD model

N	**2D-DPD Memory Depth M**									
	Odd Only Orders					**Even Odd Orders**				
	M=1	M=2	M=3	M=4	M=5	M=1	M=2	M=3	M=4	M=5
1	1	2	3	4	5	1	2	3	4	5
3	3	6	9	12	15	6	12	18	24	30
5	6	12	18	24	30	15	30	45	60	75
7	10	20	30	40	50	28	56	84	112	140
9	15	30	45	60	75	45	90	135	180	225

2.2.2.3 Tri-Band Power Amplifier Behavioral Models

In the literature, dual-band behavioral model/DPD had a lot of intention in the recent years, whereas fewer researches dealt with tri- and higher band. In fact, generalizing SISO model to MIMO cases or pruning a general Volterra formulation increases noticeably the required number of coefficients and/or the complexity. As an example, authors in [113] extended the MP formulation to the case of tri-band signal (3D-DPD). This model achieved more than 15 dB reduction in the ACPR; however, the number of coefficients needed was over 100 coefficients. In addition, the 3D-DPD is expected to suffer from the same limitation of the 2D-DPD, namely numerical instability and sensitivity to time delay misalignment. Alternatively, authors in [112] proposed a phase-aligned pruned Volterra series and demonstrated about 2 dB better performances than the 3D-DPD. However, the required number of coefficients was even higher than the 3D-DPD case.

In this thesis, to construct a MIMO behavioral model/DPD, a novel strategy is adopted. First, a block diagram that describe the behavior of single-, dual-, tri- and multi-band PAs are established. These block diagrams will be the root of developing a single-, dual-, tri- and multi-band GMP models. The derivations of these models are detailed in Chapters 5 and 6, whereas a general multi-band behavioral model/DPD is inspired by the special

cases. It will be shown that the resulting models are inherently low complexity and able to taken into account cross modulation and correct for time-delay misalignment.

Chapter 3

Wideband Digital Predistortion Pruning

Reduction of the PA behavioral modeling and DPD complexity is a challenging task, even when dealing with single narrow-band signals. Indeed, for accurate modeling and good linearization performances, the Volterra series model requires large nonlinearity order and a large memory depth. This yields to large number of coefficients as it was shown in Table 2.1. In this context, a new technique based on the Wiener G-Functionals [73, 75] is proposed to prune the Volterra series. This pruning technique doesn't need the estimation or restriction of the nonlinearity order or the memory depth of each kernel. This approach is measurement-based where the pertinence of each coefficients in the Volterra series formulation is evaluate off-line. Therefore, it allows the identification of dominant kernels while eliminating unnecessary and negligible ones.

The Wiener G-Functionals are originally developed by Wiener [93] as a more general formulation of the Volterra series. They are orthogonal; representing the minimum set of approximating functions. Additionally, when they are identified using a white Gaussian noise input, they are convergent in their means. These characteristics, namely orthogonality and convergence in the mean, significantly alleviate the complexity of Wiener G-Functionals identification as it is shown hereafter.

In the following, it is shown that the kernels of the Wiener G-Functionals under white Gaussian noise are determined using an explicit formula. Unlike when using LSE, all the kernels can be computed separately. Consequently, there is no need for inverting a large matrix that involves all the possible powered delayed input signals as it is used conventionally. When using the Wiener G-Functionals, increases in the nonlinearity order and the memory taps can be envisioned without an exponential increase of the number of coefficients.

This chapter outlines the Wiener G-Functional formulation and its use in pruning the Volterra series. The technique is illustrated first by identifying a known dynamic nonlinearity and then for modeling and linearizing PA output signals.

3.1 Overview of Wiener G-Functionals

Consider a nonlinear dynamical system that is described by a conventional Volterra series, where $x(t)$ and $y(t)$ are the input and output signals, respectively; and, h_n are the Volterra series kernels:

$$\begin{aligned} y(t) &= \sum_{n=0}^{\infty} K_n[x(t)] \\ &= \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} k_n(\tau_1, \cdots, \tau_n)\, x(t-\tau_1) \cdots x(t-\tau_n)\, d\tau_1 \cdots d\tau_n \end{aligned} \tag{3.1}$$

The Wiener G-Functionals aim to reformulate the Volterra series equation into an equivalent set of equations where the orthogonality property between the basis can be exploited. This reformulation builds on the orthogonality of the white Gaussian noise to determine a simplified expression of the Volterra series kernels as given in [93]. Let's consider a white Gaussian noise input signal, $x(t)$, with an autocorrelation function, $\phi_{xx}(\tau) = Au_0(\tau)$, where $u_0(\tau)$ is the unit impulse function. The p^{th}-degree nonhomogeneous G-Functionals is defined in (3.2), where $K_{n(p)}[x(t)] = \int \cdots \int k_{n(p)}(\tau_1, \cdots, \tau_i)\, x(t-\tau_1) \cdots x(t-\tau_n)\, d\tau_1 \cdots d\tau_n$ are the modified Volterra kernels of order p. The modified kernels are applied to an input signal with a degree of nonlinearity n, and are generated such that the orthogonal property of (3.3) holds.

$$\begin{aligned} g_p\left[k_p, k_{p-1(p)}, \cdots, k_{0(p)}; x(t)\right] &= \sum_{n=0}^{p} K_{n(p)}[x(t)] \\ &= k_{0(p)} \\ &\quad + \sum_{n=1}^{p} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} k_{n(p)} x(t-\tau_1) \cdots x(t-\tau_n) d\tau_1 \cdots d\tau_n \end{aligned} \tag{3.2}$$

$$\overline{K_m[x(t)]\, g_n\left[k_n, k_{n-1(n)} \cdots, h_{0(n)}; x(t)\right]} = 0 \quad \forall m < n \tag{3.3}$$

$$\Rightarrow \sum_{n=0}^{p} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} k_n(\sigma_1, \cdots, \sigma_m) k_{n(p)}(\tau_1 \cdots \tau_n) \begin{matrix} x(t-\sigma_1) \cdots x(t-\sigma_m) d\sigma_1 \cdots d\sigma_m \\ x(t-\tau_1) \cdots x(t-\tau_n)\ d\tau_1 \cdots d\tau_n \end{matrix} = 0$$

The expression of p^{th}-degree nonhomogeneous Volterra functionals, g_n, contains all the Volterra series kernel, k_n, as shown in (3.2). Thus, it can be concluded that g_n is equivalent to the Volterra series kernels, k_n. Thus, the expression of the output signal, $y(t)$, in (3.1) expressed as a function of the Volterra series kernel, k_p, can be expressed using the p^{th}-degree nonhomogeneous G-Functionals. In other words, the output signal, $y(t)$, in (3.1) is rewritten using the Wiener G-Functionals, while satisfying the orthogonality property of (3.3). Note that, the set of functionals derived to be orthogonal are denoted in capital letter as $G_n[k_n; x(t)]$. In addition, only the kernels k_n are the same for the Volterra series and the Wiener G-Functionals, whereas the kernels $k_{n-1(n)}, k_{n-2(n)}, \cdots, k_{1(n)}, and\ k_{0(n)}$ are determined uniquely from the kernel k_n. In [93], the kernel k_n was called nth-order Wiener kernel and the kernels $k_{n-1(n)}, k_{n-2(n)}, \cdots, k_{1(n)}, and\ \ k_{0(n)}$ are referred to as the derived Winer kernels of the Winer G-Functional, $G_n[k_n; x(t)]$.

$$y\left(t\right) = \sum_{n=0}^{\infty} G_n\left[k_n; x(t)\right] = \sum_{n=0}^{\infty} g_n\left[k_n, k_{n-1(n)}, \cdots, k_{0(n)}; x(t)\right] \tag{3.4}$$

In the following, identification of Wiener G-Functionals expressions is given. First, the zeroth and the first degree Wiener G-Functionals are as follows:

$$G_0\left[k_0; x(t)\right] = k_0 \tag{3.5}$$

$$g_1\left[k_1, k_{0(1)}; x(t)\right] = K_1[x(t)] + K_{0(1)}[x(t)] = \int k_1(\tau_1)x(t-\tau_1)d\tau_1 + k_{0(1)} \tag{3.6}$$

Since $\overline{H_0[x(t)]g_1\left[k1, k_{0(1)}; x(t)\right]} = 0$ according to (3.3) and $x(t)$ is a white Gaussian input signal, then:

$$\overline{H_0[x(t)]g_1\left[k1, k_{0(1)}; x(t)\right]} = 0 \tag{3.7}$$

$$\Rightarrow h_0 \int k_1(\tau_1)\overline{x(t-\tau_1)}d\tau_1 + h_0 k_{0(1)} = 0 \tag{3.8}$$

$$\Rightarrow h_0 k_{0(1)} = 0 \tag{3.9}$$

Because $h_0 k_{0(1)} = 0$ for any constant h_0, then $k_{0(1)} = 0$. Consequently:

$$G_1\left[k_1; x(t)\right] = \int k_1(\tau_1)x(t-\tau_1)d\tau_1 \tag{3.10}$$

Using the same procedure, all the Wiener G-Functionals can be identified iteratively. Few of these are given in the following and in Table 3.1.

$$G_2\left[k_2;x(t)\right] = \int\int k_2(\tau_1,\tau_2)\prod_{i=1}^{2}x(t-\tau_i)d\tau_1 d\tau_2 - A\int k_2(\tau_1,\tau_1)d\tau_1 \tag{3.11}$$

$$\begin{aligned} G_3\left[k_3;x(t)\right] &= \int\int\int k_3(\tau_1,\tau_2,\tau_3)\prod_{i=1}^{3}x(t-\tau_i)d\tau_1 d\tau_2 d\tau_3 \\ &\quad + \int k_{1(3)}(\tau_1)x(t-\tau_1)d\tau_1 \end{aligned} \tag{3.12}$$

$$\begin{aligned} G_4\left[k_4;x(t)\right] &= \int\int\int\int k_4\left(\tau_1,\tau_2,\tau_3,\tau_4\right)\prod_{i=1}^{4}x(t-\tau_i)d\tau_1 d\tau_2 d\tau_3 d\tau_4 \\ &\quad - 6A\int k_4\left(\tau_1,\tau_2,\tau_3,\tau_3\right)d\tau_3 + 3A^2\int\int k_4(\tau_1,\tau_1,\tau_2,\tau_2)d\tau_1 d\tau_2 \end{aligned} \tag{3.13}$$

$$\begin{aligned} G_5\left[k_5;x(t)\right] &= \int\int\int\int\int k_5\left(\tau_1,\tau_2,\tau_3,\tau_4,\tau_5\right)\prod_{i=1}^{5}x(t-\tau_i)d\tau_1 d\tau_2 d\tau_3 d\tau_4\tau_5 \\ &\quad - 10A\int k_5\left(\tau_1,\tau_2,\tau_3,\tau_4,\tau_4\right)d\tau_4 + 15A^2\int\int k_5(\tau_1,\tau_2,\tau_2,\tau_2,\tau_3)d\tau_2 d\tau_3 \end{aligned} \tag{3.14}$$

Table 3.1: Relation between the Wiener and Volterra kernels [93]

G-Functional	0^{th}**-order kernels**	1^{st}**-order kernels**	2^{nd}**-order kernels**	3^{rd}**-order kernels**	4^{th}**-order kernels**	5^{th}**-order kernels**
G_0	k_0					
G_1		k_1				
G_2	$k_{0(2)}$		k_2			
G_3		$k_{1(3)}$		k_3		
G_4	$k_{0(4)}$		$k_{2(4)}$		k_4	
G_5		$k_{1(5)}$		$k_{3(5)}$		k_5

To determine an explicit formula of the Wiener kernels using the white Gaussian input signal, consider an n-dimensional delay function, $D_n\left[x(t)\right] = x\left(t-\sigma_1\right)\cdots x(t-\sigma_n)$. This function is represented by a Volterra series with the kernel $d_n\left[\tau_1,\cdots,\tau_n\right] = u_0\left(t-\sigma_1\right)\cdots u_0(t-\sigma_n)$. The average product of the output signal with the n-dimensional delay system is given in (3.15):

$$\overline{y(t)D_n[x(t)]} = \sum_{p=1}^{\infty}\overline{G_m\left[k_m;x(t)\right]D_n[x(t)]} \tag{3.15}$$

According to (3.3), $\overline{D_n[x(t)]G_m[k_m;x(t)]} = 0\ \forall n < m$. Thus, if $G_n[d_n;x(t)] = D_n[x(t)] + F[x(t)]$, where $F[x(t)]$ is a sum of homogeneous functionals of degrees lower than n, the crosscorrelation for the previous term where $m = n$ is:

$$\overline{G_n[k_n;x(t)]D_n[x(t)]} = \overline{G_n[k_n;x(t)]G_n[d_n;x(t)]} - \overline{G_n[k_n;x(t)]F[x(t)]} \tag{3.16}$$

Since $G_n[k_n;x(t)]$ is orthogonal to all functionals of degree lower and n, then:

$$\overline{G_n[k_n;x(t)]F[x(t)]} = 0 \tag{3.17}$$

$$\overline{G_n[k_n;x(t)]D_n[x(t)]} = \overline{G_n[k_n;x(t)]G_n[d_n;x(t)]} \tag{3.18}$$

Therefore,

$$\begin{aligned}\overline{G_n[k_n;x(t)]G_n[d_n;x(t)]} &= n!A^n \int \cdots \int k_n(\tau_1,\cdots,\tau_n)d_n(\tau_1,\cdots,\tau_n)d\tau_1\cdots d\tau_n \\ &= n!A^n \int \cdots \int k_n(\tau_1,\cdots,\tau_n)u_0(\tau_1-\sigma_1)\cdots u_0(\tau_n-\sigma_n)d\tau_1\cdots d\tau_n \\ &= n!A^n k_n(\sigma_1,\cdots,\sigma_n)\end{aligned} \tag{3.19}$$

Consequently,

$$\overline{G_n[k_n;x(t)]D_n[x(t)]} = n!A^n k_n(\sigma_1,\cdots,\sigma_n) \tag{3.20}$$

Finally, it can be deduced that the n^{th}-degree Volterra kernel is given in (3.21) [93]:

$$k_n(\sigma_1,\cdots,\sigma_n) = \frac{1}{n!A^n}\overline{\left[y(t) - \sum_{m=0}^{n-1} G_m[k_m;x(t)]\right] D_n[x(t)]} \tag{3.21}$$

3.2 Volterra series Pruning Technique

Using the explicit expression of the Volterra kernels found in the last section, i.e., (3.21), one can identify the dominant kernels required for nonlinear system modeling. These dominant kernels are determined as those that lie within a given distance or norm from the maximum value of that order. This later generates only the required Volterra kernels, irrespective of the kernel order or memory depth. In addition, the formula (3.21) allows

the dominant kernels identification without the need to estimate the nonlinearity order and memory depths for each kernel. Conversely, the kernels with small contributions to the target are ignored.

The resulting set of dominant kernels needs not to be a uniform or dense selection in the total kernel space. This is in contrast to the Volterra series DDR approaches, where the number of kernels for a high dynamic order is large because the complete sequence of coefficients for the polynomial is included for each order. Thus, at each dynamic order, an exponential growth of coefficients is still present. The G-Functional method provides a non-uniform selection of dominant kernels that are essential for nonlinear dynamic system modeling and linearization. Therefore, the high dynamics orders can be used without suffering from the exponential growth of the number of coefficients.

The abovementioned approach can be applied in PA behavioral modeling or DPD synthesis. First, the PA is driven with a white Gaussian noise to determine the set of dominant kernels as explained above. The same PA is then driven with the realistic signal. The resulting PA input and output are then used to construct a pruned Volterra series model where only the dominant kernels are used. It is important to note that the PA should exhibit the same distortions during the phase of dominant kernels selection and when driven with the desired realistic signal. Therefore, the PA should be operated at the same carrier frequency and input average power in both cases. In addition, the white Gaussian noise and desired realistic signal should have similar characteristics that would dictate the distortions of the PA. Thus, the Gaussian noise has to be preprocessed to generate a signal with similar bandwidth and PAPR to those of the modulated signal.

The operation of band limiting the white Gaussian noise yields non zero autocorrelation for non zero lag time σ. Figure 3.1 shows the autocorrelation of various filtering scenarios of the white Gaussian noise. It is clear from Figure 3.1 that as the filter bandwidth decreases, the resulting autocorrelation deviates from the ideal Gaussian noise. In the following, the realistic signal used in the measurement has a bandwidth of 20 MHz. However, as shown clearly in Figure 3.1, the 20 MHz filtered noise autocorrelation is experiencing significant correlation compare to the white Gaussian one. Therefore, the filter bandwidth was set to 30 MHz when generating the signal used to identify to set of dominant kernels.

The abovementioned approach is summarized in Figure 3.2. It starts with the determination of the set of dominant kernels when the PA is driven with white Gaussian noise. The same PA is then excited with a realistic signal. The PA input and output signals are then used to construct a pruned Volterra series model where only the previously chosen set of kernels is used.

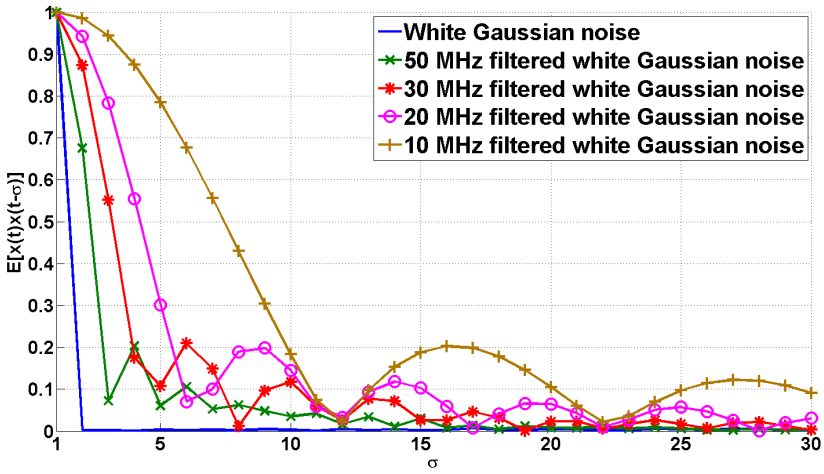

Figure 3.1: Autocorrelation of filtered Gaussian white noise

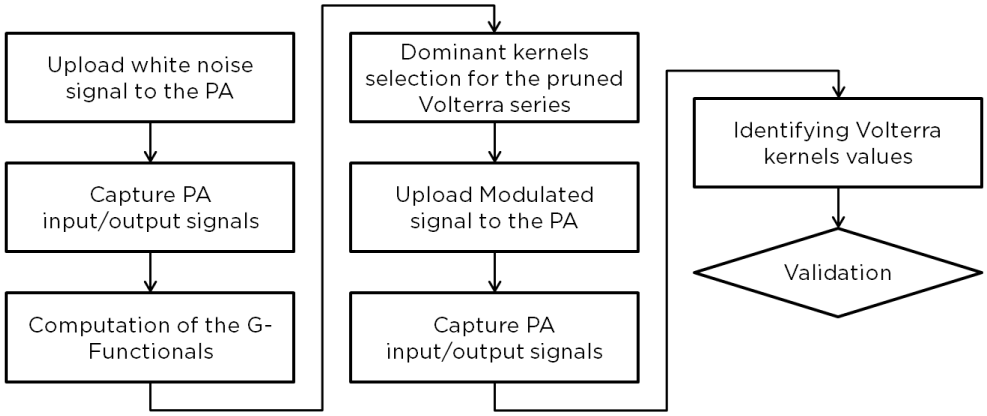

Figure 3.2: Volterra series pruning using the Wiener G-Functionals for Volterra kernels selection

Volterra Kernels Selection using G-Functionals for a Predefined Dynamic Nonlinear System

For initial validation, the previously described approach is used here to determine the dominant kernels of a predefined dynamic nonlinear system given in (3.22), where α_i are preset constant values, $\forall\, i \in \{1, \cdots, 6\}$.

$$\begin{aligned} y(n) = {} & \alpha_1 x(n) + \alpha_2 x(n-2) + \alpha_3 x^2(n) x^*(n) + \alpha_4 x(n) x(n-4) x^*(n-4) \\ & + \alpha_5 x^3(n) x^*(n) x^*(n-5) + \alpha_6 x^2(n) x(n-2) x^{*2}(n-5) \end{aligned} \tag{3.22}$$

The application of (3.21) to the predefined nonlinear system led to the following set of

dominant kernels: $\{h_1(0), h_1(2), h_1(4), h_3(0,0,0), h_3(0,4,4), h_5(0,0,0,0,5), h_5(0,0,0,5,5)\}$, which have been found to accurately approximate the predefined dynamics of the system.

3.3 Behavioral Modeling and Digital Predistortion using G-Functionals Pruned Volterra series

In this section, the G-Functionals pruning technique is applied to model a 250 W Doherty PA driven with 20 MHz 4C WCDMA signal, and to linearizer a 45 W gallium nitride (GaN) PA [105] driven with 20 MHz 1001 WCDMA signal. Both PAs are operating at the carrier frequency $f_c = 2.14\ GHz$.

3.3.1 Behavioral Modeling using G-Functionals Pruned Volterra series

In this subsection, modeling the output response of the 250 W Doherty PA is investigated. First, the measured output signal of the Doherty PA driven with the filtered white Gaussian noise was used to compute the Wiener G-Functionals. Subsequently the set of non-uniform dominant kernels is deduced. It is worth mentioning that the Wiener G-Functionals computation was performed using a nonlinearity order and memory depth equal to 7 and 9, respectively. These large values were chosen to cover the largest possible combination of Volterra kernels. Nevertheless, no major computation complexity was involved thanks to the explicit formula given in (3.21). Furthermore, the Wiener G-Functionals were computed only once with the objective being the set of dominant kernels determination necessary to accurately model the PA behavior.

Following the dominant kernels identification, the pruned Volterra is reduced to small number of coefficients. Table 3.2 shows the selected kernels for the nonlinearity order of 3 which are a non-uniform selection. Indeed, Table 3.2 shows a number of gaps in the selection of h_3. Comparing this selection with the Volterra series DDR model leads to the following conclusions. On one hand, if a small number of coefficients are targeted, i.e., when using the first order reduction, i.e., $r = 1$, the DDR results in a selection of kernels that are less general than those found by the Wiener G-Functional based technique. For example, kernel $k_3(1,2,3)$ will not be selected, since it represents cross terms of the third order. On the other hand, if a higher order of deviation, i.e., $r = 3$, and a larger number of memory taps, i.e., $M = 9$, are used in the DDR, the number of coefficients will be extremely large, higher than 500 coefficients for only the third order. This demonstrates

that the G-Functional pruning method selects only the dominant kernels irrespective of the nonlinearity order or memory depth. Thus, it allows high nonlinearity order and memory taps without increasing exponentially the number of coefficients.

Table 3.2: Pruned kernels of order using the G-Functionals

Kernel 3					
$h_3(0,0,0)$	$h_3(0,3,2)$	$h_3(0,7,8)$	$h_3(0,0,1)$	$h_3(0,3,3)$	$h_3(0,8,8)$
$h_3(0,0,2)$	$h_3(0,4,4)$	$h_3(0,8,9)$	$h_3(0,1,0)$	$h_3(0,4,5)$	$h_3(0,9,9)$
$h_3(0,1,1)$	$h_3(0,5,5)$	$h_3(1,1,0)$	$h_3(0,1,2)$	$h_3(0,5,6)$	$h_3(1,1,1)$
$h_3(0,1,3)$	$h_3(0,6,5)$	$h_3(1,1,2)$	$h_3(0,2,0)$	$h_3(0,6,6)$	$h_3(1,1,3)$
$h_3(0,2,1)$	$h_3(0,6,7)$	$h_3(1,2,1)$	$h_3(0,2,2)$	$h_3(0,7,6)$	$h_3(1,2,2)$
$h_3(0,2,3)$		$h_3(0,7,7)$		$h_3(1,2,3)$	

Following the dominant kernel selection, pruned Volterra series identification using the LSE algorithm is performed. Figure 3.3 demonstrates the capacity of the pruned Volterra series in predicting the PA output signal to deduce the AM/AM characteristics. It is clear that the prediction is in very good agreement with the measured output signal. The pruned Volterra model also shows similar agreement for the AM/PM characteristics. The overlap between the measured and modeled power spectrum densities (PSDs) of the output signal, as shown in Figure 3.4, and a normalized mean square error (NMSE) equal to -43 dB, confirm the modeling capability of the pruned Volterra series. These results were obtained using only 73 kernels, with a nonlinearity order and a memory depth equal to 7 and 9, respectively.

3.3.2 Digital Predistortion using G-Functionals Pruned Volterra series

In this subsection, the 45 W GaN PA output response linearization is investigated. Similarly to the previous subsection, the measured PA output signal driven with modified white Gaussian noise was used to compute the Wiener G-Functionals. Subsequently, a set of non-uniform dominant kernels is deduced. Following the dominant kernels selection, the pruned Volterra series identification using the LSE algorithm was performed. Figure 3.5 demonstrates the pruned Volterra series capacity in reducing the output signal ACPR to lower than -50 decibels relative to the carrier (dBc). According to Figure 3.5, the proposed pruned Volterra series model shows similar results than the Volterra series DDR when the

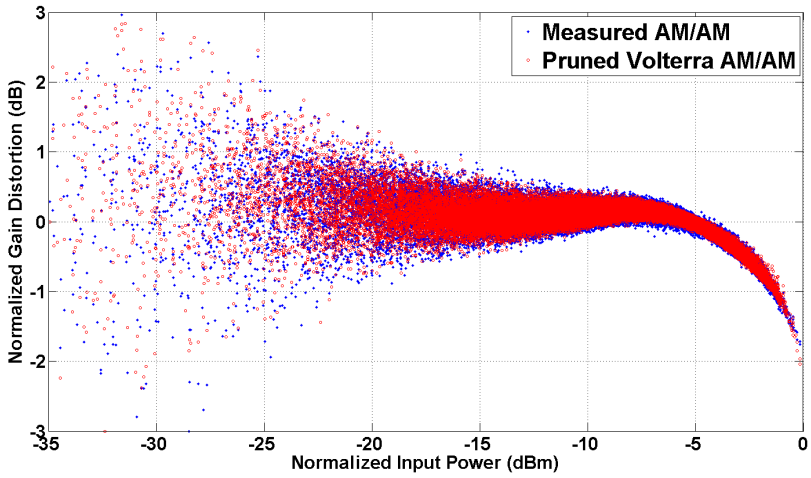

Figure 3.3: Measured and modeled AM/AM characteristic of the power amplifier

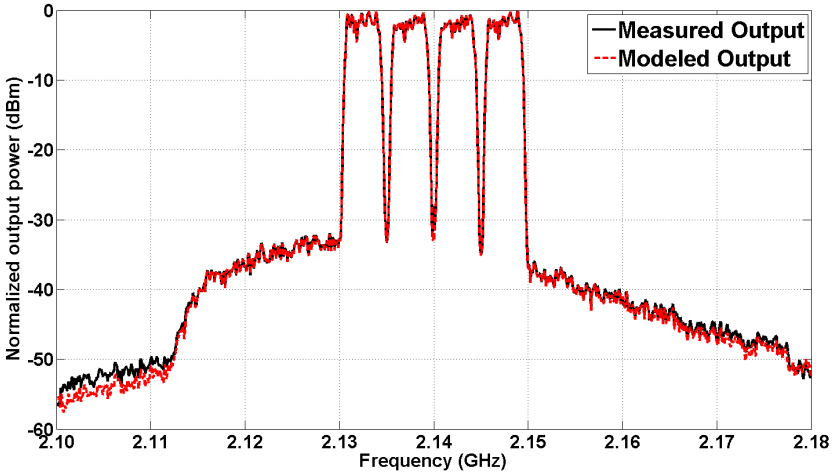

Figure 3.4: Measured and modeled spectrums of the power amplifier output signal

reduction order was set to 2, i.e., $r = 2$. However, Table 3.3 reports better EVM with fewer number of coefficients for the proposed pruned Volterra series. This is attributed to the adequate non-uniform dominant kernels selection when adopting the G-Functionals pruning technique. Furthermore, according to Figure 3.6 and Table 3.3, the G-Functionals pruning coefficients values range is smaller than when using the DDR approach. This results in easier hardware implementation, e.g., when using field programmable gate array (FPGA). In fact, smaller coefficient values range requires lesser bits to preserve the computational accuracy. For instance, in Table 3.3, the coefficient values range is reduced by a factor of 100 in the case of the G-Functionals versus DDR, resulting in a 7-bits, i.e., $log_2(100)$, reduction to achieve the same computational accuracy. Therefore, the G-Functional approach achieves

similar ACPR while improving the EVM and the ease of hardware implementation.

Table 3.3: Linearization and modeling performances comparison

	EVM (dB)	Number of coefficients	Conditioning number	$\frac{max(\|h_p\|)}{min(\|h_p\|)}$
Without DPD	-25.45	N.A.	N.A.	N.A.
Volterra series DDR DPD	-35.05	91	3.8e7	4.4e4
G-Functionals pruned DPD	-38.15	90	4.2e5	4e2

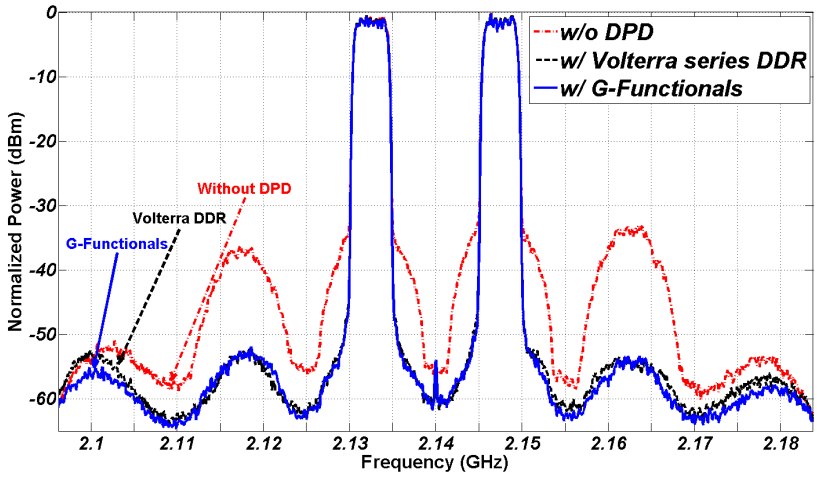

Figure 3.5: Output spectrum of the power amplifier before and after digital predistortion

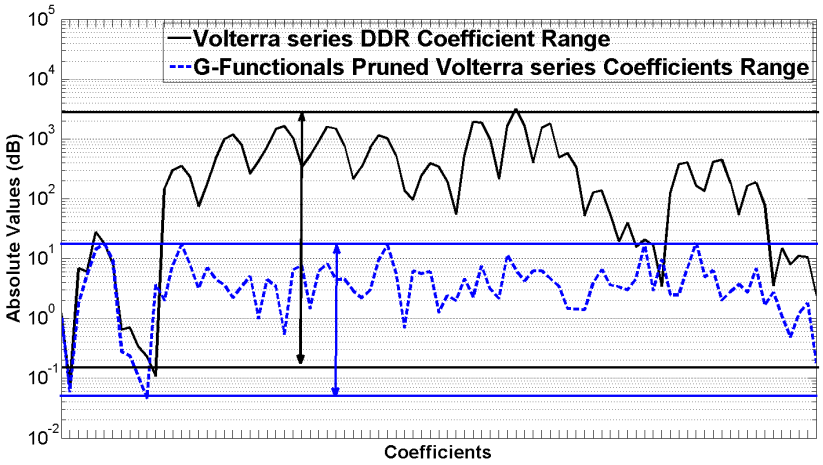

Figure 3.6: Absolute value of Coefficients for the Volterra series DDR and G-Functionals

3.4 Conclusion

This chapter introduces a novel pruning technique to reduce the number of kernels in the Volterra series formulation using G-Functionals. With memory taps and highest kernel order set to 9 and 7, respectively, a pruned Volterra series which involves only 73 coefficients was constructed to linearize a nonlinear PA output signal. Excellent EVM and ACPR reduction results demonstrate the reliability of the Wiener G-Functionals based pruning of the Volterra series. In addition the Volterra pruning allowed for significant improvements of the numerical computation (through smaller conditioning number and smaller range of coefficient values) when compared to the Volterra series DDR. It is shown in Chapter 4 that this pruning is very important when reducing the TOR bandwidth as it selects only the dominant kernels. This technique is, however, not suitable for the pruning of multi-band PA. Alternatively, Chapters 5 and 6 propose a new procedure to construct accurate, yet, complexity reduced, multi-band behavioral models/DPD schemes for multi-band PA.

Chapter 4

Wideband Digital Predistortion Using Reduced Transmitter Observation Receiver Bandwidth

Deployment of wideband signals necessitates the use of wideband TOR with high ADCs and flat frequency response over a wide range of frequencies. The significant power and cost overheads of high-speed ADCs reduce the DPD+PA cascade overall efficiency. Consequently, they limit the usefulness of the DPD approach to enhance the trade-off between efficiency and linearity.

The goal of this chapter is to relax the TORs bandwidth stringent requirements. It starts with demonstrating that GST is insufficient on its own to reduce the TOR bandwidth. Then, a new approach is devised to identify the DPD predistorter which utilizes an appropriate nonlinear function to pre-process the PAs input signal. G-Functionals, introduced in Chapter 3, are used to prune a low pass equivalent Volterra series. Consequently, the nonlinear function structure, i.e., dominant kernels, nonlinearity order, and memory depth, that approximates the PA behavior inverse is determined. It will be shown that using only a portion of the PA output signal spectrum, the coefficients of the DPD are iteratively updated to linearize its output response. It is also demonstrated that the TOR bandwidth reduction extent is proportional to the agreement between the nonlinear function structure and the actual structure of the PA inverse. Simulations and measurements conducted using different PA architectures and signals demonstrate the validity and generality of the proposed approach for linearizing and modeling wideband signals of up to 60 MHz. The modeling and linearization will be evaluated inside and outside of the reduced TOR bandwidth. The band-limited Volterra series [114] is used as a model reference to evaluate the proposed approach performances. Following the linearization of output of the

PA, it will be shown that the GST is enabled to model a wideband PA output signal using reduced TOR bandwidth.

4.1 General Sampling Theorem for Digital Predistortion Identification

The application of DPD to linearize a PA requires capturing the output signal using a TOR that has a bandwidth equal to at least five times the input signal modulation bandwidth. As the latter widens, this requirement becomes less manageable. The GST was suggested as a potential solution to model nonlinear systems with reduced observation bandwidth, however, it will be demonstrated in this section that GST is not suitable for DPD identification. As explained in Chapter 2, the GST [123] provides a solution to capture the output signal of a nonlinear system, e.g., a PA at a sampling rate lower than the ONR, ultimately at the INR.

In the PA post-distortion cascade case, if $f(.)$ and $g(.)$ represent the PA model and its corresponding post-distortion function, respectively, then $g(f(.))$ reduces the output signal bandwidth to that of the input signal. Therefore, the GST theoretically supports a reduction in the bandwidth of the TOR needed to capture the PA output signal. This signal can be obtained by applying the inverse function of $g(.)$ as given in (2.21); $g(.)$ needed to be known a-priori. However, if a post-distorter exists, then a pre-distorter exists which is exactly equal to the post-distorter (according to the p^{th}-order inverse theorem [91]). Therefore, application of the GST as formulated in [123] to reduce the bandwidth requirement when capturing the PA output signal would require a priori knowledge of the pre-distortion function itself. Since the exact DPD, i.e., the $g(.)$ that meets the GST conditions, is not available beforehand, GST is not a suitable approach to use to reduce the TOR bandwidth requirement.

Below, examples are provided to illustrate that using the GST alone is insufficient to obtain the output of a nonlinear PA. Using a 20 MHz input signal as a test signal for a nonlinear PA yields an output signal of wider bandwidth, i.e., 100 MHz, as shown in Figure 4.1. Therefore, a sampling rate of 2×100 MHz is needed to capture the PA output signal envelope without loss of information according to Nyquist-Shannon theorem [92]. To apply the GST and reduce the TOR bandwidth, knowledge of the exact post-distortion function, $g(.)$, is required so that the output signal of the cascade composed by the PA and the post-distorter can be sampled at a sampling rate equal to 2×20 MHz, as shown in Figure 4.2. Given the equality between the pre-distortion and post-distortion functions,

the GST may be applied to reduce the TOR bandwidth at the PA output as depicted in Figure 4.3. However, in the absence of an exact pre- or post- distortion function, the output signal of the cascade will not experience a bandwidth reduction, as shown in Figure 4.4. In this latter case, the output signal will need to be sampled at the normal sampling rate of $2 \times 100\ MHz$. Note that when using the p^{th}-order inverse, the pre-distorter and the post-distorter are equivalent and will not affect the applicability of the GST.

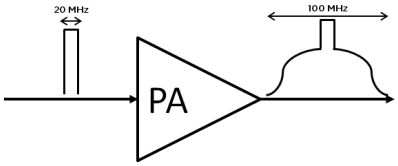

Figure 4.1: Signal representation of a nonlinear power amplifier

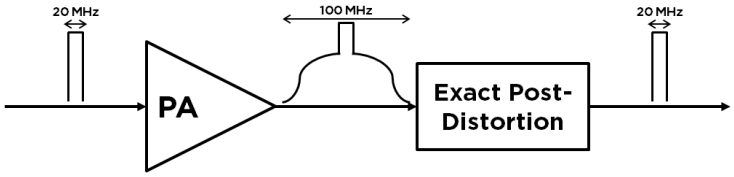

Figure 4.2: Signal representation of a nonlinear power amplifier with exact post-distortion

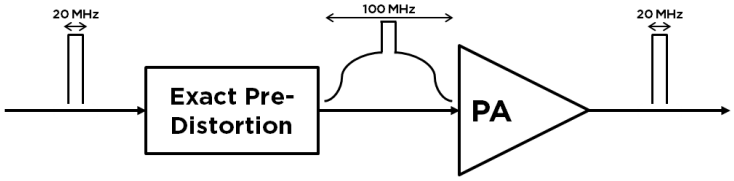

Figure 4.3: Signal representation of a nonlinear power amplifier with exact pre-distortion

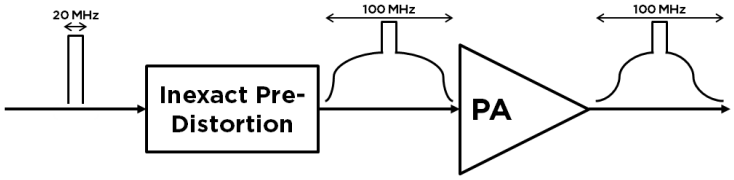

Figure 4.4: Signal representation of a nonlinear power amplifier with inexact pre-distortion

4.2 Bandwidth Requirement in the Transmitter Observation Receiver

In order to enable the GST to sample the output spectrum using reduced rate, a technique is proposed to enable the configuration of Figures 4.1 and 4.4 to attain the conditions satisfied by Figures 4.2 and 4.3. In the following sections, the PA will always be in the configuration of Figure 4.1 or Figure 4.4 unless otherwise stated. It will be shown that a reduced sampling rate of $2 \times INR$ can be used for the linearization of the PA instead of the conventional $5 \times INR$. Theoretical explanations, simulation and measurement results are reported to support the sampling rate reduction claim.

Transmitter Observation Receiver Bandwidth Requirement

Let us consider a PA with a static nonlinear transfer function as given in (4.1), where $x(n)$ and $y(n)$ are the PA input and output signals respectively, P is the nonlinearity order, and a_i are known memoryless coefficients of the PA model. Since the nonlinearity order dictates the bandwidth expansion, a memoryless model was selected for simplicity rather than a model with memory effects. For reasons of clarity, the nonlinearity order will be set to 3, i.e., $P = 3$, initially. Later, it will be shown that the results can be generalized to any order of nonlinearity. Using the p^{th}-order inverse [91], (4.1) suggests that the DPD model can be expressed in the same form, i.e., (4.2) , with b_i being the unknown memoryless coefficients, $u(n)$ the DPD input signal and $x(n)$ the DPD output signal.

$$y(n) = \sum_{p=0}^{\frac{P-1}{2}} a_{2p+1} \left|x(n)\right|^{2p} x(n) \tag{4.1}$$

$$x(n) = \sum_{p=0}^{\frac{P-1}{2}} b_{2p+1} \left|u(n)\right|^{2p} u(n) \tag{4.2}$$

In order to identify the DPD models unknown coefficients, i.e., b_i, when $P = 3$, at least two independent equations are needed, i.e., one for each unknown value. Figures 4.5 and 4.6 illustrate the frequency domain of the PAs input and output signals, with and without the presence of a DPD. In Figure 4.6, while the DPD must be a nonlinear function, it does not have to be exactly the inverse of the PA (Figure 4.4). In either case, the PA input signal is a multi-tone. Figure 4.5 suggests that one of the equations for identifying b_i can be deduced from the nonlinear amplification of the fundamental tones. A second equation

can be determined by the identification of the IMD3 generated at the output of the PA. Note that, theoretically the fundamental tones are amplified differently and thus lead to two equations toward the identification of b_i [89]. However, in practice the theoretical difference between the amplitude of the different amplified tones can be very small and its practical determination can be subject to measurement noise and limited resolution (please refer to Appendix A). Thus, the inclusion of the IMD3 measurement is used to mitigate this problem. The same conclusions remain true for any nonlinearity order. For example, for $P = 7$, measurement of IMD3, fifth order IMD (IMD5) and seven order IMD (IMD7) is required to accurately determine the DPD model coefficients when driven by a multi-tone signal; in other words five equations. However, in practice, while the nonlinearity order of the DPD model is usually set higher than 5, only IMD3 and IMD5 are captured to identify the DPD coefficients. This discrepancy is discussed later.

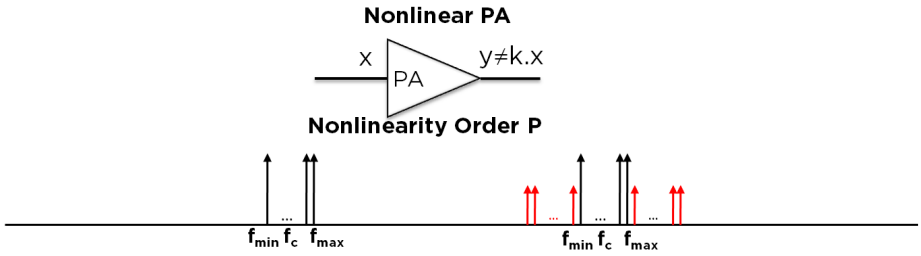

Figure 4.5: Frequency domain representation of the input and output signals of a memoryless nonlinear power amplifier driven with a multi-tone signal

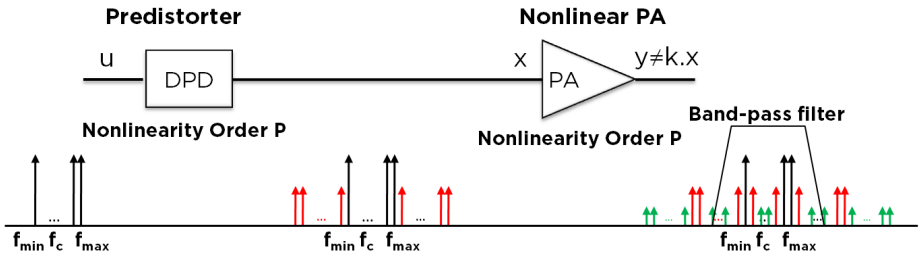

Figure 4.6: Frequency domain representation of the input and output signals of a memoryless nonlinear digital predistortion and power amplifier cascade driven with a multi-tone signal

Returning to the case of $P = 3$, Figures 4.5 and 4.6 show that the output signal bandwidth is larger than the input due to the PA nonlinearity that generates IMDs. However, the same figures show a major difference in the location of the IMDs at the PA output signal. Figure 4.5 shows that all IMD3 are generated outside of the output in-band signal: $[f_{min} \cdots f_{max}]$, where f_{min} and f_{max} are the smallest and highest frequency amplified tones

respectively. Figure 4.6, however, shows that even though the PA nonlinearity is only 3, up to 5 times the input bandwidth is observed at the output of the PA and the IMDs location is now in both the in-band and out-of-band regions, i.e., inside and outside $[f_{min} \cdots f_{max}]$. The presence of IMDs inside the output signal in-band region is due to the presence of tones outside the $[f_{min} \cdots f_{max}]$ band at the input of the PA. The latter generates in-band IMDs at the output of the PA when passing through the nonlinear PA. Therefore, by only measuring the in-band output signal (the amplified multi-tone and the in-band IMDs), more than one independent equation can be deduced. In other words, with an observation of only the in-band output signal, all b_i can be identified and a DPD model can be computed. This allows for a reduction of the required TOR bandwidth at the output of the PA. The previous demonstrates that more IMDs are generated inside and close by the in-band when preceding the PA with a nonlinear function. Therefore more equations can be produced for the identification of the DPD coefficients with reduced TOR bandwidth.

With a large enough number of tones, only in-band observation would be required to characterize the PA nonlinearity and to identify the DPD coefficients. This is due to the large number of IMDs generated in the output in-band region which provide a large number of independent equations to identify the DPD coefficients, b_i. However, in practice, this is not sufficient and an observation of up to twice the input signal bandwidth is usually needed. This is explained in detail in Section 5.3.

Reduction of the Bandwidth Requirement for a Memoryless Nonlinear Power Amplifier

In the proposed technique, the output signal is captured by a receiver with a bandwidth equal to twice the input one. This is still a considerable reduction of the TOR bandwidth from five times to only two times the input signal bandwidth. Several iterations were needed to identify the DPD coefficients when using the reduced TOR bandwidth. For the remainder of this chapter, all linearization performances have been calculated after iterating the DPD coefficients identification ten times.

A simulation case study of a memoryless nonlinear PA was conducted using a three uniformly spaced tones input signal. The nonlinearity of the model was set equal to 9 and even and odd orders were used. The model is given in (4.3).

$$y(n) = \sum_{p=1}^{9} a_p \left|x(n)\right|^{p-1} x(n) \tag{4.3}$$

$$\begin{array}{lll} a_1 \approx -0.56 + 0.84j & a_2 \approx -1.14 + 0.57j & a_3 \approx 15.94 - 8.80j \\ a_4 \approx -95.11 + 50.22j & a_5 \approx 303.09 - 136.99j & a_6 \approx -50.68 + 202j \\ a_7 \approx 575.1 - 168.66 & a_8 \approx -321.89 + 76.1j & a_9 \approx 74.9 - 14.66j \end{array}$$

The frequency components which were far from the carrier frequency (both baseband and harmonics) were neglected. To determine the DPD coefficients b_i, (4.4) was expanded and minimized as follows:

$$e(n) = y(n) - u(n) \tag{4.4}$$

$$where \quad x(n) = \sum_{q=1}^{9} b_q \left|u(n)\right|^{q-1} u(n) \tag{4.5}$$

When (4.4) is expanded using MAPLE software, within a TOR bandwidth equal to twice the bandwidth of the input signal, a total of 11 independent equations were found in the case of Figure 4.5 and 25 in the case of Figure 4.6. One can conclude that a reduced TOR path is more likely to be feasible for the identification of b_i due to the presence of the nonlinear non-convergent DPD function prior to the PA. Therefore, identifying the DPD coefficients is possible using a reduced TOR path in the case of Figure 4.6.

It is important to note that while the TOR bandwidth is reduced at the output of the PA, the input signal of the PA should still be sampled at full rate (sampling equal to $5 \times INR$). In fact, the generation of the IMDs at the input of the PA is what makes the sampling rate reduction possible in the TOR path. It allows the generation of IMDs inside the output in-band and close to the output in-band region. In addition, the nonlinear predistorter generates IMDs at the input of the PA both inside of and outside the observable band. As the PA is also a nonlinear function, some of the IMDs outside of the observable band will generate new sets of IMDs in the observable band, as shown Figure 4.6. Therefore, by capturing only the observed band, information about the PA behavior outside the observable band can be found. Consequently, modeling and linearization using reduced TOR bandwidth is possible even when the PA is non-consistent outside the observable band. In other words, a reduced TOR bandwidth is only possible when using a full speed DAC ($5 \times INR$) at the input of the PA. This is in contrast to the band-limited Volterra series [114] where both the input and the output of the PA are sampled at the

reduced rate. Actually, the reduced rate of the input signal in the band-limited Volterra series yields to linearization only within the observed band; no linearization is possible elsewhere. However, using the proposed approach, linearization will also be possible outside of the observed band. This hypothesis is supported by simulation and measurement results reported in the following sections.

4.3 Necessity of Twice the Input Nyquist Rate at the Transmitter Observation Receiver

Above, it was shown that within a bandwidth equal to that of the input signal, enough signal information at the output is present to linearize the PA output signal. However, due to potential modeling error in selecting the right kernels for the DPD (selecting kernels that do not represent the inverse of the PA model), the sampling rate should be set to $2 \times INR$ instead of $1 \times INR$. In this section, an explanation of this claim is provided.

The DPD structure is thought to be correct/appropriate when the right kernels are selected for the DPD to represent exactly the inverse of the PA. Conversely, the incorrect DPD structure means that either different, more or fewer kernels have been selected for the DPD model. As an example, if a PA model is given in (4.6), the correct DPD model should be similar to (4.7). Thus, when selecting the kernels of (4.7) to be identified, it is said that the correct DPD structure is selected. Whenever any other, fewer or more kernels are selected to be identified as a DPD model, e.g., as in (4.8) or (4.9), it is said that the incorrect DPD structure has been selected. Note that, the harmonics are not considered in the DPD system as the focus is on the signal around the carrier frequency and the harmonics have been previously filtered.

$$y(n) = a_1 x(n) + a_3 x(n-1)x(n-2)x(n) \tag{4.6}$$

$$x(n) = b_1 u(n) + b_3 u(n-1)u(n-2)u(n) \tag{4.7}$$

$$x(n) = b_1 u(n) + b_5 u(n-1)u(n-2)u^3(n) \tag{4.8}$$

$$x(n) = b_1 u(n) + b_{3,1} u(n-1)u^2(n) + b_{3,2} u^3(n) \tag{4.9}$$

In the case where a correct DPD structure is selected, a sampling rate equal to $1 \times INR$ can be used for the linearization of the PA output as explained before. The coefficients identified using the in-band information will ensure the construction of a PA inverse model that covers the out-of-band region too. However, if an incorrect DPD structure is used,

the coefficients identified using the in-band information will not guarantee good accuracy in the out-of-band region or the exact PA inverse model. In such a case, the constructed DPD will only fit the inverse of the PA within the in-band region and deviates outside, (please refer to appendix B for an illustration with an example). In practice, it is almost impossible to come up with the correct DPD kernels for the following reasons:

- Complexity: only a limited number of kernels are allowed, i.e., there will be always some kernels missing that perfectly describe the inverse of the PA.
- Modeling error: due to measurement noise and test setup calibration.

In a simulation environment (modeling error free mode) and using a memoryless model (reduced complexity mode), it has been shown that when a correct DPD structure is found, a sampling rate equal to $1 \times INR$ suffices to linearize the output of the PA model. Let us consider the case of a memoryless PA as was given in (4.3) with $P = 7$. The DPD structure is easy to find and is equal to a memoryless polynomial with $P \approx 7$. In this case, identification of b_i can be done while sampling at a rate equal to $1 \times INR$ and linearization of the PA output is achieved. Figure 4.7 shows the case of a nonlinear PA with and without DPD where the input signal is a 20 MHz 4C WCMDA. The DPD was constructed using the conventional approach with the sampling rate is equal to $5 \times INR$, and with a reduced sampling rate equal to $1 \times INR$. It is clear from Figure 4.7 that both DPD approaches were able to linearize the output signal of the PA. It can be concluded then that when the structure of the DPD is known a priori, sampling the output signal at $1 \times INR$ is sufficient.

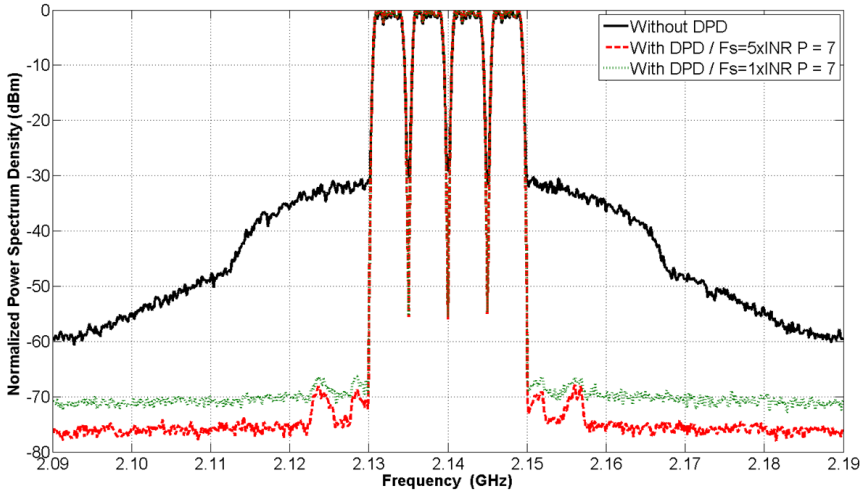

Figure 4.7: Power amplifier output spectrum without digital predistortion and with a memoryless digital predistortion of nonlinearity order equal to 7 with $Fs = 1 \times INR$ and $Fs = 5 \times INR$

When the DPD structure is not chosen correctly, i.e., in the cases of over/under modeling or incorrect choice of the set of kernels, and the sampling rate is set to $1 \times INR$, it is anticipated that the linearization capability will be very limited. Conversely, if the sampling rate is increased to $2 \times INR$, even if the correct DPD structure has not been selected, good linearization will be achieved. Figure 4.8 shows the linearization capabilities of a memoryless DPD model when $N = 13$ (incorrect DPD structure due to over modeling). It is clear from Figure 4.8 that when the sampling rate at the output of the PA was set to $5 \times INR$, a good reduction of the output spectrum was observed. However, when the output sampling rate was reduced to $1 \times INR$, the linearization capabilities degrade as there were uncorrelated IMDs inside and outside the in-band region. In this case, due to over modeling, accurate linearization can occur only for the inside IMDs. Since more degrees of freedom than needed are present in the model, there is no guarantee that the out of band IMDs will be deduced. Consequently, limited linearization was recorded. However, when the output sampling rate was increased to $2 \times INR$, some of the outside IMDs were captured and accurate linearization was achieved inside and outside the observation path as shown in Figure 4.8. It is important to note that, when the TOR bandwidth is set equal to the targeted linearized band, increasing the nonlinearity order will not degrade the linearization capability.

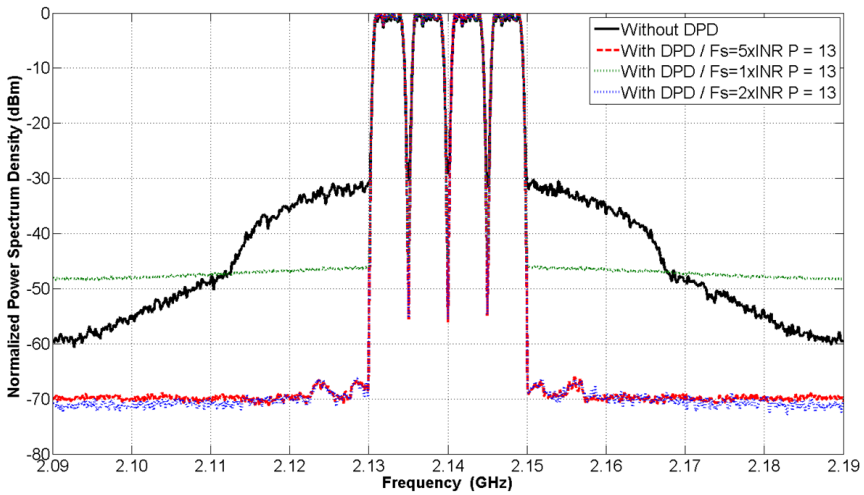

Figure 4.8: Power amplifier output spectrum without digital predistortion and with a memoryless digital predistortion of nonlinearity order equal to 13 with $Fs = 1 \times INR$, $Fs = 2 \times INR$ and $Fs = 5 \times INR$

Figure 4.9 shows the ACLR performance for the memoryless model against the TOR bandwidth (output sampling rate) and nonlinearity order of the DPD. It can be deduced that when a correct DPD structure is selected, ($P \approx 7$), a reduction of the sampling rate to $1 \times INR$ is possible without trading-off the linearization capabilities. However, as

soon as the nonlinearity order begins to drift from $P \approx 7$, the linearization capabilities are degraded if the sampling rate is decreased to less than $2 \times INR$. Note that, for cases where $P < 7$, the performance is limited even for high sampling rates due to under modeling.

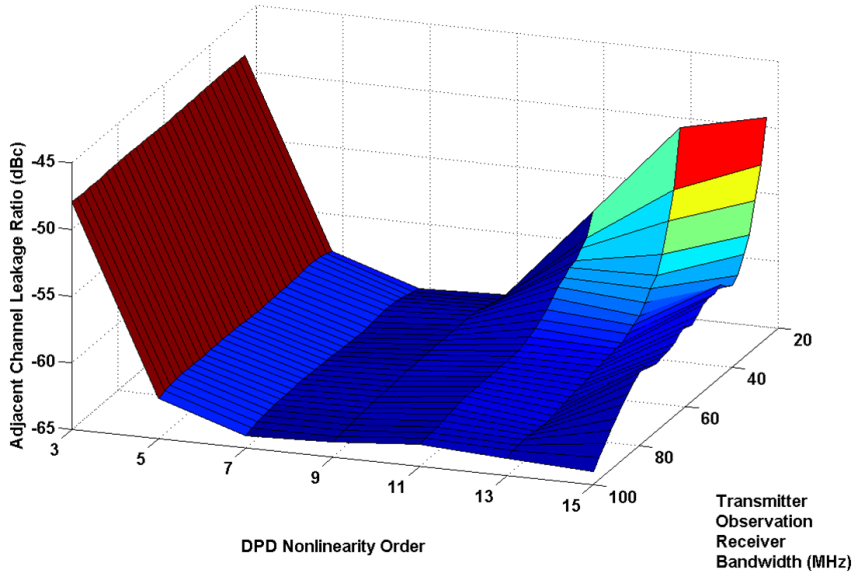

Figure 4.9: Adjacent channel leakage ratio versus the digital predistortion nonlinearity order and the transmitter observation receiver bandwidth

In conclusion, Figure 4.9 demonstrates that if the correct DPD structure is selected, any sampling rate equal to or greater than $1 \times INR$ can be deployed without trading-off the linearization capabilities. However, when the correct DPD structure is not precisely known, the sampling rate has to be increased to at least $2 \times INR$ in order to achieve good linearization. For realistic PAs, memory effects are introduced in addition to the static nonlinearity and therefore a correct DPD structure is more challenging and complex to determine. There is no guarantee that the correct DPD kernels can be identified, especially when the PA is driven with wideband signals and thus exhibiting strong memory effects. For this reason, in our analysis, the reduced sampling rate was relaxed to $2 \times INR$ to improve the linearization capabilities. This is still a considerable reduction of the ADC speed requirement as a $2/5$ reduction rate is achieved compared with the conventional approach.

Since, it is established that selecting the right kernels for the DPD is crucial for the reduction of the TOR bandwidth, the G-Functionals pruned Volterra series described in Chapter 3 is used for a better approximation of the correct DPD structure. Consequently, the DPD structure that describes best the inverse of the PA can be found and a reduction of the TOR bandwidth will be more likely possible.

4.4 Digital Predistortion with Reduced Transmitter Observation Receiver Bandwidth Identification Algorithm

In this section, the pruned Volterra series using G-Functionals is used for the validation of the reduced TOR bandwidth approach in simulation using multi-tone and modulated signals. First, three-, four- and five-tones signals were used as input signals to the memoryless nonlinear PA model given in (4.3). The sampling rate of the output of the PA was set to $2 \times INR$ instead of the usual $5 \times INR$. In other words, if the input signal bandwidth was set to 20 MHz, the required ADC sampling rate would be only 80 Msps (40 MHz bandwidth observation) instead of 200 Msps (100 MHz bandwidth observation). The output signal was then up-sampled by a factor of $5/2$ to match the bandwidth of the input of the PA (the predistorted signal) and the DPD coefficients were identified. Ten iterations of this process were performed. After sufficient iterations, the DPD aligned with the inverse of the PA behavior and the results are discussed hereafter. The previous statements are summarized in the algorithm 1 and Figure 4.10.

initialize $g_0(u(n))$ to a random nonlinear expansion function ;
for *k from 1 to 10* **do**
compute $x_k(n) = g_{k-1}(u(n))|F_s = 5 \times INR$;
modulate discrete $x_k(n)$ to continuous $x_k(t)$;
up-convert $x_k(t)$ to RF ;
down-convert the PA output $y_k(t)$ to IF ;
filter $y_k(t)$ by a band-pass filter having a bandwidth equal to twice the bandwidth of the input signal ;
sample continuous $y_k(t)$ to discrete $y_k(t)|F_s = 2 \times INR$;
up-sample $y_k(n)$ by $5/2$;
identify DPD parameters with LSE ;
$B = \left(Y^T Y\right)^{-1} Y^T X = \left[h_1^i(\sigma_1) \cdots h_P^i(\sigma_1, \cdots, \sigma_P)\right]^T$;
where $X = \left[x_k(n) \cdots x_k(n-M)\right]^T$ — M: memory depth ;

$$Y = \begin{bmatrix} y_k(n-\sigma_1) & \cdots & y_k(n-\sigma_1)\cdots y_k(n-\sigma_{n_p}) \\ \vdots & \ddots & \vdots \\ y_k(n-M-\sigma_1) & \cdots & y_k(n-M-\sigma_1)\cdots y_k(n-M-\sigma_{n_p}) \end{bmatrix}$$

end

Algorithm 1: Identification algorithm for digital predistortion with reduced transmitter observation receiver bandwidth

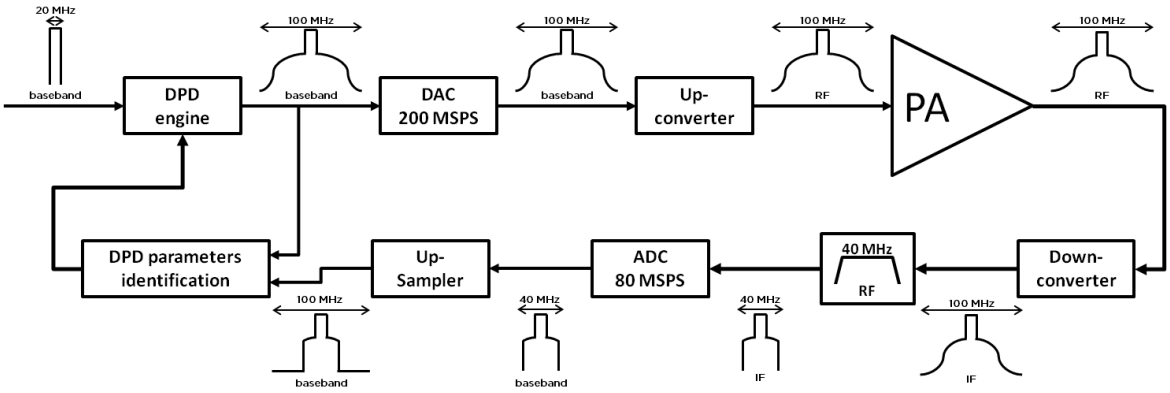

Figure 4.10: Identification process of the digital predistortion function using a reduced transmitter observation receiver bandwidth

Figure 4.10 shows the block diagram of the DPD+PA cascade using a reduced TOR bandwidth. The example is provided for a 20 MHz bandwidth input signal. The full rate DPD generates a 100 MHz predistorted signal from a five times oversampled input signal. A 200 MHz DAC converts the digital signal to analog and an up-converter makes it RF. The PA produces an output signal of a bandwidth equal to 100 MHz. Since the purpose of this chapter is to demonstrate the use of a reduced TOR bandwidth, the signal goes through a band-pass filter with a bandwidth equal to 40 MHz. This allows the capture of the output signal without aliasing. Indeed, although the signal has a bandwidth of 100 MHz, the ADC speed is only set to 80 Msps allowing a capture of only 40 MHz. To avoid any aliasing problem, the signal is first filtered to remove any unnecessary spectrum then sent to the ADC. Up-sampling of the digitized output signal is carried out to ensure the matching between the sampling rates of the samples representing the PA input and output signals (2×100 Msps). This does not recover the filtered spectrum band but rather augments the sampling rate of the output signal to match the full speed rate of the DPD. The input and output of the PA are then used to identify the DPD using indirect learning. Finally, the coefficients of the DPD are updated and another iteration of the previous steps is begun.

Figures 4.11 and 4.12 show the worst ACLR for the equally spaced and non-uniformly spaced multi-tone input signals for both the DPD with reduced TOR bandwidth and a conventional DPD with full bandwidth. It is clear that the linearization performance starting with four tones for equally spaced multi-tone signals, and beginning at three for non-uniformly spaced multi-tone signals, is the same for both the DPD with full bandwidth and DPD with reduced TOR bandwidth. In fact, as the number of tones is increased, more IMDs are generated in the in-band and close to the in-band output spectrum region. Thus, more equations for identifying b_i are generated. This suggests the applicability of

the reduced TOR bandwidth technique for modulated signals, as these signals are as rich as multi-tone signals. Next, simulation and measurement results are presented for 20 MHz and 60 MHz modulated wideband signals.

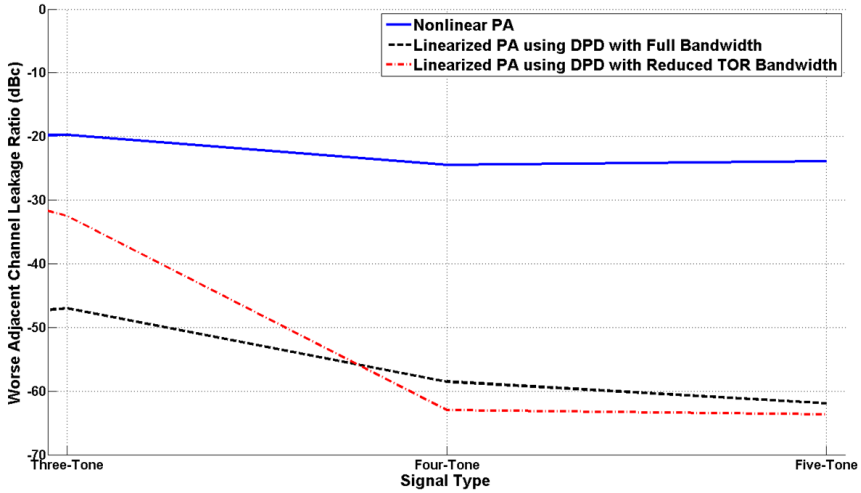

Figure 4.11: Adjacent channel leakage ratio of the output signal of the power amplifier in Figure 4.6 under equally spaced multi-tone stimuli

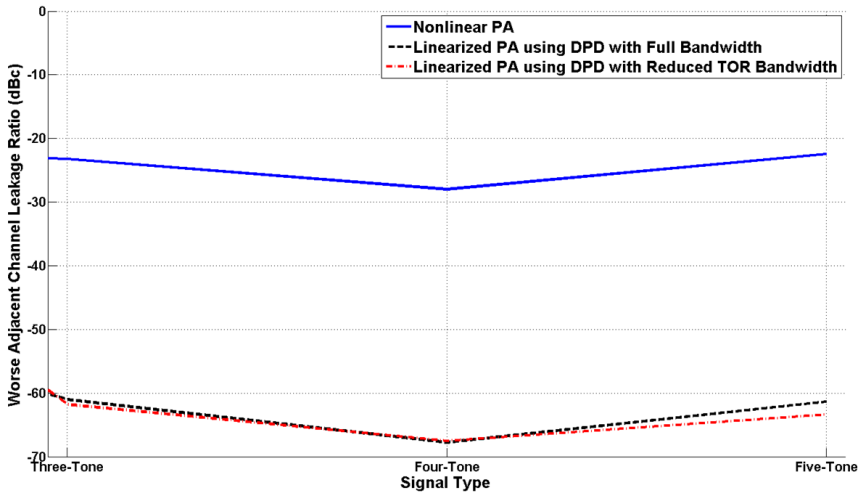

Figure 4.12: Adjacent channel leakage ratio of the output signal of the power amplifier in Figure 4.6 under non-uniformly spaced multi-tone stimuli

4.5 Digital Predistortion with Reduced Transmitter Observation Receiver Bandwidth Validation

4.5.1 Simulation Results

Simulation with a nonlinear dynamic PA, with an output signal having an ACLR of about -30 dBc, was used. The input signal was set to 20 MHz 4C WCDMA. The linearization of this PA was performed with an output observation path set to 40 MHz (the proposed reduced TOR bandwidth). For comparison, a 40 MHz band-limited Volterra series was used in addition to the conventional DPD with full and limited-bandwidth. The four DPD configurations are given in Table 4.1. To contextualize the results outlined in Table 4.1, the DPD with full-bandwidth uses full rate ADC (speed equal to $5 \times INR$) to capture the output signal with a bandwidth equal to five times the bandwidth of the input signal. It also uses a full rate DAC (speed equal to $5 \times INR$) to generate a predistorted signal with a bandwidth equal to five times the original input signal. Finally, the model identifies the DPD coefficients when the time step is $T = {}^{1}/_{F_s} = {}^{1}/_{(5 \times INR)}$.

Table 4.1: Different linearization digital predistortion schemes

	DAC Speed	ADC Speed	DPD Speed
Conv DPD with full bandwidth	$5 \times INR$	$5 \times INR$	$5 \times INR$
Conv DPD with limited bandwidth	$2 \times INR$	$2 \times INR$	$2 \times INR$
Band-Limited Volterra series DPD	$2 \times INR$	$2 \times INR$	$5 \times INR$
DPD with reduced TOR bandwidth	$5 \times INR$	$2 \times INR$	$5 \times INR$

Figure 4.13 shows the output PSD of the nonlinear PA with and without DPDs. The four DPDs were iterated ten times for convergence, especially relevant when using the reduced TOR bandwidth of 40 MHz. The band-pass filter appears in the figures that follow, when applicable, to illustrate the TOR bandwidth. On the one hand, after ten iterations, the reduced TOR bandwidth DPD approach allowed an adjacent channel power ratio (ACPR) of less than -50 dBc, i.e., met the WCDMA mask [1]. On the other hand, the DPD with an output signal sampled at 100 MHz had faster convergence and allowed an ACPR reduction of about -60 dBc. However, the band-limited Volterra series was able to linearize the PA only within the observed band; no linearization was achieved elsewhere. Finally, the band-limited conventional DPD showed no linearization outside the observed

band and poor linearization within. Thus, it can be concluded that under simulation the reduced TOR bandwidth succeeded in linearizing the PA output signal within and outside the observation path but achieved slightly less ACPR reduction than the DPD with full bandwidth. Nevertheless, it should be pointed out that the sampling rate was lowered by more than half, thereby trading-off some ACPR reduction while still passing the WCDMA mask. These results were replicated under measurement conditions which are reported in the next subsection.

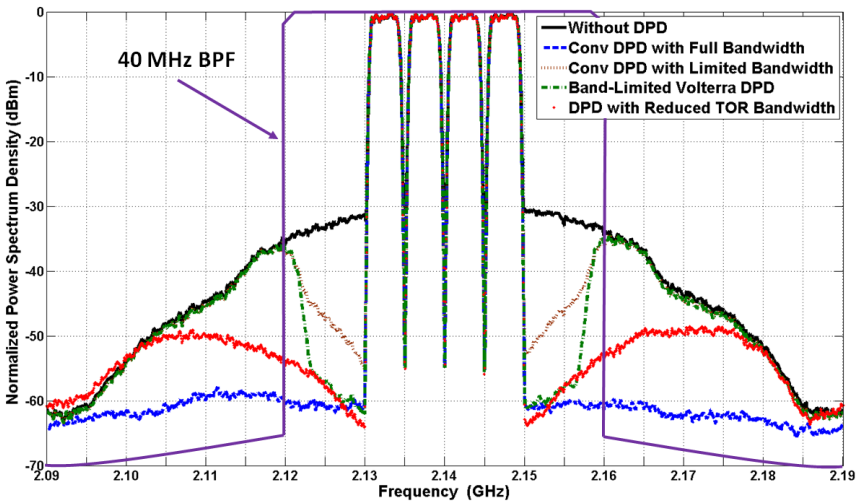

Figure 4.13: Output spectrum of the power amplifier driven with a 20 MHz 4C WCDMA signal with and without the different digital predistortion schemes of Table 4.1

The same conclusions were found in the case where the input signal bandwidth was 60 MHz as shown in Figure 4.14. The conventional DPD with full bandwidth and the DPD with reduced TOR bandwidth were able to linearize the PA for more than -50 dBc, meeting the spectrum mask. Similarly, the band-limited Volterra series could only linearize inside the observed band. The conventional DPD with limited bandwidth achieved poor linearization performance even within the observed band. Therefore, one can conclude that using the DPD with reduced TOR bandwidth, linearization of a given PA is possible inside and outside the observed band without a severe degrading of the linearization performance.

Note that the DPD with reduced TOR bandwidth and the band-limited Volterra series model achieved better ACPR next to the in-band signal. In fact, when sampling at large time step (small TOR bandwidth), both the receiver noise figure and flatness are better than when using small time step (large TOR bandwidth). Therefore, the sampling of the output signal close to the in-band region was better in the case of the proposed model with reduced TOR bandwidth and the band-limited Volterra series DPD when compared to the case of conventional DPD with full bandwidth. Consequently, better linearization

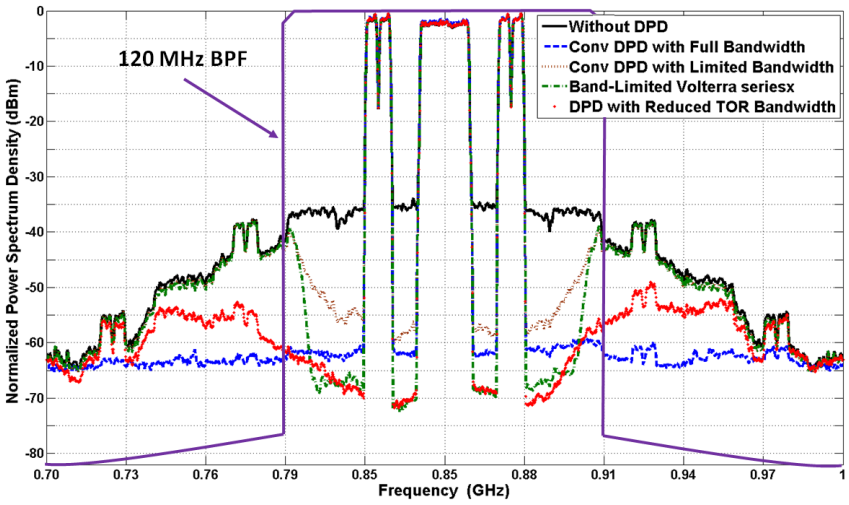

Figure 4.14: Output spectrum of the power amplifier driven with a 60 MHz mixed LTE signal with and without the different digital predistortion schemes of Table 4.1

was recorded in that region.

4.5.2 Measurement Results

In this subsection, measurement results are presented to prove the applicability of the proposed approach in reducing the required bandwidth of the TOR. Two different PA architectures were used (GaN single-ended [105] and GaN Doherty PA [106]), with two different signal bandwidths (20 MHz and 60 MHz), in order to prove the generalizability of the DPD with reduced TOR bandwidth. As was done for the simulation results, the DPDs with reduced TOR bandwidth performance was compared with the band-limited Volterra series and conventional DPDs with limited and full bandwidth. ACPR and EVM measurements were used to evaluate the different DPD performances. The ACPR is given in (4.10), where P_{ch} is the average power in the channel (the in-band power) and $P_{adj}(offset)$ is the average power of the adjacent channel centered at $f_c \pm offset$. Note that for the 20 MHz signal, the average power is computed over 5 MHz bandwidth, while for the 60 MHz signal the average power is computed over a 15 MHz bandwidth. The EVM is given in (4.11) where $u(n)$ and $y(n)$ are the PA input and output signals respectively, N is the number of samples (set to $100e3$), and SSG is the small signal gain of the PA stages (main PA and drivers). Note that different sets of samples are used during the DPD coefficients identification and performance evaluation (EVM and ACPR).

$$\text{ACPR} = 10log_{10}\left(\frac{P_{adj}(\text{Offset})}{P_{ch}}\right) \tag{4.10}$$

$$EVM = 10log_{10}\left(\frac{\sum_{n=1}^{N}\left|u(n) - \frac{y(n)}{SSG}\right|^2}{\sum_{n=1}^{N}|u(n)|^2}\right) \tag{4.11}$$

Figure 4.15 illustrates the measurements setup that allows for PA DUT modeling/linearization up to 300 MHz. It included an I/Q modulation generator of 1 Gsps (R&S®AFQ100A) and an RF source up to 12.75 GHz (R&S®SGS100A) on the transmission side. A signal and spectrum analyzer of 320 MHz analysis bandwidth (R&S®FSW) was used in the feedback path. All equipments were locked to the same 10 MHz reference and the transmitter path triggered the feedback path. A computer equipped with MATLAB controlled the instruments and performed time alignment, up-sampling, DPD coefficients identification, predistorted signal generation and linearization performance calculations.

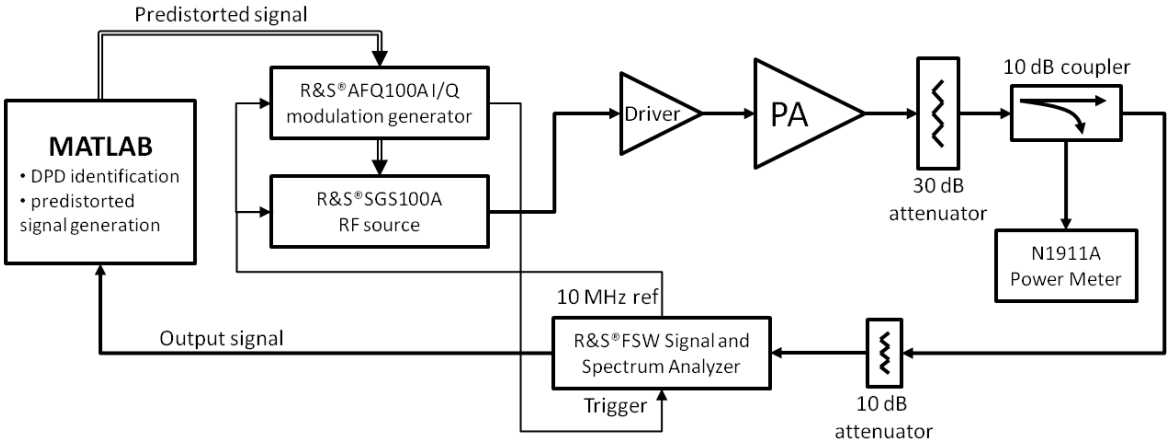

Figure 4.15: Characterization setup of the digital predistortion with reduced transmitter observation receiver bandwidth

Figure 4.16 depicts the linearized output spectrum of the GaN single-ended PA, driven with a 20 MHz 4C WCDMA signal, under the different DPD schemes as given in Table 4.1. Figure 4.16 shows a significant reduction in the out-of-band spectrum emission when deploying the conventional DPD with full bandwidth. About -42 dB of EVM was achieved when deploying the conventional DPD with full bandwidth as shown in Table 4.2, and less than -51 dBc of ACPR was recorded according to Table 4.3. It is clear from Figure 4.16 and Table 4.3 that when using the conventional DPD with limited-bandwidth, the linearization performance dropped dramatically, especially in terms of ACPR. For instance,

at the second offset, the DPD did not reduce the out-of-band spectrum at all (-34 dBc). The band-limited Volterra series reduced the spectrum regrowth within the observation bandwidth; however, there was still spectrum regrowth outside the observed band. Actually, the band-limited Volterra DPD almost matched the performance of the conventional DPD with full bandwidth inside the observed band, but did not improve the signal quality outside of the band. The DPD with reduced TOR bandwidth, however, showed linearization performance close to those of the conventional DPD with full bandwidth. In fact, about -50 dBc ACPR was maintained within and outside the observed bandwidth.

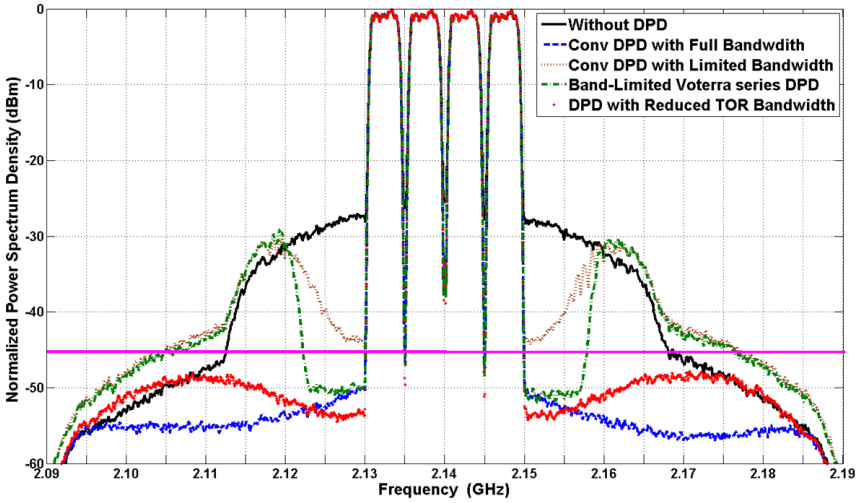

Figure 4.16: Output spectrum of the GaN single-ended power amplifier driven with a 20 MHz 4C WCDMA signal with and without the different digital predistortion schemes of Table 4.1 when $Fs = 2 \times INR$

The sampling rate $2 \times INR$ allowed linearization inside and outside the observable band. The linearized output signal respected the spectrum mask [1] and achieved better performance than the band-limited and conventional DPD with limited bandwidth. In the following section, an attempt to reproduce the exact performance of the full bandwidth DPD is provided with a small increase of the sampling rate to $2.5 \times INR$. Figure 4.17 shows a comparison between the PA output PSDs when deploying the different DPD linearizers when the output sampling rate is $2.5 \times INR$. It can be concluded that the conventional DPD with limited-bandwidth failed to linearize the PA output outside the observed band and only achieved limited linearization inside the observed band. The band-limited Volterra series DPD reduced the spectrum regrowth only within the observation bandwidth and did not improve the signal quality outside of the observed band. However, the DPD with reduced TOR bandwidth linearization capabilities were the same than those of the DPD with full bandwidth as shown in Table 4.3. About -53 dBc was maintained

Table 4.2: Error vector magnitude at the output signal of the power amplifiers

Model Type \ Signal type		20 MHz	60 MHz
Without DPD		-20.92	-21.39
Conv DPD with Full Bandwidth	$5.0 \times INR$	-41.93	-38.64
Conv DPD with Limited Bandwidth	$2.0 \times INR$	-29.01	-33.66
Conv DPD with Limited Bandwidth	$2.5 \times INR$	-35.29	-35.32
Band-Limited Volterra Series DPD	$2.0 \times INR$	-29.96	-34.29
Band-Limited Volterra Series DPD	$2.5 \times INR$	-35.70	-37.15
DPD with Reduced TOR Bandwidth	$2.0 \times INR$	-39.13	-34.31
DPD with Reduced TOR Bandwidth	$2.5 \times INR$	-39.93	-37.54

Table 4.3: Measured adjacent channel power ratio at the output of the GaN single-ended power amplifier Driven with a 20 MHz 4C WCDMA signal

Model Type \ Offset		12.5 MHz	17.5 MHz	22.5 MHz	27.5 MHz
Without DPD		-27.09	-29.28	-32.69	-42.84
Conv DPD with Full Bandwidth	$5.0 \times INR$	-50.42	-52.30	-53.64	-54.64
Conv DPD with Limited Bandwidth	$2.0 \times INR$	-41.97	-34.18	-30.83	-39.19
Conv DPD with Limited Bandwidth	$2.5 \times INR$	-47.19	-46.17	-40.65	-39.93
Band-Limited Volterra Series DPD	$2.0 \times INR$	-49.70	-44.25	-30.41	-39.69
Band-Limited Volterra Series DPD	$2.5 \times INR$	-52.19	-52.22	-44.92	-39.08
DPD with Reduced TOR Bandwidth	$2.0 \times INR$	-52.70	-51.33	-49.28	-47.92
DPD with Reduced TOR Bandwidth	$2.5 \times INR$	-52.17	-52.53	-53.46	-54.68

across 100 MHz at the output of the PA although only 50 MHz of output signal was captured and used during the DPD coefficients identification. This demonstrates that increasing the sampling rate beyond $2.5 \times INR$ is not necessary.

To demonstrate the applicability of the DPD with reduced TOR bandwidth to wideband signals, a 60 MHz bandwidth signal was used, namely a mixed LTE and WCDMA signal as given in Figure 4.18. The PAPR of the signal was 10.5 dB and the GaN Doherty PA

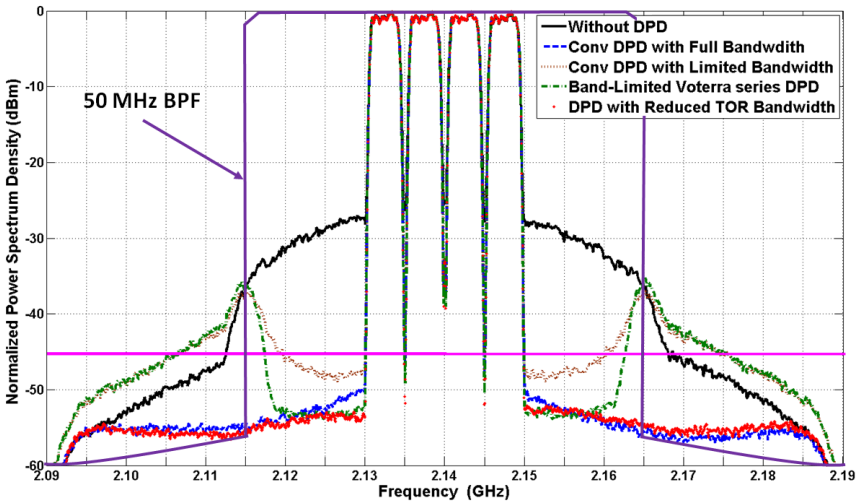

Figure 4.17: Output spectrum of the GaN single-ended power amplifier driven with a 20 MHz 4C WCDMA signal with and without the different digital predistortion schemes of Table 4.1 when $Fs = 2.5 \times INR$

was driven to 2 dB compression point. The PA carrier frequency was set to 850 MHz. Figure 4.18 shows the PA output spectrums with and without DPDs. The different DPD schemes, shown in Table 4.1, were used to linearize the PA and the results are reported in Figures 4.18 and 4.19 and Tables 4.2 and 4.4. First, the TOR bandwidth was set to 120 MHz (twice the input signal bandwidth) and then increased to 150 MHz (2.5 times the input signal bandwidth).

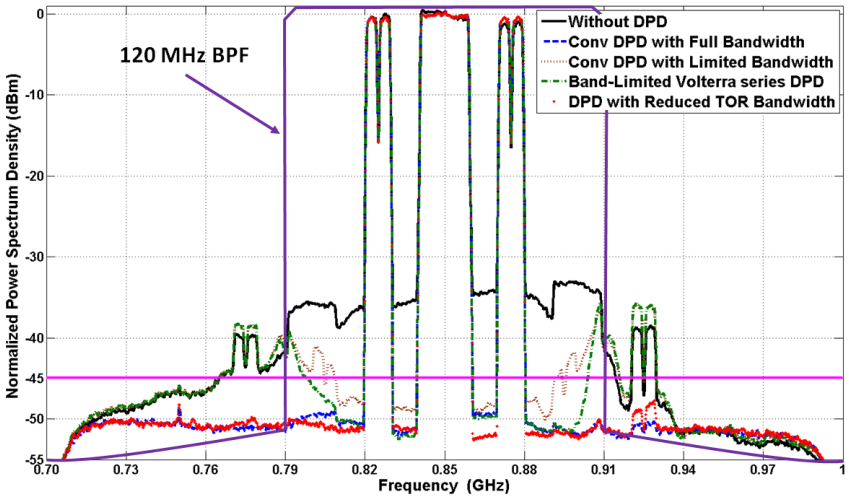

Figure 4.18: Output spectrum of the GaN Doherty power amplifier driven with a 60 MHz mixed LTE signal with and without the different digital predistortion schemes of Table 4.1 when $Fs = 2.5 \times INR$

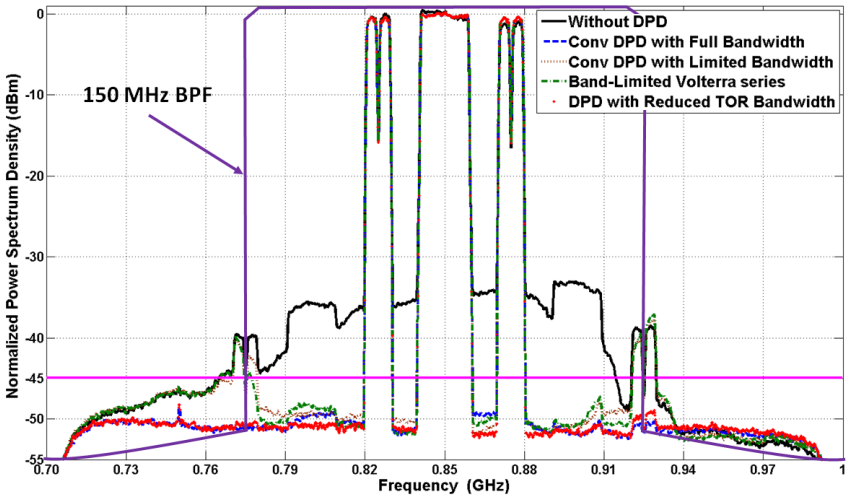

Figure 4.19: Output spectrum of the GaN Doherty power amplifier driven with a 60 MHz mixed LTE signal with and without the different digital predistortion schemes of Table 4.1 when $Fs = 2.5 \times INR$

Table 4.4: Measured adjacent channel power ratio at the output of the GaN Doherty power amplifier driven with a 60 MHz 4C WCDMA signal

Model Type \ Offset		37.5 MHz	52.5 MHz	67.5 MHz	82.5 MHz
Without DPD		-35.82	-35.13	-43.07	-45.14
Conv DPD with Full Bandwidth	$5.0 \times INR$	-50.03	-50.52	-51.18	-51.03
Conv DPD with Limited Bandwidth	$2.0 \times INR$	-46.58	-41.62	-41.50	-44.31
Conv DPD with Limited Bandwidth	$2.5 \times INR$	-49.79	-49.57	-46.92	-45.35
Band-Limited Volterra Series DPD	$2.0 \times INR$	-49.82	-44.81	-41.33	-44.31
Band-Limited Volterra Series DPD	$2.5 \times INR$	-49.95	-49.04	-48.18	-44.94
DPD with Reduced TOR Bandwidth	$2.0 \times INR$	-50.90	-50.83	-50.64	-50.45
DPD with Reduced TOR Bandwidth	$2.5 \times INR$	-50.71	-50.91	-50.93	-50.60

Figure 4.18 shows that the conventional DPD with full bandwidth linearized the output spectrum for less than -50 dBc along the 300 MHz bandwidth. Similarly to the previous experiments, the DPD with limited bandwidth failed to linearize the output spectrum PA outside the observed band and achieved limited linearization within the observed band. The band-limited Volterra series succeeded in linearizing the PA within the observed band but did not reduce the spectrum regrowth beyond 120 MHz (outside the observed band).

However, it is clear from Figure 4.18 and Table 4.4 that the DPD with reduced TOR bandwidth linearized the output of the PA for the entire 300 MHz bandwidth at about 50 dBc. Therefore, using only reduced ADC speed, the DPD with reduced TOR bandwidth was able to show similar performance to the conventional DPD with full bandwidth. The DPD with reduced TOR bandwidth resulted in an EVM of about -35 dB. It can be concluded therefore, that the proposed technique to identify the DPD coefficients succeeded at fully linearizing the PA output IMDs even though only limited spectrum is observed at the output of the PA. Unlike the band-limited Volterra series, where a band-pass filter is necessary to remove all the remaining distortion, the DPD with reduced TOR bandwidth does not involve any additional linearization steps.

When the sampling rate was increased to $2.5 \times INR$, in an effort to improve the linearization performance of the DPD models, Figure 4.19 shows that the DPD with reduced TOR bandwidth now matches the conventional DPD with full bandwidth performance in both EVM and ACPR (shown in Tables 4.2 and 4.4). The band-limited Volterra series and conventional DPD were still not linearizing outside the observed band and the band-limited conventional DPD had poor linearization even within the 150 MHz observed band.

4.6 Nonlinear System Behavioral Modeling Using Reduced Transmitter Observation Receiver Bandwidth

Traditionally, to construct a PA behavioral model, the PA is first driven with a modulated signal while a TOR is used to capture the output signal. Then, a low pass equivalent Volterra series behavioral model is identified to describe the input to output relationship. This model can then be used with CAD tools to predict both the response of the PA and its impact on overall RF system performances. However, since the PA is a nonlinear system, its output signal has a bandwidth that is typically five times that of the input signal. Therefore, in order to construct the corresponding low pass equivalent Volterra series model, the TOR is required to have a bandwidth equal to at least five times the input signal bandwidth. This requirement is more stringent when deploying wider bandwidth signals. For example, to accurately model the PA when using advanced-LTE (where the modulation bandwidth can reach 100 MHz), a TOR with a bandwidth equal to at least 500 MHz would be required. Not only would this result in a frequency dispersive receiver across the signal bandwidth, the cost and power consumption of the TOR would be excessive and not feasible for certain applications.

In the Section 4.1, it was shown that the GST is a potential solution for capturing the output signal of a nonlinear PA with a received that has a bandwidth less than the output signal bandwidth under the condition of identifying an appropriate function $g(.)$ that compresses the output spectrum of the PA. Throughout Sections 4.2-4.4, it was shown that a DPD with reduced TOR bandwidth is possible to linearize the output spectrum of the PA and thus determine the function $g(.)$. Combining these two results, construction of an accurate behavioral model using a reduced TOR bandwidth is possible and could predict the output response of the PA inside and outside the TOR bandwidth as it is shown hereafter. First, a DPD is identified using a reduced TOR bandwidth. This determine the function $g(.)$ and thus enables the application of the GST and the output of the PA can be accurately predicted using (2.21). Two DPD schemes are used, namely the band-limited Volterra series and the proposed reduced TOR bandwidth Volterra series pruned using G-Functionals. The Behavioral model identification process is illustrated in the flow chart of Figure 4.20 [72].

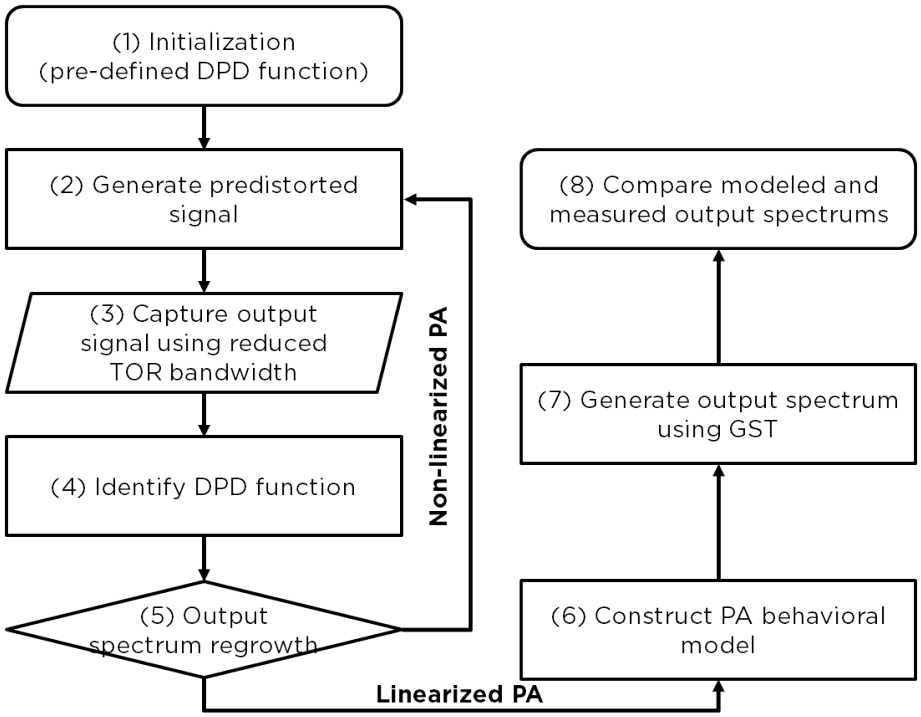

Figure 4.20: Behavioral model identification approach

A 250 W laterally diffused metal oxide semiconductor (LDMOS) Doherty PA was used as the DUT. The PA was driven with a 20 MHz 4C WCDMA input signal into compression resulting in an output signal with a bandwidth of about 92.16 MHz. An N9030A PXA signal analyzer was used to capture the output signal with analysis bandwidths set to 27.65, 36.86, 55.30, and 92.16 MHz.

The TOR bandwidth was first set to 92.16 MHz to prove the validity of the low pass equivalent Volterra series model in predicting the output spectrum. It is clear from Figure 4.21 that the NMSE and the power spectrum density error (PSDE) showed excellent modeling capability. In fact, with or without DPD, the modeling performances are equal since the full output PA spectrum is captured. However, as the TOR bandwidth decreased, the modeling of the PA response became inaccurate and the modeled spectrum get far from the actual spectrum, especially outside of the observed band. For example, Figure 4.21 shows an NMSE of only -31 dB and a PSDE of only -40 dB when the TOR bandwidth was set to 27.65 MHz. This illustrates again that the GST cannot be applied unless the DPD is first identified. To support this assertion, the application of (2.21) when identifying first the DPD, i.e., the function $g(.)$, will guarantee an accurate recovery of the output spectrum shown hereafter.

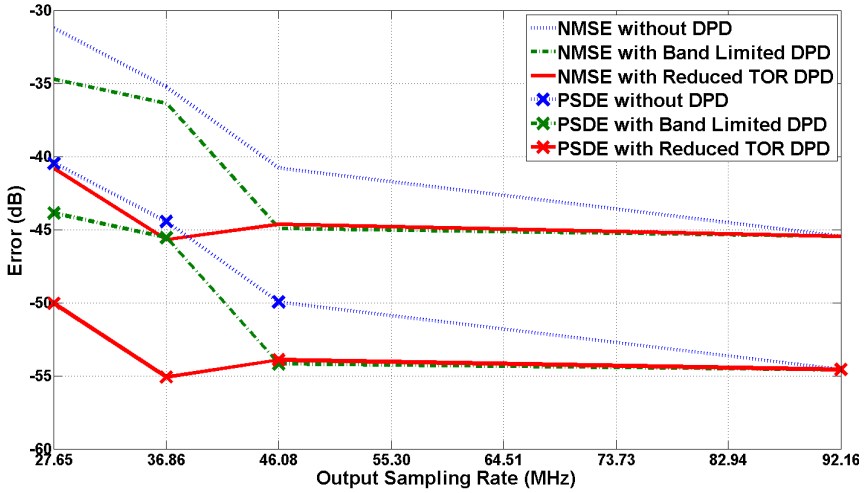

Figure 4.21: NMSE and PSDE between the modeled and the measured output spectrum

When applying the band-limited Volterra series DPD, better modeling performances were recorded when compared to the case where no DPD is used as illustrated in Figure 4.21. Nevertheless, as the TOR bandwidth is reduced below 46.08 MHz, the NMSE and the PSDE degrades as well. This is explained by the fact that the band-limited Volterra series allows linearization only inside the TOR bandwidth and no linearization is observed outside as explained previously. Therefore, the band-limited Volterra series doesn't satisfy the GST condition. However, when the proposed reduced TOR bandwidth DPD is used, linearization inside and outside the TOR bandwidth is achieved and therefore the condition of the GST are met. Consequently, as illustrated by Figure 4.21, the NMSE and the PSDE are constant versus the TOR bandwidth even for very small output sampling rate.

4.7 Conclusion

This chapter demonstrates that GST is not applicable for DPD identification. Alternatively, it is shown that a DPD can be constructed using only a portion of the output spectrum, i.e., when using a reduced TOR bandwidth. This is possible by lowering only the sampling rate of the PA output signal and conserving the full sampling rate of the PA input one. Moreover, it is shown that the reduction extend of the TOR is depending on the DPD structure used. For that, the G-Fucntionals pruning technique introduced in Chapter 3 is deployed. Excellent linearization capabilities are achieved for wideband signals, up to 60 MHz, using a reduced TOR bandwidth. The proposed technique is compared to the band-limited Volterra series. It is shown that using the proposed technique, linearization is possible inside and outside the TOR bandwidth, where only linearization inside the TOR bandwidth is achieved in the band-limited Volterra series case. The proposed DPD linearization capabilities enabled the GST application for PA behavioral modeling. This allows the TOR bandwidth reduction from the usual five times to as low as one time the original signal bandwidth.

Chapter 5

Single-Band Digital Predistortion

This chapter presents a novel strategy for constructing a SISO behavioral model/DPD scheme. The process starts with establishing a block diagram that describes the output signal of single-band PAs. Then, a closed-form formulation is derived. It will be shown that the resulting model is inherently low complexity and takes into account linear and nonlinear memory effects. The Volterra series DDR and the GMP [76] models are used as model references. It will be shown that the proposed model exhibit similar or better performances than the previously mentioned reference models while deploying fewer coefficients.

5.1 Power Amplifier Physical Inspiration

A good understanding of PA behavior is a crucial step in developing a suitable model that can closely mimic or linearize the PA output response. This section describes the PA behavior. Then, the subsequent sections explain the development process of PA/DPD model schemes.

In [4, 25, 85], a simplified field-effect transistor (FET) based PA circuit was proposed, as shown in Figure 5.1. The PA circuit is composed of a voltage input source, $V_s(t)$; an input band-pass filter, representing the linear dynamic input matching network; a nonlinear current source, i_{DS}, representing the nonlinear transistor; an output band-pass filter, representing the linear dynamic output matching network; and the output voltage, $V_O(t)$. Figure 5.1 is used to derive the block diagram of the PA given in Figure 5.2. The PA is represented by two band-pass filters representing the input and output matching networks, a memoryless nonlinear function representing the nonlinear transistor, and a low pass filter representing the bias network that is responsible for reflecting the baseband component at the output of the transistor. The mechanism of the PA block diagram is illustrated with a

two-tone test input signal. First, the signal passes through the linear input band-pass filter without any nonlinear distortion, i.e., no generation of new frequency components. Second, the transistor amplifies the two tones and generates IMDs, harmonics, and baseband components. Finally, the baseband signal is reflected back to the input of the transistor, the harmonics are filtered out, and the amplified two-tone and the IMDs pass through the output band-pass filter, i.e., the output matching network. It is safe to consider the response of the two band-pass filters to be flat and thus contribute little to the combined distortions.

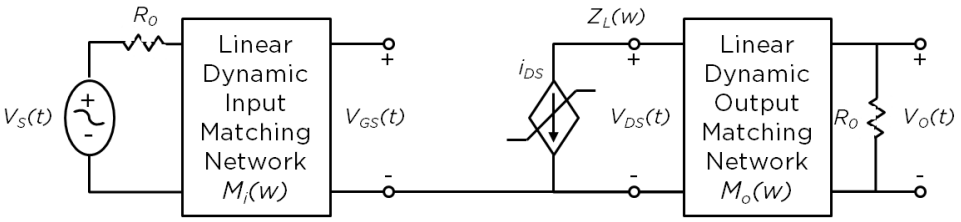

Figure 5.1: Simplified field-effect transistor based power amplifier circuit

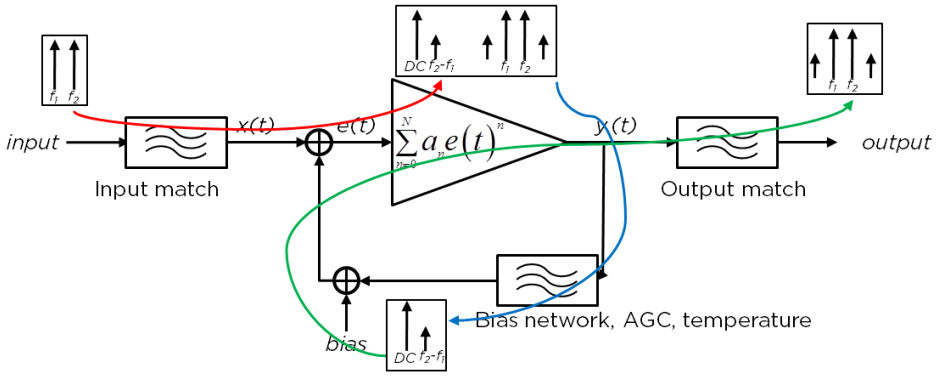

Figure 5.2: Power amplifier block diagram illustration with two-tone test input signal

The resulting block diagram (Figure 5.2) is the root used in the next section to derive a single-band PAs. This block diagram is also used in the next chapter for the multi-band PA case.

5.2 Single-Band Generalized Memory Polynomial

5.2.1 Single-Band Model Derivation

The physically based model presented in the previous section is used to formulate a closed-form expression for the nonlinearity and memory effects present in an RF PA. The adopted

PA block diagram is a simplification of the one given in Figure 5.2, where the input and output matching networks are considered flat. Figure 5.3 shows a simple representation of the behavior of the PA in the envelope domain: the signals $x(n)$ and $y(n)$ are real discrete-time sampled at $T_s = 1/2f_{max}$, where f_{max} is the highest RF frequency of the input and output signals.

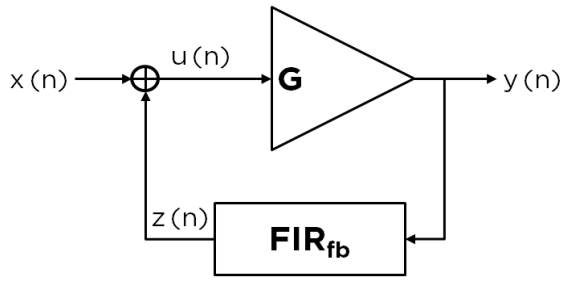

Figure 5.3: Simplified physical model of a power amplifier

The closed-form expression of Figure 5.3 encompasses a static nonlinear function G that accounts for the memoryless nonlinear behavior of the PA, along with a finite impulse response (FIR) filter, FIR_{fb}, which captures the frequency-dependent mechanism behind the memory effects. In Figure 5.3, the discrete real output signal, $y(n)$, of the RF PA circuit is expressed in terms of the discrete real input, $x(n)$, and its past samples as in (5.2), where γ_j is the j^{th} instance of the discrete time domain representation of the feedback FIR filter, FIR_{fb}, and M denotes the memory depth. The feedback FIR, FIR_{fb}, allows the modeling of the biasing and matching networks dispersive behavior at the signals' envelope and harmonic frequencies.

$$y(n) = G\left(x(n) + z(n)\right) \tag{5.1}$$

$$y(n) = G\left(x(n) + \sum_{j=1}^{M} \gamma_j y(n-j)\right) \tag{5.2}$$

The static nonlinear behavior of the transistor is characterized by a memoryless nonlinear function, G, defined in (5.3), where N is the memoryless nonlinear order.

$$G(x) = p_1 x + p_2 x^2 + \cdots + p_N x^N \tag{5.3}$$

Expanding the expression of G in (5.2), the following formulation is obtained:

$$y(n) = p_1\left(x(n) + \sum_{j=1}^{M} \gamma_j y(n-j)\right) + \cdots + p_N\left(x(n) + \sum_{j=1}^{M} \gamma_j y(n-j)\right) \tag{5.4}$$

Equation (5.4) suggests that if memory effects are not taken into account, i.e., $\gamma_j = 0$, the model reduces to the memoryless nonlinear transistor expression, i.e., G. It is important at this stage to examine the summation of reflected signals in (5.2); the terms of which are expressed as: $\gamma_j y(n-j)\ \ j \geq 1$. Expanding each term of that summative expression, using $N = 3$ as an example, yields the following expression in terms of the real input signal, $x(n)$:

$$\begin{aligned}
\gamma_j y(n-j) = {} & p_1\gamma_j x(n-j) + \underline{p_1\gamma_j \sum_{k=1}^{M} \gamma_k y(n-j-k)} \\
& + p_2\gamma_j x^2(n-j) + \underline{p_2\gamma_j \left(\sum_{k=1}^{M} \gamma_k y(n-j-k)\right)^2} \\
& + \underline{2p_2\gamma_j x(n-j) \sum_{k=1}^{M} \gamma_k y(n-j-k)} \\
& + p_3\gamma_j x^3(n-j) + \underline{3p_3\gamma_j x^2(n-j) \sum_{k=1}^{M} \gamma_k y(n-j-k)} \\
& + \underline{3p_3\gamma_j x(n-j) \left(\sum_{k=1}^{M} \gamma_k y(n-j-k)\right)^2} + \underline{p_3\gamma_j \left(\sum_{k=1}^{M} \gamma_k y(n-j-k)\right)^3}
\end{aligned} \tag{5.5}$$

In (5.5), the underlined terms similar to $\gamma_j x^Q(n-j)\left(\sum_{k=1}^{M} \gamma_k y(n-j-k)\right)^K$ where $Q, K \in \{0,1,2\}$ with $Q+K > 1$, can be rewritten as $x^Q(n-j)\left(\sum_{k=1}^{M} \gamma_j^{\frac{1}{K}} \gamma_k y(n-j-k)\right)^K$. These terms represent multiple cycles through the feedback path outlined in Figure 5.3, and become negligible when the total order is higher than one, that is, when past samples of the input are multiplied by a factor of $\gamma_j^a \gamma_k^b$, where $a + b > 1$ [76]. Consequently, such terms can be omitted from the calculations, and the previous expression of $\gamma_j y(n-j)$ substituted for all $j \geq 1$; thus, (5.5) is simplified to:

$$\gamma_j y(n-j) = \gamma_j \left(p_1 x(n-j) + p_2 x^2(n-j) + p_3 x^3(n-j)\right) \tag{5.6}$$

$$\gamma_j y(n-j) = \gamma_j G(x(n-j)) \ \ \forall j \geq 1 \tag{5.7}$$

This simplification is explained by the fact that a realistic PA can be modeled using a finite memory feed forward model. doing so requires the recursive model to have a finite output memory span (fading memory system [97]). Therefore, if we assume that the overall memory of the PA model is equal to M, where M is set by definition to be the length of

the feedback FIR according to (5.2), the components at the output of the nonlinear block can only enter the filter once, since the second and following entries would be out of the output's memory span. Thus, the delayed output signal includes only delayed input terms, as shown in (5.7), (Equation (6) in [122]). It is worth mentioning that (5.7) does not mean that the system (memoryless block + feedback) is a memoryless function of the input signal. Conversely, (5.7) is only true for the delayed output signal and not for the current sample of the output signal; i.e., (5.7) is only for $j \geq 1$.

The substitution of (5.7) into (5.4) leads to the following:

$$\begin{aligned} y(n) &= G\left(x(n) + \sum_{j=1}^{M} \gamma_j y(n-j)\right) \\ &= G\left(x(n) + \sum_{j=1}^{M} \gamma_j G\left(x(n-j)\right)\right) \\ &= p_1 x(n) + p_2 x^2(n) + \cdots + p_N x^N(n) \\ &+ \left(p_1 + 2p_2 x(n) + \cdots + N p_N x(n)^{N-1}\right) \sum_{j=1}^{M} \gamma_j G(x(n-j)) \\ &+ \textit{terms including higher order of } \gamma_j \ \left(i.e.; \gamma_j^a; a > 1\right) \end{aligned} \tag{5.8}$$

Based on (5.3), γ_j represents the memory effects contribution of signal $y(n-j)$ added to the current input $x(n)$ toward the output signal $y(n)$. Therefore, to ensure a stable system, the absolute value of γ_j can be considered to be less than one. Consequently, in (5.8) higher powers of γ_j are considered negligible and can be omitted; that is, $|\gamma_j| < 1 \Rightarrow |\gamma_j|^a \ll 1$. Using the last approximation, (5.8) leads to the expression given in (5.9), where $a_{i,j,k} = (i+1)p_{i+1}\gamma_j p_k$.

$$y(n) = G\left(x(n)\right) + \sum_{i=0}^{N-1} \sum_{j=1}^{M} \sum_{k=1}^{N} a_{i,j,k} x^i(n) x^k(n-j) \tag{5.9}$$

In DPD technique, the envelope signal is deployed to construct the predistorted signal in baseband. Therefore, (5.9) should be transformed to its baseband equivalent. In order to accomplish this, the complex envelope signal is denoted hereafter by $\tilde{x}(n)$ such that $2x(n) = \tilde{x}(n)e^{j2\pi f_0 n} + \tilde{x}^*(n)e^{j2\pi f_0 n}$, where f_0 is the carrier frequency of the PA. Usually, only components around the carrier frequency are significant to the PA output formulation [76]. Therefore, all of the terms that lead to spectral components far away from the carrier frequency are discarded. Writing the expression (5.9), using the baseband input and output signals, leads to:

$$
\begin{aligned}
\tilde{y}(n) = & \sum_{\substack{i=1 \\ i\ odd}}^{N} p_i \tilde{x}^{\frac{i+1}{2}} \tilde{x}^{*\frac{i-1}{2}} \\
& + \sum_{\substack{i=1 \\ i\ odd}}^{N-2} \sum_{m=1}^{M} \sum_{\substack{k=2 \\ k\ even}}^{N-1} \alpha_{i,m,k} \tilde{x}^{\frac{i+1}{2}}(n) \tilde{x}^{*\frac{i-1}{2}}(n) \tilde{x}^{\frac{k}{2}}(n-m) \tilde{x}^{*\frac{k}{2}}(n-m) \\
& + \sum_{\substack{i=0 \\ i\ even}}^{N-1} \sum_{m=1}^{M} \sum_{\substack{k=1 \\ k\ odd}}^{N} \beta_{i,m,k} \tilde{x}^{\frac{i}{2}}(n) \tilde{x}^{*\frac{i}{2}}(n) \tilde{x}^{\frac{k+1}{2}}(n-m) \tilde{x}^{*\frac{k-1}{2}}(n-m)
\end{aligned} \tag{5.10}
$$

Equation (5.10) presents a pruned Volterra series formulation denoted as single-band generalized memory polynomial (1D-GMP), where with a high degree of nonlinearity and a large value for memory depth, the number of kernels is much lower than that of the full Volterra series formulation. In addition, the 1D-GMP offers a comparable or lower number of coefficients than does the Volterra series DDR model with $r = 2$, where r is the order of dynamics, as shown in Table 5.1.

This formulation allows the usage of the Volterra series' proven capabilities to model highly nonlinear PAs. When using large values for N and M, no exponential growth occur in the number of kernels, thanks to the physical inspiration of the 1D-GMP. In fact, this model leads to the elimination of unnecessary kernels from the full Volterra series without the need to restrict the nonlinearity order or the memory depth.

Equation (5.10) reveals an order of nonlinearity equal to $2N - 1$ in the closed-form expression, where N is the memoryless nonlinearity order of the transistor (Figure 5.3). In fact, as shown with the above expressions, when dealing with memory effects, the nonlinearity order is higher than the memoryless nonlinearity and equal to $P = 2N - 1$, where N is the actual memoryless nonlinear order of the system being modeled.

The 1D-GMP model exhibits similarities to the GMP model [76], where the cross terms combining the delayed and current input signals are shared. In this 1D-GMP model, the sources of its formulation are given and are based on the physical behavior of the PA. This process leads to a systematic technique to identify, separately, both the nonlinearity order and memory depth. A comparison of the proposed approach to both the Volterra series DDR and the GMP is included to demonstrate that the 1D-GMP affords similar or better results with less complexity, i.e., fewer coefficients.

Table 5.1: Number of kernels in different Volterra series schemes

Volterra series scheme	**Nonlinearity order**	**Memory depth**	**Number of coefficients**
Full	5	4	605
DDR r=1	5	4	23
DDR r=2	5	4	85
GMP	$K_a = [0:4]$ $L_a = [0:3]$ $K_b = [1:3]$ $L_b = [0:3]$ $M_b = [1:3]$		61
1D-GMP	5	4	57
Full	5	5	1308
DDR r=1	5	5	28
DDR r=2	5	5	123
GMP	$K_a = [0:4]$ $L_a = [0:3]$ $K_b = [1:4]$ $L_b = [0:3]$ $M_b = [1:3]$		73
1D-GMP	5	5	70
Full	7	3	944
DDR r=1	7	3	25
DDR r=2	7	3	82
GMP	$K_a = [0:5]$ $L_a = [0:4]$ $K_b = [1:4]$ $L_b = [0:3]$ $M_b = [1:3]$		83
1D-GMP	7	3	80
Full	7	4	3055
DDR r=1	7	4	32
DDR r=2	7	4	130
GMP	$K_a = [0:5]$ $L_a = [0:4]$ $K_b = [1:5]$ $L_b = [0:3]$ $M_b = [1:4]$		115
1D-GMP	7	4	103

5.2.2 Experimental Validation Conditions

This subsection describes the test conditions, circuit, and signal tests. Two technologies of PAs exist nowadays, namely, GaN and LDMOS and two different architectures, namely single-ended or multiple transistors configuration, e.g., Doherty and envelop tracking PA. The experimental results in the following subsections were conducted on four PAs as it follows, to span a wide range of the PA selection available in the market.

- 45 W GaN single-ended PA [105]
- 90 W GaN Doherty PA [106]
- 250 W LDMOS single-ended PA from Freescale Semiconductor
- 250 W LDMOS Doherty PA from Freescale Semiconductor.

Both 20 MHz 1001 WCDMA and 20 MHz LTE test signals, with PAPR of 7.12 dB and 7.74 dB, respectively, were synthesized in Agilent's advanced design system (ADS) and used as test signals. The four PAs were driven into the nonlinear region (output peak power $\approx$ P1dB + 1.3 dB) to ensure the presence of highly dynamic nonlinearity requiring correction, where P1dB is the 1 dB gain compression.

The performance of the PA behavioral modeling and DPD were evaluated based on different figures of merit. For the forward model, the NMSE was used to evaluate the accuracy of the PA behavioral model by capturing the output of the PA. For the reverse model DPD, the ACPR and the EVM were used as indicators of the linearization capability.

5.2.3 Behavioral Modeling Forward Validation

To assess the accuracy of the 1D-GMP model for predicting the behavior of the PA DUT, several tests were conducted.

Parameters Identification

The 1D-GMP has two structure parameters that must be determined; namely the memoryless nonlinearity degree N and the memory depth M. Therefore, two tests were conducted to determine whether the approached achieves a good approximation for these numbers. First, a memoryless polynomial model, as shown in (5.11), was used to determine the unknown N. Tables 5.2 and 5.3 show the memoryless models NMSE versus the nonlinearity

degree for the four PAs driven with the two previously mentioned signals. From these tables, the value of N is chosen as the smallest value that corresponds with the NMSE plateau.

$$y_{memoryless}(n) = \sum_{i=1}^{N} a_i x(n) \left|x(n)\right|^{i-1} \tag{5.11}$$

Table 5.2: Normalized mean square error versus nonlinearity order for the memoryless model for the GaN power amplifiers

N	**NMSE (dB)**				**NMSE Change (dB)**			
	Single-Ended		**Doherty**		**Single-Ended**		**Doherty**	
	WCDMA	**LTE**	**WCDMA**	**LTE**	**WCDMA**	**LTE**	**WCDMA**	**LTE**
1	-21.46	-20.05	-23.26	-23.80	N.A.	N.A.	N.A.	N.A.
2	-30.00	-24.76	-28.31	-29.97	-8.53	-4.71	-5.05	-6.17
3	-32.15	-25.32	-29.40	-31.32	-2.15	-0.56	-1.09	-1.35
4	-33.10	-25.53	-29.94	-32.52	-0.95	-0.22	-0.54	-1.2
5	-33.27	-25.57	-30.44	-33.15	-0.16	0.03	-0.51	-0.64
6	-33.26	-25.57	-30.44	-33.23	0	0	0	-0.08
7	-33.27	-25.58	-30.46	-33.44	0	0.01	-0.02	-0.21
8	-33.27	-25.58	-30.46	-33.46	0	0	0	-0.02
9	-33.27	-25.58	-30.46	-33.47	0	0	0	0

From Tables 5.2 and 5.3, we chose a value of $N = 5$ for the memoryless nonlinearity, as the relative change between $N = 5$ and $N = 7$ is very small. The total nonlinearity degree for the Volterra series is then equal to $P = 2N - 1 = 9$ when including memory as previously explained.

In the second step, the memory depth of the model (5.10) was determined. The value of M is increased for the model expressed in (5.10); and, the optimal value is the smallest M that corresponds to the plateau in NMSE. Tables 5.4 and 5.5 show a summary of the results of these tests when using the optimal nonlinearity degree, $N = 5$. One can observe that starting at $M = 4$ or $M = 5$, only modest changes occurred in the NMSE; consequently, M was selected as 5 in all cases except for the GaN PA driven with an LTE signal where M was selected to be equal to 4.

Table 5.3: Normalized mean square error versus nonlinearity order for the memoryless model for the LDMOS power amplifiers

N	NMSE (dB)				NMSE Change (dB)			
	Single-Ended		Doherty		Single-Ended		Doherty	
	WCDMA	LTE	WCDMA	LTE	WCDMA	LTE	WCDMA	LTE
1	-24.38	-25.48	-16.90	-20.40	N.A.	N.A.	N.A.	N.A.
2	-34.47	-36.58	-25.74	-27.66	-10.1	-11.2	-8.85	-7.26
3	-35.25	-38.09	-28.09	-29.86	-0.78	-1.41	-2.34	-2.20
4	-35.56	-38.95	-29.22	-30.50	-0.31	-0.86	-1.13	-0.63
5	-35.61	-39.02	-30.27	-30.65	-0.04	-0.16	-0.11	-0.15
6	-35.62	-39.16	-30.33	-30.69	-0.01	-0.05	-0.06	0.04
7	-35.62	-39.22	-30.35	-30.69	0	-0.06	0.02	0
8	-35.62	-39.25	-30.35	-30.69	0	-0.03	0	0
9	-35.62	-39.25	-30.35	-30.69	0	0	0	0

Table 5.4: Normalized mean square error versus memory depth for the 1D-GMP model for the GaN power amplifiers

M	NMSE (dB)				NMSE Change (dB)			
	Single-Ended		Doherty		Single-Ended		Doherty	
	WCDMA N=5	LTE N=5	WCDMA N=5	LTE N=5	WCDMA N=5	LTE N=5	WCDMA N=5	LTE N=5
1	-41.64	-39.23	-35.22	-35.64	N.A.	N.A.	N.A.	N.A.
2	-42.50	-39.23	-37.52	-36.24	-0.87	0	-2.30	-0.6
3	-42.57	-39.27	-38.43	-36.47	-0.07	-0.04	-0.92	-0.24
4	-43.41	-39.40	-38.88	-36.61	-0.84	-0.13	-0.45	-0.13
5	-43.56	-39.47	-39.05	-36.66	-0.16	-0.06	-0.17	-0.06
6	-43.62	-39.53	-39.22	-36.74	-0.05	-0.06	-0.17	-0.07
7	-43.67	-39.57	-39.29	-36.78	-0.05	-0.04	-0.07	-0.06
8	-43.72	-39.62	-39.34	-36.87	-0.05	-0.05	-0.05	-0.07

Table 5.5: Normalized mean square error versus memory depth for the 1D-GMP model for the LDMOS power amplifiers

M	NMSE (dB)				NMSE Change (dB)			
	Single-Ended		Doherty		Single-Ended		Doherty	
	WCDMA N=5	LTE N=5	WCDMA N=7	LTE N=5	WCDMA N=5	LTE N=5	WCDMA N=5	LTE N=5
1	-44.53	-43.42	-41.40	-42.19	N.A.	N.A.	N.A.	N.A.
2	-45.61	-44.90	-40.15	-40.91	-1.08	-1.49	1.25	1.28
3	-45.96	-45.12	-40.10	-41.81	-0.35	-0.22	0.05	-0.90
4	-46.75	-45.20	-41.45	-42.63	-0.79	-0.07	-0.32	-0.82
5	-46.83	-45.33	-41.80	-42.81	-0.08	-0.13	-0.34	-0.18
6	-46.88	-45.43	-41.94	-42.92	-0.05	-0.10	-0.15	-0.11
7	-46.96	-45.54	-42.09	-43.08	-0.08	-0.12	-0.15	-0.16
8	-46.99	-45.60	-42.10	-43.20	-0.04	-0.06	-0.01	-0.12

Forward Validation

To evaluate the performance of the 1D-GMP approach further, its modeling accuracy was compared to the GMP the Volterra series DDR. The GMP parameters were set such that its number of coefficients is about the same than the 1D-GMP, and the Volterra series DDR was used with the first and second-order dynamics and $M = 5$ up to the fifth kernel. Tables 5.6–5.9 show the NMSE between the modeled and measured signals for the different PAs driven. From these tables, one can observe that although the Volterra series DDR (r=2) number of coefficients is a little higher than that of the 1D-GMP, a slightly better NMSE is observed in the case of the proposed model. Moreover, the proposed model shows similar NMSE results to the GMP, yet, the linearization performances differ.

5.2.4 Digital Predistortion Reverse Validation

The 1D-GMP model is used here to linearize the previously mentioned four PAs when driven with 20 MHz 1001 WCDMA and 20 MHz LTE signals. Tables 5.10–5.13 summarize the values obtained for the 1D-GMP parameters, under the different scenarios, and the obtained optimal EVMs. In each case, the obtained 1D-GMP DPD performances were compared with that of the Volterra series DDR r=1 and r=2 models, using the same

Table 5.6: Normalized mean square error of the 1D-GMP versus the Volterra series DDR and GMP for the single-ended GaN power amplifier

Model Name	**NMSE (dB)**		**Number of Coefficients**	
	WCDMA	LTE	WCDMA	LTE
Volterra series DDR r=1	-42.08	-38.95	28	25
Volterra series DDR r=2	-43.50	-39.36	123	87
GMP	-43.51	-39.47	73	61
1D-GMP	-43.56	-39.41	70	57

Table 5.7: Normalized mean square error of the 1D-GMP versus the Volterra series DDR and GMP for the GaN Doherty power amplifier

Model Name	**NMSE (dB)**		**Number of Coefficients**	
	WCDMA	LTE	WCDMA	LTE
Volterra series DDR r=1	-43.47	-38.95	28	25
Volterra series DDR r=2	-43.85	-39.36	123	87
GMP	-44.23	-39.47	73	61
1D-GMP	-44.45	-39.65	70	57

Table 5.8: Normalized mean square error of the 1D-GMP versus the Volterra series DDR and GMP for the single-ended LDMOS power amplifier

Model Name	**NMSE (dB)**		**Number of Coefficients**	
	WCDMA	LTE	WCDMA	LTE
Volterra series DDR r=1	-45.29	-44.53	28	25
Volterra series DDR r=2	-46.47	-44.90	123	87
GMP	-46.53	-44.86	73	73
1D-GMP	-46.83	-45.33	70	57

nonlinearity degree and memory depth order, and the GMP, using about the same number of coefficients.

Figures 5.4 and 5.5 describe the output spectrum of the GaN single-ended PA, driven with 20 MHz 1001 WCDMA and LTE signals without DPD and with the previously men-

Table 5.9: Normalized mean square error of the 1D-GMP versus the Volterra series DDR and GMP for the LDMOS Doherty power amplifier

Model Name	**NMSE (dB)**		**Number of Coefficients**	
	WCDMA	LTE	WCDMA	LTE
Volterra series DDR r=1	-38.69	-37.97	28	25
Volterra series DDR r=2	-40.57	-41.17	123	87
GMP	-41.42	-41.17	73	73
1D-GMP	-41.12	-41.53	70	57

Table 5.10: Error vector magnitude with and without DPDS for the GaN single-ended power amplifier

Model Name	**EVM (dB)**		**Number of Coefficients**	
	WCDMA	LTE	WCDMA	LTE
Without	-23.92	-24.87	N.A.	N.A.
Volterra series DDR r=1	-34.34	-35.20	28	25
Volterra series DDR r=2	-34.66	-35.17	123	87
GMP	-33.73	-34.81	73	61
1D-GMP	-34.58	-34.98	70	57

Table 5.11: Error vector magnitude with and without DPDS for the GaN Doherty power amplifier

Model Name	**EVM (dB)**		**Number of Coefficients**	
	WCDMA	LTE	WCDMA	LTE
Without	-20.39	-19.55	N.A.	N.A.
Volterra series DDR r=1	-37.29	-35.65	28	25
Volterra series DDR r=2	-37.10	-35.74	123	87
GMP	-38.64	-35.36	73	61
1D-GMP	-38.06	-35.24	70	57

tioned DPDs. It is clear from these figures that the four schemes were able to linearize the Doherty PA and achieved a reduction of more than 20 dB in the ACPR. Figure 5.4

Table 5.12: Error vector magnitude with and without DPDS for the LDMOS single-ended power amplifier

Model Name	**EVM (dB)**		**Number of Coefficients**	
	WCDMA	LTE	WCDMA	LTE
Without	-26.16	-26.91	N.A.	N.A.
Volterra series DDR r=1	-38.20	-38.81	28	25
Volterra series DDR r=2	-36.86	-43.79	123	123
GMP	-42.57	-41.86	73	73
1D-GMP	-42.55	-41.07	70	57

Table 5.13: Error vector magnitude with and without DPDS for the LDMOS Doherty power amplifier

Model Name	**EVM (dB)**		**Number of Coefficients**	
	WCDMA	LTE	WCDMA	LTE
Without	-20.39	-19.55	N.A.	N.A.
Volterra series DDR r=1	-37.29	-35.65	50	28
Volterra series DDR r=2	-37.10	-35.74	131	123
GMP	-38.64	-35.36	115	73
1D-GMP	-38.06	-35.24	103	57

shows that the 1D-GMP slightly outperformed the other DPDs models at reducing the out-of-band spectrum, by 3 dB. The 1D-GMP used about the same number of coefficients as the GMP and fewer than the Volterra DDR r=2. Similar performances were found in the case of GaN Doherty PA. Figures 5.6 and 5.7 show the linearization performances of the GaN Doherty PA using the 1D-GMP, the DDR and the GMP DPDs.

Figures 5.8 and 5.9 describe the LDMOS Doherty PA output spectrum, driven with 20 MHz 1001 WCDMA and LTE signals without DPD and with the previously mentioned DPDs. As before, all four schemes were able to linearize the PA and achieved a reduction of more than 20 dB in the ACPR. However, the 1D-GMP linearization capability is now more pronounced when using the 1001 WCDMA signal. Actually, better than 5 dB reduction was achieved (Figure 5.8) as compared to the other DPD models. Similar performances were found in the case of LDMOS single-ended PA as given by Figures 5.10 and 5.11.

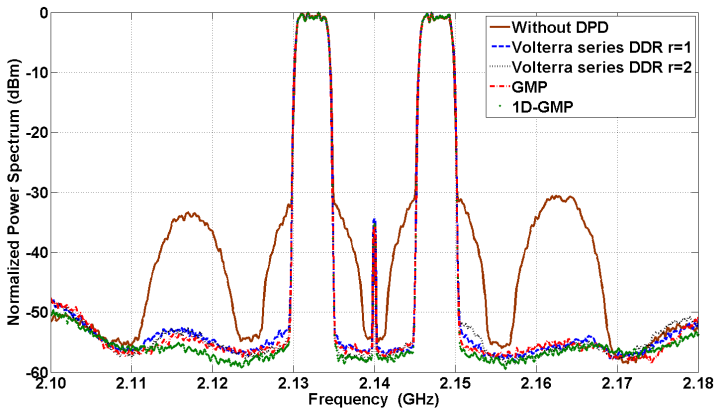

Figure 5.4: PSD of the GaN single-ended PA driven with a 20 MHz 1001 WCDMA signal without DPD and when linearized with the Volterra series DDR r=2, r=1, GMP and 1D-GMP

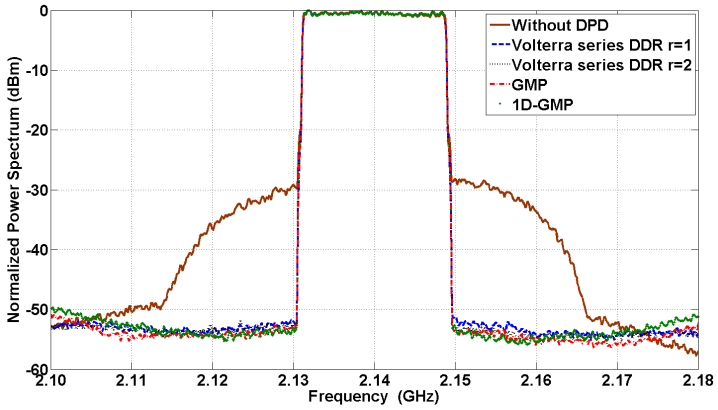

Figure 5.5: PSD of the GaN single-ended PA driven with a 20 MHz LTE signal without DPD and when linearized with the Volterra series DDR r=2, r=1, GMP and 1D-GMP with N=5

The number of coefficients reduction required in the proposed DPD formulation has two advantages. With fewer kernels or coefficients, the DPD identification and implementation is less demanding in terms of both time and memory computation. Therefore, the proposed model can be implemented more easily in a real-time DPD application. In addition, with fewer coefficients, the problem of "over fitting" and the problem conditioning are improved. Thus, the 1D-GMP demonstrated similar or better performances when compared to the proven scheme of the Volterra series DDR r=2, as was reported in the above-mentioned measurements.

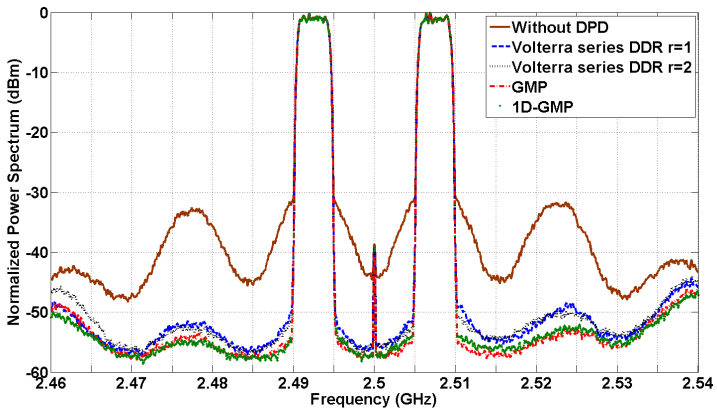

Figure 5.6: PSD of the GaN Doherty PA driven with a 20 MHz 1001 WCDMA signal without DPD and when linearized with the Volterra series DDR r=2, r=1, GMP and 1D-GMP

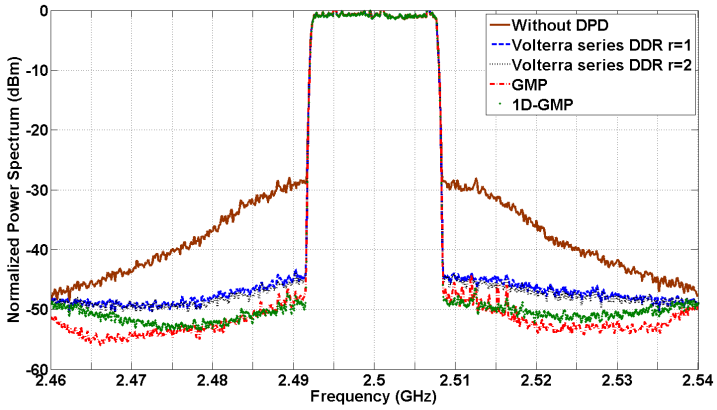

Figure 5.7: PSD of the GaN Doherty PA driven with a 20 MHz LTE signal without DPD and when linearized with the Volterra series DDR r=2, r=1, GMP and 1D-GMP

5.3 Conclusion

This chapter presents a closed-form deterministic approach to identifying a single-band PA/DPD modeling scheme, i.e., 1D-GMP. It starts with proposing PA block diagram that describes the PA behavior. Derivations from the feedback block diagram structure to a feed-forward GMP model is proposed and carefully justified. The resulting 1D-GMP shows good performance in modeling and linearization, when compared to the proven Volterra series DDR and GMP models. In addition, the 1D-GMP has the advantages of both fewer coefficients and a deterministic method to identify its structure parameters. The proposed DPD scheme is used to model and linearize two different PAs with two different 20 MHz WCDMA signals. The linearization performances are discussed in terms of EVM and

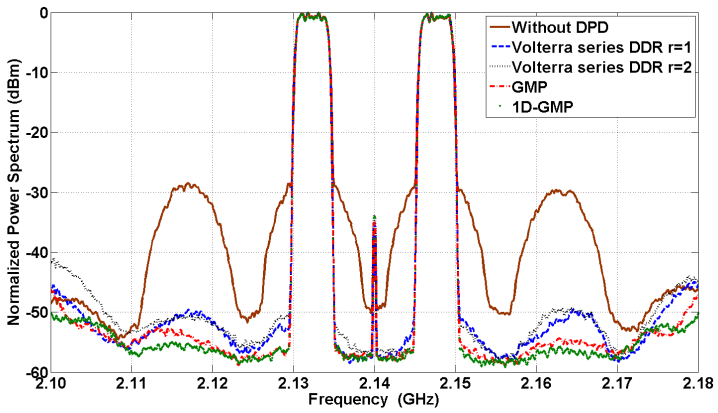

Figure 5.8: PSD of the LDMOS Doherty PA driven with a 20 MHz 1001 WCDMA signal without DPD and when linearized with the Volterra series DDR r=2, r=1, GMP and 1D-GMP

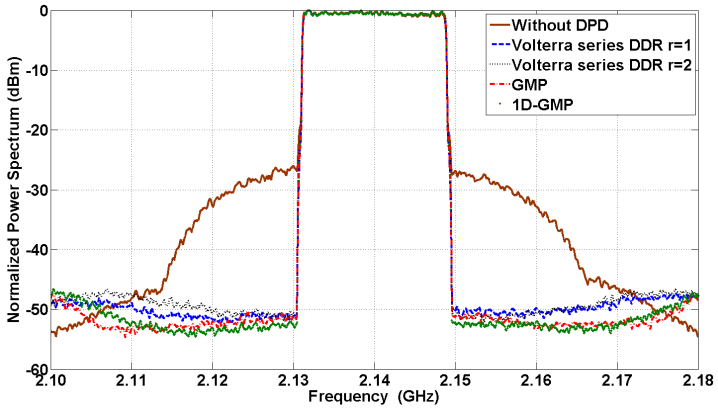

Figure 5.9: PSD of the LDMOS Doherty PA driven with a 20 MHz LTE signal without DPD and when linearized with the Volterra series DDR r=2, r=1, GMP and 1D-GMP

ACPR reduction. It is reported that the 1D-GMP model provides similar or better results than the Volterra series DDR and GMP models while boasting reduced complexity.

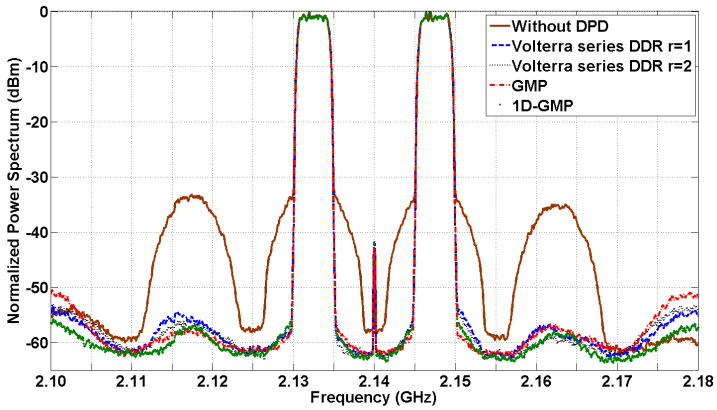

Figure 5.10: PSD of the LDMOS single-ended PA driven with a 20 MHz 1001 WCDMA signal without DPD and when linearized with the Volterra series DDR r=2, r=1, GMP and 1D-GMP

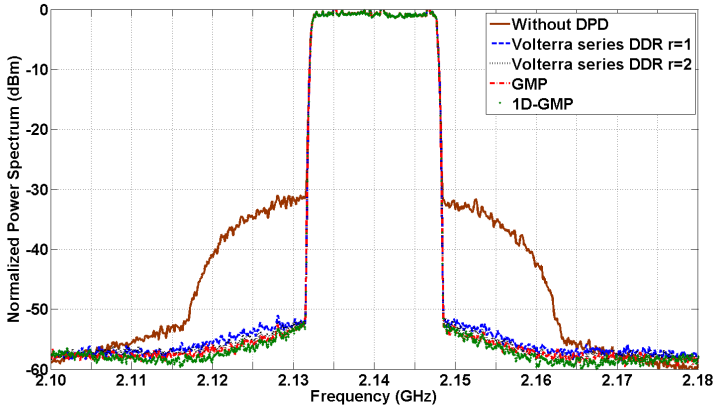

Figure 5.11: PSD of the LDMOS single-ended PA driven with a 20 MHz LTE signal without DPD and when linearized with the Volterra series DDR r=2, r=1, GMP and 1D-GMP with N=5

Chapter 6

Multi-Band Digital Predistortion

In the previous chapter, the 1D-GMP was introduced for a single-band PA. This model enabled a reduced complexity Volterra series like model and offered comparable results to the Volterra series DDR model while using fewer coefficients. The nonlinearity order and memory depth were allowed to reach high values without excessively augmenting the number of coefficients. The inherently low complexity and excellent modeling/linearization capabilities were mainly attributed to the model's construction procedure. In this chapter, the construction of dual-, tri- and multi-band models will be performed using the same procedure.

6.1 Dual-Band Generalized Memory Polynomial

Multi-band PA/DPD models have been proposed to mimic the behavior of a PA under multi-band stimulus. PA distortions in multi-band PAs are more complex than for single-band PAs due to the cross modulation distortion [8]. these cross modulation are a result from the mix between the multiple bands. Thus, the MIMO model, given in Figure 6.1, is adopted for linearizing the multi-band PA output signal, where $x_i(n)$, $u_i(n)$, and $y_i(n)$ are the input, predistorted and output signals of the (DPD+PA) cascade, respectively for each band.

6.1.1 Dual-Band Power Amplifier Block Diagram

Chapter 5 introduced Figure 5.3 as a simplified block diagram of a single-band PA. The same block diagram can be used for the dual-band case. In fact, consider dual-band signals modulated around f_1 and f_2 and a PA with a carrier frequency of $(f_1+f_2)/2$. If $G(.)$

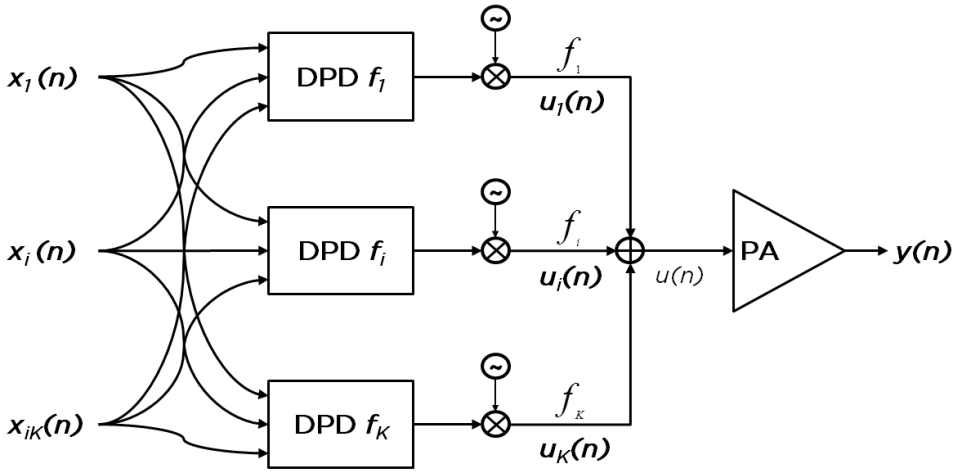

Figure 6.1: Digital predistortion for a multi-band power amplifier

is a gain characteristic of the PA around $(f_1+f_2)/2$, and FIR_{fb} is a filter that illustrates the feedback of low frequencies, Figure 5.3 illustrates the behavior of the dual-band PA. Unfortunately, function $G(.)$ must be very wideband to cover the two band around f_1 and f_2. Alternatively, two different functions, $G_1(.)$ and $G_2(.)$, may be considered to describe the PA gain characteristics around the two carrier frequencies, f_1 and f_2.

Figure 6.2 represents a special case of Figure 5.3 for a typical dual-band PA. The block diagram of Figure 6.2 is composed of two nonlinear memoryless functions and one feedback filter: $x_1(n)$ and $x_2(n)$ are the two input RF signals, and $y_1(n)$ and $y_2(n)$ their respective output RF signals sampled at $T_s = 1/2f_{max}$, where f_{max} is the highest RF frequency component. The FIR_{fb} filter is maintained as it continues to illustrate the feedback of low frequencies. Figure 6.2 is described by (6.1)-(6.3), where γ_j is the j^{th} instance of the feedback FIR discrete time domain representation, FIR_{fb}, and M denotes the memory depth.

$$y(n) = y_1(n) + y_2(n) \tag{6.1}$$

$$y_i(n) \underset{i\in\{1,2\}}{=} G_i\left[u_i(n)\right] = G_i\left[x_1(n) + x_2(n) + z(n)\right] \tag{6.2}$$

$$z(n) = FIR_{fb}\left[y(n)\right] = \sum_{j=1}^{M} \gamma_j y(n-j) \tag{6.3}$$

6.1.2 Dual-Band Power Amplifier Model Derivations

In this subsection, derivations of the dual-band generalized memory polynomial (2D-GMP) model are given using (6.1)-(6.3). Gain characteristics $G_1(.)$ and $G_2(.)$ are represented by

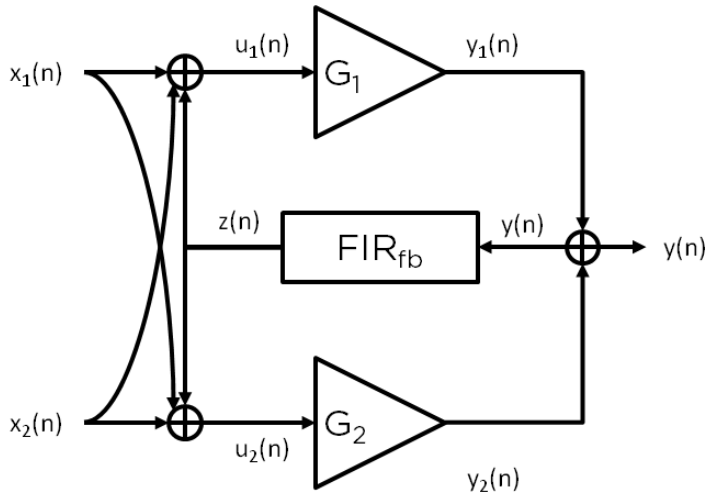

Figure 6.2: Block diagram of a dual-band power amplifier

a memoryless polynomial given in (6.4), where N_1 and N_2 are the memoryless nonlinear orders for each band, respectively.

$$G_i(u) \underset{i\in\{1,2\}}{=} p_{i,1}u + p_{i,2}u^2 + \cdots + p_{i,N_i}u^{N_i} \tag{6.4}$$

Substituting (6.3) and (6.4) into (6.2), the following is obtained, where $i \in \{1, 2\}$:

$$y_i(n) = \sum_{q=1}^{N_i} p_{i,q}\left[x_1(n) + x_2(n) + \sum_{j=1}^{M}\gamma_j y(n-j)\right]^q \tag{6.5}$$

The feedback signals in (6.5), i.e., $\gamma_j y(n-j)$ where $j \geq 1$, can be simplified, as given in (6.6) and illustrated in the previous chapter. In fact, $\gamma_j y(n-j)$ include terms that represent multiple cycles through the feedback path, i.e., $\gamma_j\left[\sum_{k=1}^{M}\gamma_k y(n-j-k)\right]^s$ where $s \geq 1$. Since the multiple cycles become negligible [76], the feedback signal is simplified to (6.6). In fact, a realistic PA is modeled using a finite memory feed-forward model. This requires the recursive model to have a finite output memory span (fading memory system [97]). Therefore, if the overall memory of the PA model is equal to M, the second and following entries of the delayed output signal to the FIR filter would be out of the output's memory span. Consequently, the delayed output signal expression includes only delayed input terms. Similarly to the single-band case, the previous does not mean that the system (memoryless gains + feedback) is a memoryless system. In fact, (6.6) is only true for the delayed output signals, $\gamma_j y(n-j); j \geq 1$, that contribute to the current output expression and not for the current output signal sample, $y(n)$, as shown hereafter.

$$\gamma_j y_q(n-j) = \gamma_j G_q\left[x_1(n-j) + x_2(n-j)\right] \tag{6.6}$$

Substituting (6.6) into (6.5) leads to (6.7).

$$\begin{aligned} y_i(n) = & \sum_{k=1}^{N_i} p_{i,k}\left[x_1(n) + x_2(n)\right]^k \\ & \left(p_{i,1} + \cdots + N_i p_{i,N_i}\left(x_1(n) + x_2(n)\right)^{N_i-1}\right) \sum_{j=1}^{M} \gamma_j \left[\begin{pmatrix} G_1[x_1(n-j) + x_2(n-j)] + \\ G_2[x_1(n-j) + x_2(n-j)] \end{pmatrix}\right] \\ & + terms\ with\ higher\ power\ of\ \gamma_j\ \left(\gamma_j^a; a > 1\right) \end{aligned} \tag{6.7}$$

Based on Figure 6.2, γ_j represents the ME contribution of the signal $y(n-j)$, added to the current input signals, $x_1(n)$ and $x_2(n)$, resulting in the output signal $y_i(n)$. Therefore, to ensure a stable system, the absolute value of γ_j should be considered to be less than one, i.e., $|\gamma_j| < 1$. Thus, higher powers of γ_j are considered to be negligible and may be omitted in (6.7), i.e., $|\gamma_j| < 1 \Rightarrow |\gamma_j|^a \ll 1$ where $a > 1$. Using the last approximation, (6.7) leads to the following expression where repetitive terms are combined:

$$\begin{aligned} y_i(n) = & \sum_{k=1}^{N_i}\sum_{r=0}^{k} \delta x_1^{k-r}(n) x_2^r(n) + \sum_{j=1}^{M}\sum_{k=1}^{N_i}\sum_{q=1}^{max(N_1,N_2)}\sum_{r=0}^{k-1}\sum_{s=0}^{q} \begin{pmatrix} \theta x_1^{k-1-r}(n) x_2^r(n) \\ x_1^{q-s}(n-j) x_2^s(n-j) \end{pmatrix} \\ & where\ \ \delta = \binom{k}{r} p_{i,k}\ \ \theta = \binom{k-1}{r}\binom{q}{s} k \gamma_j p_{i,k} p_{p,q} \end{aligned} \tag{6.8}$$

In the DPD technique, the envelope signal is deployed to construct the predistorted signal in the digital domain. Therefore, (6.8) is transformed here to its baseband equivalent form, where the complex envelope signal is denoted by $\widetilde{x}_i(n)$. This signal is given such that $2x_i(n) = \widetilde{x}_i(n)e^{jw_i n} + \widetilde{x}_i^*(n)e^{-jw_i n}$, where w_i is the carrier angular frequency of the dual-band signal, $i \in \{1, 2\}$. According to [76], only components around the two carrier frequencies are significant to the PA output formulation in general. Therefore, all of the terms that lead to spectral components far from the two carrier frequencies are discarded.

Writing (6.8) using the baseband input and output signals leads to (6.9). This obtained formulation, i.e., (6.9), represents the 2D-GMP model. This model allows cross terms between the two delayed input signals. This feature is crucial to correct for any time delay misalignment between the two signals, as discussed hereafter. The proposed 2D-GMP is inherently reduced complexity when compared to other conventional models, such as the dual-band Volterra series model.

$$
\begin{aligned}
\widetilde{y}_i(n) = & \sum_{\substack{k=2 \\ k\ odd}}^{N_i} \sum_{\substack{r=0 \\ r\ even}}^{k} \varphi \widetilde{x}_i(n) \left|\widetilde{x}_1(n)\right|^{k-r-1} \left|\widetilde{x}_2(n)\right|^{r} \\
& + \sum_{\substack{k=2 \\ k\ even}}^{N_i-1} \sum_{\substack{q=2 \\ q\ even}}^{max(N_1,N_2)-1} \sum_{\substack{r=0 \\ r\ even}}^{k-1} \sum_{\substack{s=0 \\ s\ even}}^{q} \sum_{j=1}^{M} \alpha \widetilde{x}_i(n) \begin{pmatrix} \left|\widetilde{x}_1(n)\right|^{k-r-2} \left|\widetilde{x}_1(n-j)\right|^{q-s} \\ \left|\widetilde{x}_2(n)\right|^{r} \left|\widetilde{x}_2(n-j)\right|^{s} \end{pmatrix} \\
& + \sum_{\substack{k=1 \\ k\ odd}}^{N_i} \sum_{\substack{q=1 \\ q\ odd}}^{max(N_1,N_2)} \sum_{\substack{r=0 \\ r\ even}}^{k-1} \sum_{\substack{s=0 \\ s\ even}}^{q} \sum_{j=1}^{M} \beta \widetilde{x}_i(n-j) \begin{pmatrix} \left|\widetilde{x}_1(n)\right|^{k-r-1} \left|\widetilde{x}_1(n-j)\right|^{q-s-1} \\ \left|\widetilde{x}_2(n)\right|^{r} \left|\widetilde{x}_2(n-j)\right|^{s} \end{pmatrix}
\end{aligned} \tag{6.9}
$$

$$
\begin{aligned}
where \ \varphi &= \delta \binom{k-r}{\frac{k-r-1}{2}} \binom{r}{\frac{r}{2}} \\
\alpha &= \theta \binom{k-r-1}{\frac{k-r-2}{2}} \binom{r}{\frac{r}{2}} \binom{q-s}{\frac{q-s}{2}} \binom{s}{\frac{s}{2}} \\
\beta &= \theta \binom{k-r-1}{\frac{k-r-1}{2}} \binom{r}{\frac{r}{2}} \binom{q-s}{\frac{q-s-1}{2}} \binom{s}{\frac{s}{2}}
\end{aligned}
$$

It is important to note that the nonlinearity order N_i in (6.9), results in an order of nonlinearity equal to $N_i + max(N_1 + N_2) - 1$ in the closed-form expression. When dealing with memory effects, the memory nonlinearity order is higher than the memoryless nonlinearity (denoted hereafter as $P_i = N_i + max(N_1 + N_2) - 1$). In fact, for the second and third summation of (6.9), the nonlinearity order is equal to $k + q - 1$. For $k = N_i$ and $q = max(N_1 + N_2)$, the total nonlinearity is increased to P_i. In this context, the nonlinearity in the 2D-GMP model of (6.9) is higher than in the 2D-DPD model or any polynomial-based model. Consequently, a comparison between polynomial based models will be ineffective as the number of coefficients is high in the 2D-GMP, due to the increase of nonlinearity to P_i.

To alleviate this inequality, the pruning of (6.9) is proposed to unify the nonlinearity order of the memoryless part (first summation of (6.9)) and the memory part (second and third summation of (6.9)) to $P_i = 2N_i - 1$. Therefore, (6.9) is written as (6.10) and referred to hereafter as the modified 2D-GMP model. The resulting number of coefficients is illustrated in Table 6.1. The number of coefficients for the modified 2D-GMP is lower than that of the 2D-DPD, although the proposed model includes cross terms and the 2D-DPD does not. Since the proposed model was derived from a dual-band PA block diagram, only necessary kernels are present. Whereas, the 2D-DPD is a generalization of the MP model with no particular selection of kernels, resulting in a high number of coefficients.

$$
\begin{aligned}
\widetilde{y}_i(n) = & \sum_{\substack{k=2 \\ k\ odd}}^{2N_i-1} \sum_{\substack{r=0 \\ r\ even}}^{k} \varphi \widetilde{x}_i(n) \left|\widetilde{x}_1(n)\right|^{k-r-1} \left|\widetilde{x}_2(n)\right|^{r} \\
& + \sum_{\substack{k=2 \\ k\ even}}^{N_i-1} \sum_{\substack{q=2 \\ q\ even}}^{max(N_1,N_2)-1} \sum_{\substack{r=0 \\ r\ even}}^{k-1} \sum_{\substack{s=0 \\ s\ even}}^{q} \sum_{j=1}^{M} \alpha \widetilde{x}_i(n) \begin{pmatrix} \left|\widetilde{x}_1(n)\right|^{k-r-2} \left|\widetilde{x}_1(n-j)\right|^{q-s} \\ \left|\widetilde{x}_2(n)\right|^{r} \left|\widetilde{x}_2(n-j)\right|^{s} \end{pmatrix} \\
& + \sum_{\substack{k=1 \\ k\ odd}}^{N_i} \sum_{\substack{q=1 \\ q\ odd}}^{max(N_1,N_2)} \sum_{\substack{r=0 \\ r\ even}}^{k-1} \sum_{\substack{s=0 \\ s\ even}}^{q} \sum_{j=1}^{M} \beta \widetilde{x}_i(n-j) \begin{pmatrix} \left|\widetilde{x}_1(n)\right|^{k-r-1} \left|\widetilde{x}_1(n-j)\right|^{q-s-1} \\ \left|\widetilde{x}_2(n)\right|^{r} \left|\widetilde{x}_2(n-j)\right|^{s} \end{pmatrix}
\end{aligned} \tag{6.10}
$$

Table 6.1: Number of coefficients comparison between the modified 2D-GMP, the 2D-CR-GMP and the 2D-DPD models

P_i	**Modified 2D-GMP** (6.10)				**2D-CR-GMP** (6.11)				**2D-DPD Even Odd**			
	Memory Taps				**Memory Taps**				**Memory Taps**			
	0	1	2	3	0	1	2	3	0	1	2	3
1	1	2	3	4	1	2	3	4	1	2	3	4
3	3	6	9	12	3	4	7	10	6	12	18	24
5	6	17	28	39	6	7	14	25	15	30	45	60
7	10	34	58	82	10	11	18	36	28	56	84	112
9	15	66	117	168	15	16	23	47	45	90	135	180

To allow large value of memory taps while keeping a realistic and manageable number of coefficients, a reformulation of (6.10) is proposed in (6.11). The reformulation tolerates a large memory span for low nonlinearity orders and a small memory span for high nonlinearity orders. This is translated in different M values in the second and third summations of (6.10). This leads to (6.11) which is identified as the 2D-complexity-reduced GMP (2D-CR-GMP) model. In this case, M is no longer a single integer number but a vector of integers, i.e., $M = [M_1 \cdots M_{max(N_1,N_2)}]$.

$$
\begin{aligned}
\widetilde{y}_i(n) = & \sum_{\substack{k=2 \\ k\ odd}}^{2N_i-1} \sum_{\substack{r=0 \\ r\ even}}^{k} \varphi \widetilde{x}_i(n) \left|\widetilde{x}_1(n)\right|^{k-r-1} \left|\widetilde{x}_2(n)\right|^{r} \\
& + \sum_{\substack{k=2 \\ k\ even}}^{N_i-1} \sum_{\substack{q=2 \\ q\ even}}^{max(N_1,N_2)-1} \sum_{\substack{r=0 \\ r\ even}}^{k-1} \sum_{\substack{s=0 \\ s\ even}}^{q} \sum_{j=1}^{M(m_1)} \alpha \widetilde{x}_i(n) \begin{pmatrix} \left|\widetilde{x}_1(n)\right|^{k-r-2} \left|\widetilde{x}_1(n-j)\right|^{q-s} \\ \left|\widetilde{x}_2(n)\right|^{r} \left|\widetilde{x}_2(n-j)\right|^{s} \end{pmatrix} \\
& + \sum_{\substack{k=1 \\ k\ odd}}^{N_i} \sum_{\substack{q=1 \\ q\ odd}}^{max(N_1,N_2)} \sum_{\substack{r=0 \\ r\ even}}^{k-1} \sum_{\substack{s=0 \\ s\ even}}^{q} \sum_{j=1}^{M(m_2)} \beta \widetilde{x}_i(n-j) \begin{pmatrix} \left|\widetilde{x}_1(n)\right|^{k-r-1} \left|\widetilde{x}_1(n-j)\right|^{q-s-1} \\ \left|\widetilde{x}_2(n)\right|^{r} \left|\widetilde{x}_2(n-j)\right|^{s} \end{pmatrix}
\end{aligned} \tag{6.11}
$$

$$
\begin{aligned}
where \;\; m_1 &= \frac{(1+(k-r-2)+(q-s)+r+s)+1}{2} \\
m_2 &= \frac{(1+(k-r-1)+(q-s-1)+r+s)+1}{2}
\end{aligned}
$$

One approach to select the values for M is to start with the highest values as M_1 and decrease the values of M_k by a step equal to 1 until 0, when all consequent values are set equal to zero. This is summarized in (6.12). The resulting numbers of coefficients for (6.10) and (6.11) are summarized in Table 6.1. It is clear that the proposed 2D-CR-GMP allows a large value for M, while keeping a manageable number of coefficients. In addition, the 2D-CR-GMP model allows cross terms, resulting in better modeling and linearization capabilities especially when time delay misalignment between the two signals exists.

$$
\begin{cases} M_1 = M \\ M_{k+1} = max(0, M_{k-1} - 1) \end{cases} \quad k \underset{k\ odd}{=} \cdots max(N_1, N_2) \tag{6.12}
$$

6.2 Tri-Band Generalized Memory Polynomial

6.2.1 Tri-Band Power Amplifier Block Diagram

In order to generalize (6.12) to a tri-band PA, the same procedure that yielded the 2D-CR-GMP model is followed. A simplified block diagram of a tri-band PA is proposed, and a feed-forward model is then derived. Figure 6.3 shows a tri-band PA block diagram represented by three nonlinear memoryless functions and a feedback filter, where $x_1(n)$, $x_2(n)$ and $x_3(n)$ are the three input RF signals and $y_1(n)$, $y_2(n)$ and $y_3(n)$ are the respective RF output signals. As with the dual-band case, the three nonlinear functions can be

attributed to the PA exhibiting different gain characteristics at each frequency range. The filter simulates the feedback of low frequencies that produce memory effects and thus this feedback filter is kept the same as in Figure 5.3. Figure 6.3 is described by (6.13)-(6.16), where γ_j is the j^{th} instance of feedback filter discrete time domain representation FIR_{fb}, and M denotes the memory depth.

$$y(n) = y_1(n) + y_2(n) + y_3(n) \tag{6.13}$$

$$y_i(n) = G_i\left[u_i(n)\right] \quad i \in \{1, 2, 3\} \tag{6.14}$$

$$u_i(n) = x_1(n) + x_2(n) + x_3(n) + z(n) \quad i \in \{1, 2, 3\} \tag{6.15}$$

$$z(n) = FIR_{fb}[y(n)] = \sum_{j=1}^{M} \gamma_j y(n-j) \tag{6.16}$$

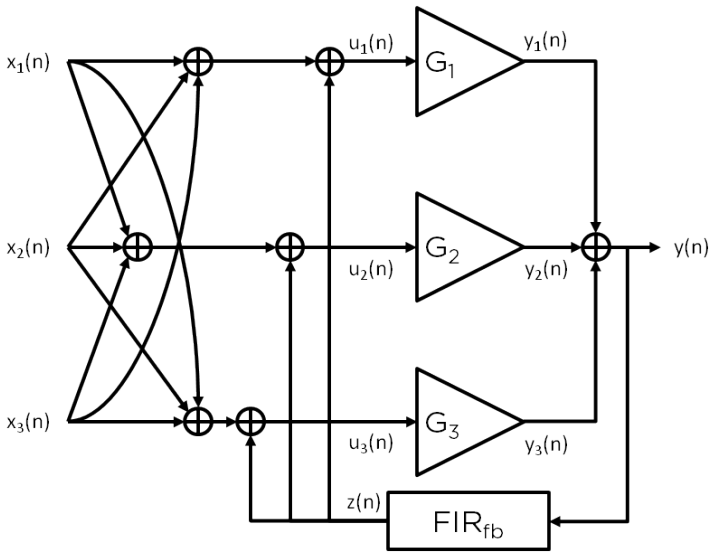

Figure 6.3: Block diagram of a tri-band power amplifier

6.2.2 Tri-Band Power Amplifier Model Derivations

The static nonlinearities $G_1(.)$, $G_2(.)$, and $G_3(.)$ are characterized by memoryless polynomials given in (6.17), where N_1, N_2 and N_3 are the memoryless nonlinear orders.

$$G_i(u) = p_{i,1}u + p_{i,2}u^2 + \cdots + p_{i,N_i}u^{N_i} \quad i \in \{1, 2, 3\} \tag{6.17}$$

Following the derivations that led to the 2D-GMP model, it is easy to show that the tri-band delayed output signals are expressed as in (6.18). Since the PA has fading memory,

the delayed output signals are allowed to enter the FIR feedback filter only once. Thus, their expressions include only the input signals.

$$\gamma_j y_1(n-j) \underset{\forall j \geq 1}{=} \gamma_j G_q \left[x_1(n-j) + x_2(n-j) + x_3(n-j) \right] \tag{6.18}$$

Substituting (6.17) and (6.18) into (6.14) leads to (6.19), where $i \in \{1, 2, 3\}$:

$$y_i(n) = \begin{bmatrix} x_1(n) + x_2(n) + x_3(n) + \\ \sum_{j=1}^{M} \sum_{q=1}^{3} \gamma_j G_q \left[x_1(n-j) + x_2(n-j) + x_3(n-j) \right] \end{bmatrix} \tag{6.19}$$

To ensure that the system of Figure 6.3 is stable, the absolute value of γ should be less than 1, i.e., $|\gamma_j| < 1$. Using (6.19) and recognizing the fact that any powered terms of the feedback FIR are neglected, $|\gamma_j| < 1 \Rightarrow |\gamma_j|^a \ll 1$, with $a > 1$, the following is obtained.

$$\begin{aligned} y_i(n) = & \sum_{k=1}^{N_i} \sum_{r=0}^{k} \sum_{t=0}^{r} \sigma x_1^{k-r}(n) x_3^{r-t}(n) x_3^{t}(n) + \\ & \sum_{k=1}^{N_i} \sum_{j=1}^{M} \sum_{p=1}^{3} \sum_{q=1}^{N_p} \sum_{r=0}^{q} \sum_{s=0}^{k-1} \sum_{t=0}^{r} \sum_{u=0}^{s} \vartheta \begin{pmatrix} x_1^{k-s-1}(n) x_2^{s-u}(n) x_3^{u}(n) \\ x_1^{q-r}(n-j) x_2^{r-t}(n-j) x_3^{t}(n-j) \end{pmatrix} \\ & \delta = \binom{k}{r} p_{i,k} \quad \sigma = \binom{r}{t} \delta \quad \vartheta = \theta \binom{r}{t} \binom{s}{u} \quad \theta = \binom{q}{r} \binom{k-1}{s} k \gamma_j p_{i,k} p_{p,q} \end{aligned} \tag{6.20}$$

Writing (6.20) as the baseband equivalent model, (6.21) is obtained and denoted as triband generalized memory polynomial (3D-GMP). Since the total nonlinearity of (6.21) is $P_i = N_i + max(N_1 + N_2 + N_3) - 1$, in order to compare this model with the polynomial based model, the memoryless nonlinearity (the first summation (6.21)) is set to $P_i = 2N_i - 1$. This results in the modified 3D-GMP model given in (6.22).

$$\begin{aligned} \widetilde{y}_i(n) = & \sum_{\substack{k=1 \\ k\ odd}}^{N_i} \sum_{\substack{r=0 \\ r\ even}}^{k-1} \sum_{\substack{t=0 \\ t\ even}}^{r} \varphi \widetilde{x}_i(n) |\widetilde{x}_1(n)|^{k-r-1} |\widetilde{x}_2(n)|^{r-t} |\widetilde{x}_3(n)|^{t} + \\ & \sum_{\substack{k=2 \\ k\ even}}^{N_i} \sum_{\substack{q=2 \\ q\ even}}^{max(N_1,N_2,N_3)} \sum_{\substack{r=0 \\ r\ even}}^{q} \sum_{\substack{s=0 \\ s\ even}}^{k-1} \sum_{\substack{t=0 \\ t\ even}}^{r} \sum_{\substack{u=0 \\ u\ even}}^{s} \sum_{j=1}^{M} \alpha \frac{\widetilde{x}_i(n) |\widetilde{x}_1(n)|^{k-s-2} |\widetilde{x}_2(n)|^{s-u} |\widetilde{x}_3(n)|^{u}}{|\widetilde{x}_1(n-j)|^{q-r} |\widetilde{x}_2(n-j)|^{r-t} |\widetilde{x}_3(n-j)|^{t}} + \\ & \sum_{\substack{k=1 \\ k\ odd}}^{N_i} \sum_{\substack{q=1 \\ q\ odd}}^{max(N_1,N_2,N_3)} \sum_{\substack{r=0 \\ r\ even}}^{q} \sum_{\substack{s=0 \\ s\ even}}^{k-1} \sum_{\substack{t=0 \\ t\ even}}^{r} \sum_{\substack{u=0 \\ u\ even}}^{s} \sum_{j=1}^{M} \beta \frac{\widetilde{x}_i(n-j) |\widetilde{x}_1(n)|^{k-s-1} |\widetilde{x}_2(n)|^{s-u} |\widetilde{x}_3(n)|^{u}}{|\widetilde{x}_1(n-j)|^{q-r-1} |\widetilde{x}_2(n-j)|^{r-t} |\widetilde{x}_3(n-j)|^{t}} \end{aligned} \tag{6.21}$$

$$\widetilde{y}_i(n) = \sum_{\substack{k=1 \\ k\ odd}}^{2N_i-1} \sum_{\substack{r=0 \\ r\ even}}^{k-1} \sum_{\substack{t=0 \\ t\ even}}^{r} \varphi\widetilde{x}_i(n)|\widetilde{x}_1(n)|^{k-r-1}|\widetilde{x}_2(n)|^{r-t}|\widetilde{x}_3(n)|^t + \sum_{\substack{k=2 \\ k\ even}}^{N_i} \sum_{\substack{q=2 \\ q\ even}}^{max(N_1,N_2,N_3)} \sum_{\substack{r=0 \\ r\ even}}^{q} \sum_{\substack{s=0 \\ s\ even}}^{k-1} \sum_{\substack{t=0 \\ t\ even}}^{r} \sum_{\substack{u=0 \\ u\ even}}^{s} \sum_{j=1}^{M} \alpha \frac{\widetilde{x}_i(n)|\widetilde{x}_1(n)|^{k-s-2}|\widetilde{x}_2(n)|^{s-u}|\widetilde{x}_3(n)|^u}{|\widetilde{x}_1(n-j)|^{q-r}|\widetilde{x}_2(n-j)|^{r-t}|\widetilde{x}_3(n-j)|^t} + \sum_{\substack{k=1 \\ k\ odd}}^{N_i} \sum_{\substack{q=1 \\ q\ odd}}^{max(N_1,N_2,N_3)} \sum_{\substack{r=0 \\ r\ even}}^{q} \sum_{\substack{s=0 \\ s\ even}}^{k-1} \sum_{\substack{t=0 \\ t\ even}}^{r} \sum_{\substack{u=0 \\ u\ even}}^{s} \sum_{j=1}^{M} \beta \frac{\widetilde{x}_i(n-j)|\widetilde{x}_1(n)|^{k-s-1}|\widetilde{x}_2(n)|^{s-u}|\widetilde{x}_3(n)|^u}{|\widetilde{x}_1(n-j)|^{q-r-1}|\widetilde{x}_2(n-j)|^{r-t}|\widetilde{x}_3(n-j)|^t} \tag{6.22}$$

From Table 6.2, it can be seen that the modified 3D-GMP requires fewer coefficients than the 3D-DPD, although cross terms are included in the proposed model. To allow a large value in the memory span, while keeping a manageable number of coefficients, the same pruning applied to obtain the 2D-CR-GMP is used here. In other words, different M values in the second and third summations of (6.22) are used. This leads to (6.23), called hereafter the tri-band complexity reduced generalized memory polynomial (3D-CR-GMP) model. M_k are selected as in (6.12). The resulting number of coefficients is summarized in Table 6.2.

$$\widetilde{y}_i(n) = \sum_{\substack{k=1 \\ k\ odd}}^{2N_i-1} \sum_{\substack{r=0 \\ r\ even}}^{k-1} \sum_{\substack{t=0 \\ t\ even}}^{r} \varphi\widetilde{x}_i(n)|\widetilde{x}_1(n)|^{k-r-1}|\widetilde{x}_2(n)|^{r-t}|\widetilde{x}_3(n)|^t + \sum_{\substack{k=2 \\ k\ even}}^{N_i} \sum_{\substack{q=2 \\ q\ even}}^{max(N_1,N_2,N_3)} \sum_{\substack{r=0 \\ r\ even}}^{q} \sum_{\substack{s=0 \\ s\ even}}^{k-1} \sum_{\substack{t=0 \\ t\ even}}^{r} \sum_{\substack{u=0 \\ u\ even}}^{s} \sum_{j=1}^{M(m_1)} \alpha \frac{\widetilde{x}_i(n)|\widetilde{x}_1(n)|^{k-s-2}|\widetilde{x}_2(n)|^{s-u}|\widetilde{x}_3(n)|^u}{|\widetilde{x}_1(n-j)|^{q-r}|\widetilde{x}_2(n-j)|^{r-t}|\widetilde{x}_3(n-j)|^t} + \sum_{\substack{k=1 \\ k\ odd}}^{N_i} \sum_{\substack{q=1 \\ q\ odd}}^{max(N_1,N_2,N_3)} \sum_{\substack{r=0 \\ r\ even}}^{q} \sum_{\substack{s=0 \\ s\ even}}^{k-1} \sum_{\substack{t=0 \\ t\ even}}^{r} \sum_{\substack{u=0 \\ u\ even}}^{s} \sum_{j=1}^{M(m_2)} \beta \frac{\widetilde{x}_i(n-j)|\widetilde{x}_1(n)|^{k-s-1}|\widetilde{x}_2(n)|^{s-u}|\widetilde{x}_3(n)|^u}{|\widetilde{x}_1(n-j)|^{q-r-1}|\widetilde{x}_2(n-j)|^{r-t}|\widetilde{x}_3(n-j)|^t} \tag{6.23}$$

$$m_1 = \frac{(1+(k-s-2)+(s-u)+u+(q-r)+(r-t)+t)+1}{2}$$
$$m_1 = \frac{(1+(k-s-1)+(s-u)+u+(q-r-1)+(r-t)+t)+1}{2}$$

Table 6.2: Comparison of number of coefficients required by modified 3D-GMP, 3D-CR-GMP and 3D-DPD Models

P_i	**Modified 3D-GMP** (6.22)				**3D-CR-GMP** (6.23)				**2D-DPD Even Odd**			
	Memory Taps				**Memory Taps**				**Memory Taps**			
	0	1	2	3	0	1	2	3	0	1	2	3
1	1	2	3	4	1	2	3	4	1	2	3	4
3	4	8	12	16	4	5	9	13	10	20	30	40
5	10	29	48	67	10	11	21	40	35	70	105	140
7	20	72	124	176	20	21	31	65	84	168	252	336
9	35	171	307	443	35	36	46	92	165	330	495	660

6.3 Multi-Band Generalized Memory Polynomial

In Chapter 5 and the previous sections, single-, dual- and tri-band models are derived from simplified PA block diagrams. The formulations given in (5.10), (6.11) and (6.23) show similarities in their derivations, which leads to the proposal of a general multi-band PA model given in (6.24). This model can be reduced to single-, dual-, and tri-band and any number of bands, depending on the number of inputs.

$$
\begin{aligned}
\widetilde{y}_i(n) = & \sum_{\substack{k_1=1 \\ k_1\ odd}}^{2N_i-1} \sum_{\substack{k_2=0 \\ k_2\ even}}^{k_1} \cdots \sum_{\substack{k_K=0 \\ k_K\ even}}^{h_K-1} \varphi\widetilde{x}_i(n)|\widetilde{x}_1(n)|^{k_1-k_2-1} \left(\prod_{l=2}^{K-1} |\widetilde{x}_l(n)|^{k_l-k_{l+1}} \right) |\widetilde{x}_K(n)|^{k_K} + \\
& \sum_{\substack{k_1=2 \\ k_1\ even}}^{N_i} \cdots \sum_{\substack{k_K=0 \\ k_K\ even}}^{h_{K-1}} \sum_{\substack{q_1=2 \\ q_1\ even}}^{\substack{max(N_i) \\ i\in\{1,\cdots,K\}}} \cdots \sum_{\substack{q_K=0 \\ q_K\ even}}^{q_{K-1}} \sum_{j=1}^{M(m_1)} \begin{pmatrix} \alpha\widetilde{x}_i(n) \\ |\widetilde{x}_1(n)|^{k_1-k_2-2}|\widetilde{x}_K(n)|^{k_K} \\ \prod_{l=2}^{K-1} |\widetilde{x}_l(n)|^{k_l-k_{l+1}} \\ |\widetilde{x}_1(n-j)|^{q_1-q_2}|\widetilde{x}_K(n-j)|^{q_K} \\ \prod_{l=2}^{K-1} |\widetilde{x}_l(n-j)|^{q_l-q_{l+1}} \end{pmatrix} + \\
& \sum_{\substack{k_1=1 \\ k_1\ odd}}^{N_i} \cdots \sum_{\substack{k_K=0 \\ k_K\ even}}^{h_{K-1}} \sum_{\substack{q_1=1 \\ q_1\ odd}}^{\substack{max(N_i) \\ i\in\{1,\cdots,K\}}} \cdots \sum_{\substack{q_K=0 \\ q_K\ even}}^{q_{K-1}} \sum_{j=1}^{M(m_2)} \begin{pmatrix} \beta\widetilde{x}_i(n-j) \\ |\widetilde{x}_1(n)|^{k_1-k_2-1}|\widetilde{x}_K(n)|^{k_K} \\ \prod_{l=2}^{K-1} |\widetilde{x}_l(n)|^{k_l-k_{l+1}} \\ |\widetilde{x}_1(n-j)|^{q_1-q_2-1}|\widetilde{x}_K(n-j)|^{q_K} \\ \prod_{l=2}^{K-1} |\widetilde{x}_l(n-j)|^{q_l-q_{l+1}} \end{pmatrix}
\end{aligned} \tag{6.24}
$$

$$m_1 = \frac{(1 + (k_1 - k_2 - 2) + k_K + \sum_{l=2}^{K-1}(k_l - k_{l+1}) + (q_1 - q_2) + q_K + \sum_{l=2}^{K-1}(1_l - q_{l+1})}{2}$$

$$m_1 = \frac{(1 + (k_1 - k_2 - 1) + k_K + \sum_{l=2}^{K-1}(k_l - k_{l+1}) + (q_1 - q_2 - 1) + q_K + \sum_{l=2}^{K-1}(1_l - q_{l+1})}{2}$$

6.4 Test Setup and Experimental Validation Conditions

Two wideband PAs were used as the DUTs for validation of the 2D-CR-GMP and 3D-CR-GMP models:

- 45 W single-ended GaN PA (1.9-2.9 GHz)
- 20 W Doherty GaN PA (700-950 MHz)

The two PAs were driven into their nonlinear regions to ensure the presence of highly dynamic nonlinearities requiring correction. Different signal standards were deployed:

- 20 MHz 4 C WCDMA signals with a PAPR equal to 7.14 dB
- 20 MHz 1001 WCDMA signals with a PAPR equal to 7.11 dB
- 20 MHz LTE signals with a PAPR equal to 7.73 dB

The test bed used to assess the proposed models' linearization performance is given in Figure 6.4. An 8 giga sample per second arbitrary waveform generator (AWG M9180A) allows the generation of signals from DC up to 4 GHz. This covers a wide range of cellular frequencies including the two DUTs' operating frequencies. Modulations and up-conversions of the multi-band signals were digitally performed using MATLAB. The resulting real signal is uploaded to the AWG. A spectrum analyzer (PXA N9030A) was used to capture the different output signals with an analysis bandwidth equal to 100 MHz.

A multi-band PA has different group delays for the different carrier frequencies. These different group delays result in time delay misalignment between the different signals, i.e., different times needed by the input signals to transform to the output signals. To correct for this time delay misalignment, some input signals should be intentionally delayed before modulation to equal the total delay between the different input and output signals. Unfortunately, these group delays vary based on many variables, e.g., PA output power, carrier frequency, temperature. Thus, there is no guarantee that the correction applied is valid over time.

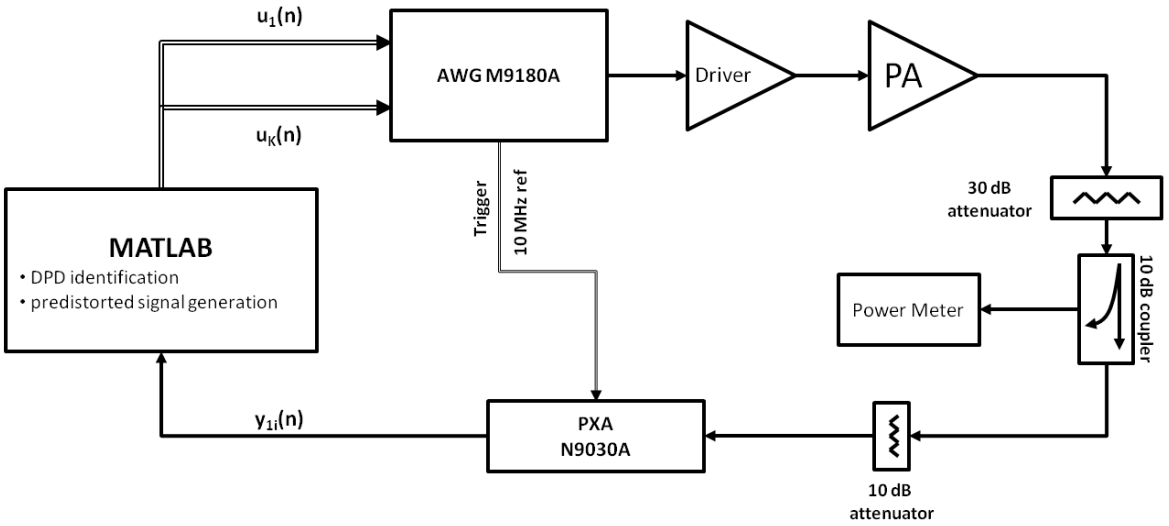

Figure 6.4: Test bed for the multi-band digital predistortion

Typically, the DPD should handle this time delay misalignment. It is shown in last section of this chapter that when cross terms exists in the DPD structure, as in the case of the 2D-CR-GMP model or 3D-CR-GMP, the linearization is not sensitive to this issue. However, in the absence of these cross terms, as in the case of 2D-DPD, the performance of the DPD degrades even for a small time delay misalignment.

6.5 Dual-Band and Tri-Band Digital Predistortion Validation

In the previous sections, behavioral modeling schemes of dual-, tri- and multi-band signals has been discussed. Derivations from PA block diagrams have demonstrated a CR-GMP models that allow cross terms, while keeping a manageable number of coefficients. The objective of this section is the validation of these models as DPD linearizers.

To begin, the DPD parameters; nonlinearity order P_i and memory depth M, were selected. The NMSE was then used to evaluate the forward DPD modeling capabilities. Finally, the linearization performance of the DPD was evaluated by observing the EVM and the output spectrum reduction through the ACPR. The first subsection examines the dual-band case, the second subsection, the tri-band case.

6.5.1 Dual-Band Digital Predistortion

The two DUTs were driven with dual input WCDMA signals. The single-ended GaN PA was driven with the 4C 20 MHz WCDMA signal around 2.14 GHz and the 20 MHz 1001 WCDMA signal around 2.8 GHz (combined signal PAPR was 9.28 dB). The Doherty PA was driven with the 20 MHz 1001 WCDMA signal around 750 MHz and the 4C 20 MHz WCDMA around 900 MHz (combined signal PAPR was 9.24 dB). The DUTs were driven to compression to ensure high nonlinearity.

To identify the 2D-CR-GMP nonlinearity and memory depth, several tests were conducted. A memoryless 2D model given in (6.25) was identified to determine the best nonlinearity degree, P_i. The nonlinearity order was chosen as the smallest value that corresponded to the NMSE plateau: P_i was selected as 7 for both DUTs. The memory depth of model (6.11) was determined with the optimal selected $P_i = 7$: M was selected to be 3 for both DUTs.

$$\widetilde{u}_i(n) = \sum_{\substack{k=1 \\ k\ odd}}^{2N_i-1} \sum_{\substack{r=0 \\ r\ even}}^{k} \delta\widetilde{y}_i(n)|\widetilde{y}_1(n)|^{k-r-1}|\widetilde{y}_2(n)|^r \tag{6.25}$$

To evaluate the performance of the 2D-CR-GMP approach with the obtained parameters, its modeling accuracy was compared to that of the 2D-DPD. For this section, the time delay misalignment was equal to zero, while in next section, the time delay misalignment is non-zero.

Table 6.3 shows a comparison between the proposed 2D-CR-GMP and the 2D-DPD modeling with odd only and even odd order capabilities. In Table 6.3, LB refers to the lower band and HB refers to the higher band. The 2D-DPD nonlinearity and memory depth were chosen to be the same as the 2D-CR-GMP. It is clear from Table 6.3 that the 2D-CR-GMP provided comparable NMSEs for both DUTs when compared to the 2D-DPD model, although the latter model requires more coefficients.

Tables 6.4 and 6.5 summarize the EVM values obtained under the different scenarios. Measured performances were compared against the 2D-DPD where odd only and even odd orders were selected. Both DPD schemes succeeded in linearizing the DUT, reducing the EVM by more than 10 dB in all cases. This is clear from Figures 6.5–6.8 where the linearized output spectrums achieved -50 dBc in all cases. Therefore, the 2D-CR-GMP showed comparable EVM and ACPR reductions and forward DPD performance as the 2D-DPD model.

Table 6.3: Normalized mean square error of the 2D-CR-GMP model versus the 2D-DPD model without time delay misalignment

	Model Name		NMSE (dB)		# of Coeffs	
			LB	HB	LB	HB
GaN PA	2D-DPD Odd Only	$(P_i = 7, M = 3)$	-36.47	-38.12	40	40
	2D-DPD Even Odd	$(P_i = 7, M = 3)$	-36.80	-38.52	112	112
	2D-DPD Even Odd	$(P_i = 5, M = 3)$	-36.80	-38.49	60	60
	2D-CR-GMP	$(P_i = 7, M = 3)$	-36.76	-38.43	36	36
Doherty PA	2D-DPD Odd Only	$(P_i = 7, M = 3)$	-37.78	-38.76	40	40
	2D-DPD Even Odd	$(P_i = 7, M = 3)$	-38.34	-39.28	112	112
	2D-DPD Even Odd	$(P_i = 5, M = 3)$	-38.16	-39.17	60	60
	2D-CR-GMP	$(P_i = 7, M = 3)$	-38.59	-39.11	36	36

Table 6.4: Error vector magnitude with and without DPD for the GaN single-ended power amplifier

Model Name		EVM (dB)		# of Coeffs	
		2.14 GHz Band	2.80 GHz Band	2.14 GHz Band	2.80 GHz Band
Without DPD		-26.31	-24.94		
2D-DPD Odd Only	$(P_i = 7, M = 3)$	-36.39	-38.64	40	40
2D-DPD Even Odd	$(P_i = 7, M = 3)$	-36.74	-39.38	112	112
2D-DPD Even Odd	$(P_i = 5, M = 3)$	-36.83	-34.87	60	60
2D-CR-GMP	$(P_i = 7, M = 3)$	-36.77	-39.12	36	36

6.5.2 Tri-Band Digital Predistortion

To examine the 3D-CR-GMP model capabilities, the single-ended GaN PA was driven with a tri-band signal composed of a 20 MHz LTE signal around 2.14 GHz, a 4C WCDMA signal around 2.6 GHz and a 1001 WCDMA signal around 2.7 GHz (combined PAPR was 10.43 dB). First, the memoryless 3D model given in (6.26) was identified to determine the nonlinearity degree P_i. This nonlinearity order is chosen to be the smallest value that corresponded to the NMSE plateau: P_i was selected to be 7. Second, the memory depth

Table 6.5: Error vector magnitude with and without DPD for the Doherty GaN power amplifier

Model Name		EVM (dB)		# of Coeffs	
		750 MHz Band	900 MHz Band	750 MHz Band	900 MHz Band
Without DPD		-25.52	-28.28		
2D-DPD Odd Only	$(P_i = 7, M = 3)$	-37.43	-38.32	40	40
2D-DPD Even Odd	$(P_i = 7, M = 3)$	-38.09	-39.06	112	112
2D-DPD Even Odd	$(P_i = 5, M = 3)$	-37.96	-38.66	60	60
2D-CR-GMP	$(P_i = 7, M = 3)$	-38.41	-38.67	36	36

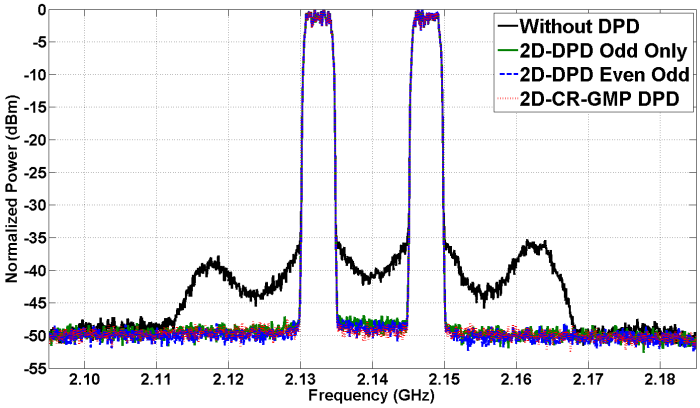

Figure 6.5: 2.14 GHz band power spectrum density of the GaN single-ended PA driven with a dual-band 20 MHz WCDMA signal, without DPD and when linearized with the 2D-CR-GMP and 2D-DPD linearizers

of the 3D-CR-GMP model, i.e., (6.23), was determined with the optimal selected $P_i = 7$: M was selected to be 2.

$$\widetilde{u}_i(n) = \sum_{\substack{k=1 \\ k\ odd}}^{2N_i-1} \sum_{\substack{r=0 \\ r\ even}}^{k} \sum_{\substack{t=0 \\ t\ even}}^{r} \varphi \widetilde{y}_i(n) |\widetilde{y}_1(n)|^{k-r-1} |\widetilde{y}_2(n)|^{r-t} |\widetilde{y}_3(n)|^t \tag{6.26}$$

To evaluate the performance of the 3D-CR-GMP approach, its modeling accuracy was compared to 3D-DPD with odd only and even odd order capabilities. The nonlinearity and memory depth of the 3D-DPD and 3D-CR-GMP were set at 7 and 2, respectively. It

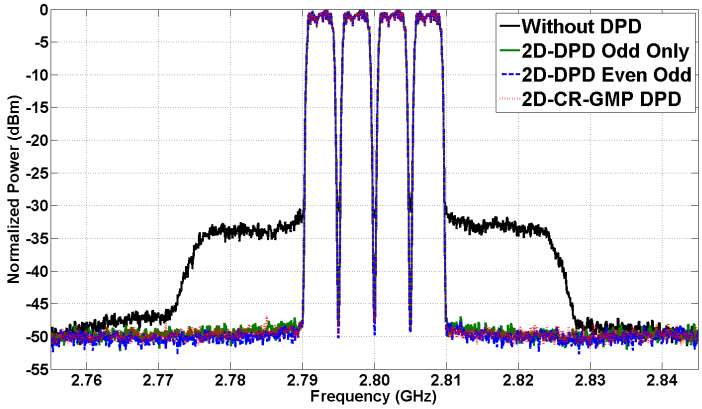

Figure 6.6: 2.80 GHz band power spectrum density of the GaN single-ended PA driven with a dual-band 20 MHz WCDMA signal, without DPD and when linearized with the 2D-CR-GMP and 2D-DPD linearizers

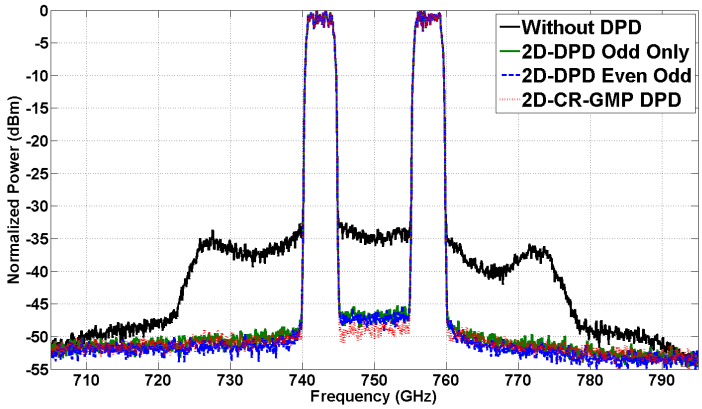

Figure 6.7: 750 MHz band power spectrum density of the GaN Doherty PA driven with a dual-band 20 MHz WCDMA signal, without DPD and when linearized with the 2D-CR-GMP and 2D-DPD linearizers

is clear from Table 6.6 that the 3D-CR-GMP provides comparable NMSE when compared to the 3D-DPD model, although the latter model used much more number of coefficients.

Table 6.7 summarizes the EVM values obtained for the two models. Both DPD schemes linearized the DUT by reducing the EVM more than 10 dB in the three bands. It can be seen in Figures 6.9–6.11 that the linearized output spectrums pass the spectrum mask [1] for all three bands. It is worth mentioning that the 3D-CR-GMP used only 31 coefficients in each band to linearize the PA output whereas the 3D-DPD required a much larger number of coefficients.

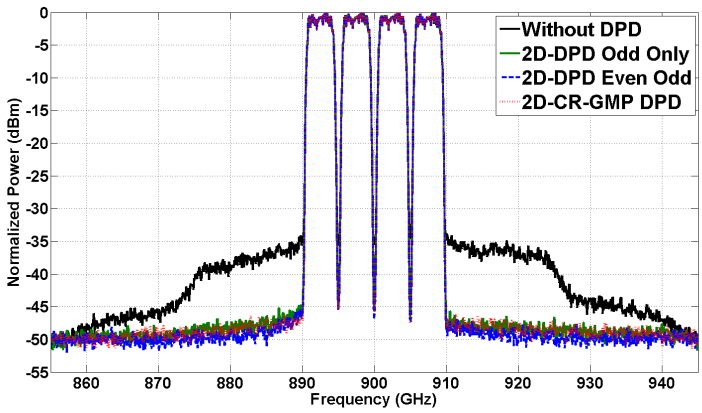

Figure 6.8: 900 MHz band power spectrum density of the GaN Doherty PA driven with a dual-band 20 MHz WCDMA signal, without DPD and when linearized with the 2D-CR-GMP and 2D-DPD linearizers

Table 6.6: Normalized mean square error of the 3D-CR-GMP model versus the 3D-DPD model

Model Name		**NMSE (dB)**			**# of**
		2.14 GHz	**2.6 GHz**	**2.7 GHz**	**Coeffs**
3D-DPD Odd Only	$(P_i = 7, M = 2)$	-36.60	-38.56	-38.38	60
3D-DPD Even Odd	$(P_i = 7, M = 2)$	-36.78	-38.77	-38.48	252
3D-DPD Even Odd	$(P_i = 5, M = 2)$	-36.80	-38.80	-38.50	105
3D-CR-GMP	$(P_i = 7, M = 2)$	-36.67	-38.44	-38.49	31

Table 6.7: Error vector magnitude of the 3D-CR-GMP model versus the 3D-DPD model

Model Name		**NMSE (dB)**			**# of**
		2.14 GHz	**2.6 GHz**	**2.7 GHz**	**Coeffs**
Without DPD		-27.53	-25.43	-24.34	
3D-DPD Odd Only	$(P_i = 7, M = 2)$	-36.64	-38.24	-38.31	60
3D-DPD Even Odd	$(P_i = 7, M = 2)$	-36.73	-38.58	-38.34	252
3D-DPD Even Odd	$(P_i = 5, M = 2)$	-33.99	-38.45	-38.39	105
3D-CR-GMP	$(P_i = 7, M = 2)$	-36.80	-38.17	-38.31	31

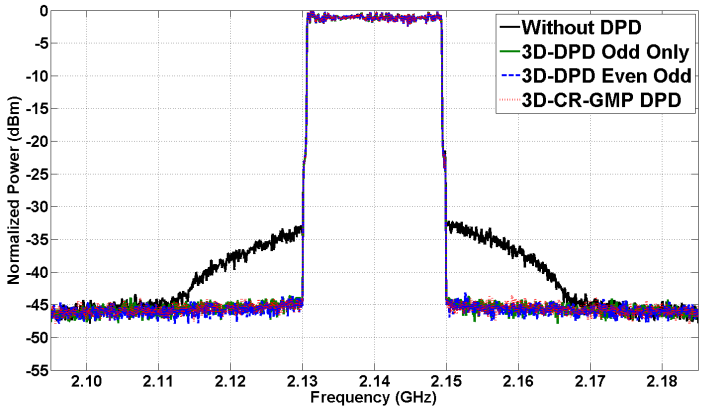

Figure 6.9: 2.14 GHz band power spectrum density of the GaN single-ended PA driven with tri-band signals, without DPD and when linearized with the 3D-CR-GMP and 3D-DPD linearizers

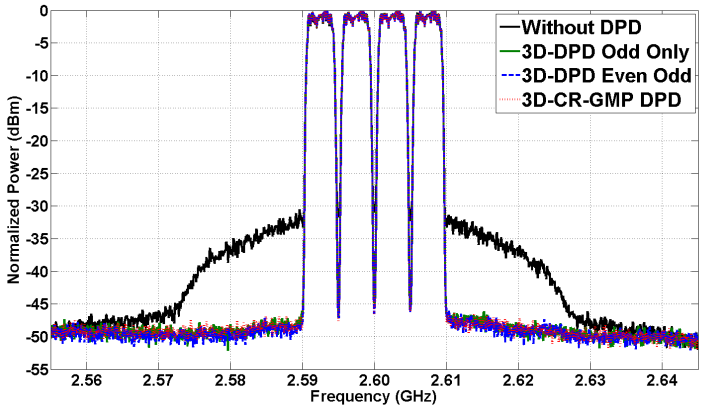

Figure 6.10: 2.6 GHz band power spectrum density of the GaN single-ended PA driven with tri-band signals, without DPD and when linearized with the 3D-CR-GMP and 3D-DPD linearizers

6.6 Multi-Band Digital Predistortion Sensitivity to Time Delay Misalignment

In this section, the sensitivity of the multi-band DPD to time delay misalignment is examined, and measurement validation for dual-band PA is proposed. As mentioned in the previous section, a multi-band PA has different group delays for the different carrier frequencies. These different group delays result in time delay misalignment between the different signals, i.e., different times needed by the input signals to transform to the output signals. Hereafter, this time is noted τ_i, where i is an index indicating one band.

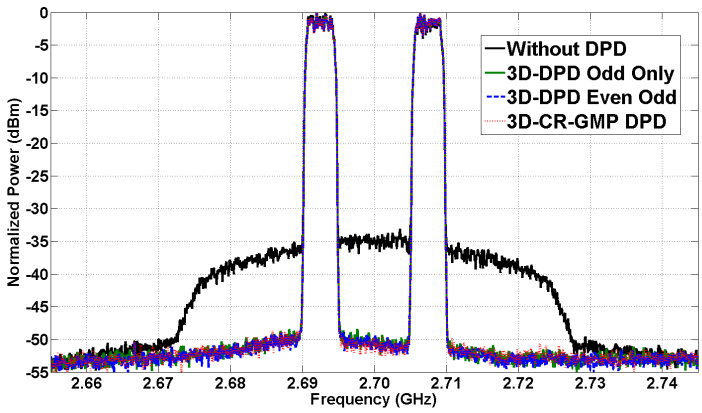

Figure 6.11: 2.7 GHz band power spectrum density of the GaN single-ended PA driven with tri-band signals, without DPD and when linearized with the 3D-CR-GMP and 3D-DPD linearizers

Consider a baseband signal $\{\widetilde{u}_1(n)\cdots\widetilde{u}_K(n)\}$. Prior to driving the PA, correction for the different group delays is applied such that τ_i is the same for all bands. This is done through delaying the baseband signal as follows: $\{\widetilde{u}_1(n+\delta_1)\cdots\widetilde{u}_K(n+\delta_K)\}$, where δ_i corrects for the different τ_i. Note that for the DPD identification, the PA input signal $\{\widetilde{u}_1(n)\cdots\widetilde{u}_K(n)\}$ is used, while the delayed version is used to drive the PA. If the different multi-band PA group delays are constant, but not necessarily equal, the resulting output signal is $\{\widetilde{y}_1(n+\delta_1+\tau_1)\cdots\widetilde{y}_K(n+\delta_K+\tau_K)\}$ such that $\delta_i+\tau_i=\tau\ \forall i\in[1,K]$. The sampled signals at the output of the PA are in the form of $\{\widetilde{y}_1(n+\tau)\cdots\widetilde{y}_K(n+\tau)\}$. The time alignment adjustment results in the following signals: $\{\widetilde{u}_1(n)\cdots\widetilde{u}_K(n)\}$ and $\{\widetilde{y}_1(n)\cdots\widetilde{y}_K(n)\}$.

DPD coefficients for a particular band, i, are derived using (6.27), where the time aligned output signals are used to map the input signal $\widetilde{u}_i(n)$. Thus, all the output signals should be time aligned with $\widetilde{u}_i(n)$. Since the corrected group delays in this case are constant, $\delta_i+\tau_i=\tau\ \forall i\in[1,K]$, the identification of the DPD coefficients is accurate and good linearization results, as shown in the previous section.

$$\widetilde{u}_i(n)=f_{DPD}\left(\widetilde{y}_1(n)\cdots\widetilde{y}_K(n)\right) \tag{6.27}$$

When the group delays change, due to changes in the temperature, input power, bandwidth, etc., the PA output signals are in the form $\{\,\widetilde{y}_1(n+\delta_1+\tau_1+\tau_1')\cdots\widetilde{y}_K(n+\delta_K+\tau_K+\tau_K')\,\}$. In this case, $\delta_i+\tau_i+\tau_i'\neq\tau'$ for some $i\in[1,K]$, where τ' is a constant. Since the delay adjustment is performed between the couple $\{\widetilde{u}_i(n),\widetilde{y}_i(n+\delta_i+\tau_i+\tau_i')\}$,

the resulting output signals are still $\{\widetilde{y}_1(n)\cdots\widetilde{y}_K(n)\}$. For specific band i using (6.27) to identify the DPD coefficients, the total delay between the output signals and the input signal $\widetilde{u}_i(n)$ should be $\delta_i + \tau_i + \tau_i'$. Thus, the output signals that must be used for (6.27) and aligned with $\widetilde{u}_i(n)$ are $\{\, \widetilde{y}_1\left(n + (\delta_1 + \tau_1 + \tau_1') - (\delta_i + \tau_i + \tau_i')\right) \cdots \widetilde{y}_K\left(n + (\delta_K + \tau_K + \tau_K') - (\delta_i + \tau_i + \tau_i')\right) \,\}$. Using $\delta_i + \tau_i = \tau \;\forall i \in [1, K]$, the output signals for (6.27) are $\{\widetilde{y}_1\left(n + \tau_1' - \tau_i'\right)\cdots\widetilde{y}_i(n)\cdots\widetilde{y}_K\left(n + \tau_K' - \tau_i'\right)\}$. The DPD formulation should be able to take into account this time delay misalignment and use signals $\{\widetilde{y}_1(n)\cdots\widetilde{y}_K(n)\}$ to generate $\{\widetilde{y}_1\left(n + \tau_1' - \tau_i'\right)\cdots\widetilde{y}_i(n)\cdots\widetilde{y}_K\left(n + \tau_K' - \tau_i'\right)\}$.

In the following, the dual-band case is studied to illustrate how the proposed 2D-CR-GMP is able to adjust for time delay misalignment while 2D-DPD cannot. The second predistorted signal is studied here, i.e., $\widetilde{u}_2(n)$, and by analogy the same is applied to the first predistorted signal, $\widetilde{u}_1(n)$. To identify the DPD coefficients of the second band, the output signals used should be $\{\widetilde{y}_1(n + \delta_1' - \delta_2'), \widetilde{y}_2(n)\}$. Thus, kernels that involve the power of $\widetilde{y}_1(n + \Delta\delta - j)$ and $\widetilde{y}_2(n - j)$ should be used where $\Delta\delta = \delta_1' - \delta_2'$.

Recalling that the adjusted delayed signals used in the DPD coefficients are $\{\, \widetilde{u}_1(n), \widetilde{u}_2(n) \,\}$ and $\{\widetilde{y}_1(n), \widetilde{y}_2(n)\}$ for the 2D-DPD formulation, all the kernels are in the form of $\widetilde{y}_2(n-j)|\widetilde{y}_1(n-j)|^a|\widetilde{y}_2(n-j)|^b$. Therefore, in the 2D-DPD formulation, it is impossible to generate or estimate kernels with different delays, $\widetilde{y}_1(n + \Delta\delta - j)$ and $\widetilde{y}_2(n - j)$. It is anticipated that the linearization performance of the 2D-DPD degrades in the case of time delay misalignment. However, in the case of the proposed 2D-CR-GMP, cross terms, such as $\widetilde{y}_2(n-j)|\widetilde{y}_1(n)|^a|\widetilde{y}_1(n-j)|^b|\widetilde{y}_2(n)|^c|\widetilde{y}_2(n-j)|^d$, are allowed. Thus, estimation of kernels with different delays for the two output signals is possible; and, it is anticipated that the 2D-CR-GMP linearization performance is not sensitive to time delay misalignment.

Measurement validation of the sensitivity to time delay misalignment of the 2D-DPD and the 2D-CR-GMP models is conducted hereafter. The single-ended and Doherty PAs were driven with dual-band signals with 4.6 ns and 10 ns time delay misalignments, respectively. These time delay misalignments are equivalent to half a sample for the single-ended PA and one sample for the Doherty PA. The 2D-CR-GMP and 2D-DPD were used to linearize the two PAs; and, the performances are summarized in Tables 6.8 and 6.9 and Figures 6.12–6.15.

Tables 6.8 and 6.9 show the EVM results without DPD and when using the 2D-CR-GMP DPD and 2D-DPD with odd only and even odd orders. When comparing these tables with those presented in previous section, it is clear that the 2D-CR-GMP maintained EVM results in the presence of time delay misalignment; whereas the 2D-DPD EVMs degraded by more than 2 dB for the higher bands of both DUTs. The 2D-CR-GMP achieved the same out-of-band reduction. This is also evident from the ACPR reduction shown in

Figures 6.12–6.15, where the 2D-DPD sometimes violates the -45 dBc mask. It is worth mentioning here that the time delay misalignment was selected to be very small; however, there is no guarantee that the misalignment is small in real applications. Thus, more severe degradation can be observed in the case of the 2D-DPD.

Table 6.8: Error vector magnitude with and without DPD for the GaN single-ended power amplifier, with 4.6 ns time delay misalignment

Model Name	**EVM (dB)**		**# of Coeffs**	
	2.14 GHz Band	2.80 GHz Band	2.14 GHz Band	2.80 GHz Band
Without DPD	-26.37	-24.96		
2D-DPD Odd Only $(P_i = 7, M = 3)$	-36.54	-36.04	40	40
2D-DPD Even Odd $(P_i = 7, M = 3)$	-36.72	-36.34	112	112
2D-DPD Even Odd $(P_i = 5, M = 3)$	-36.80	-36.32	60	60
2D-CR-GMP $(P_i = 7, M = 3)$	-36.77	-38.79	36	36

Table 6.9: Error vector magnitude with and without DPD for the Doherty GaN power amplifier, with 10 ns time delay misalignment

Model Name	**EVM (dB)**		**# of Coeffs**	
	750 MHz Band	900 MHz Band	750 MHz Band	900 MHz Band
Without DPD	-25.62	-28.47		
2D-DPD Odd Only $(P_i = 7, M = 3)$	-35.67	-34.52	40	40
2D-DPD Even Odd $(P_i = 7, M = 3)$	-35.89	-34.64	112	112
2D-DPD Even Odd $(P_i = 5, M = 3)$	-35.85	-34.63	60	60
2D-CR-GMP $(P_i = 7, M = 3)$	-38.18	-36.82	36	36

The sensitivity of the DPD to time misalignment can also be generalized for higher bands. In other words, if the multi-band DPD has proper cross terms in its formulation, the linearization is robust against time delay misalignment. This is the case of the proposed 2D-CR-GMP and 3D-CR-GMP models and the proposed multi-band DPD of (6.24). However, due to the construction of the 2D-DPD and 3D-DPD, no cross terms were al-

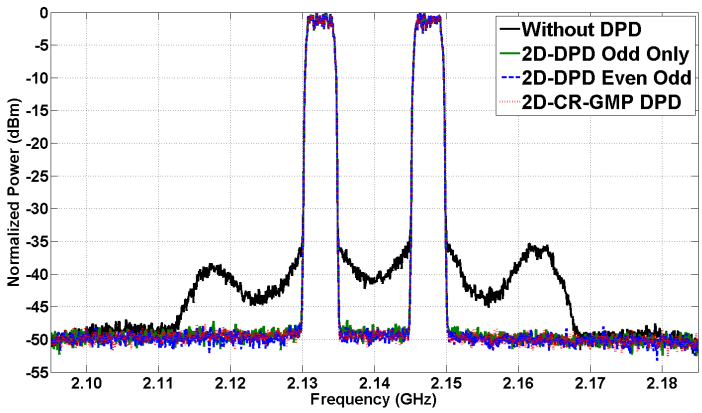

Figure 6.12: 2.14 GHz band power spectrum density of the GaN single-ended PA driven with a dual-band 20 MHz WCDMA signal, without DPD and when linearized with the 2D-CR-GMP and 2D-DPD linearizers

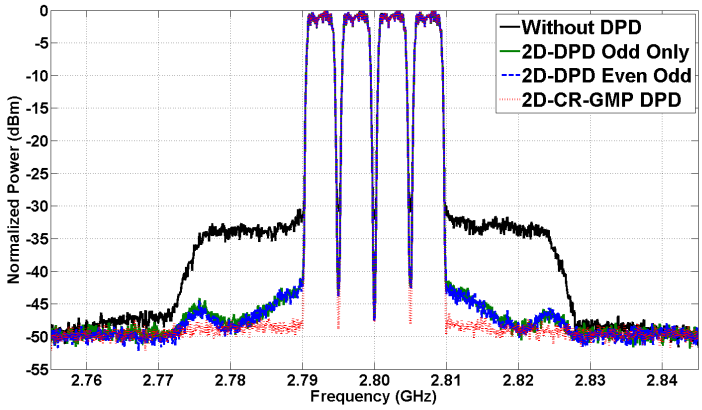

Figure 6.13: 2.80 GHz band power spectrum density of the GaN single-ended PA driven with a dual-band 20 MHz WCDMA signal, without DPD and when linearized with the 2D-CR-GMP and 2D-DPD linearizers

lowed; therefore, the performances of these models degrade when time delay misalignment occurs.

It is important to note that the 2D-CR-GMP and the other proposed multi-band models are robust against the time delay misalignment without the need of adding new coefficients and thereby increasing the complexity. In fact, since the proposed models were designed from PA block diagrams, these models are inherently robust against this issue.

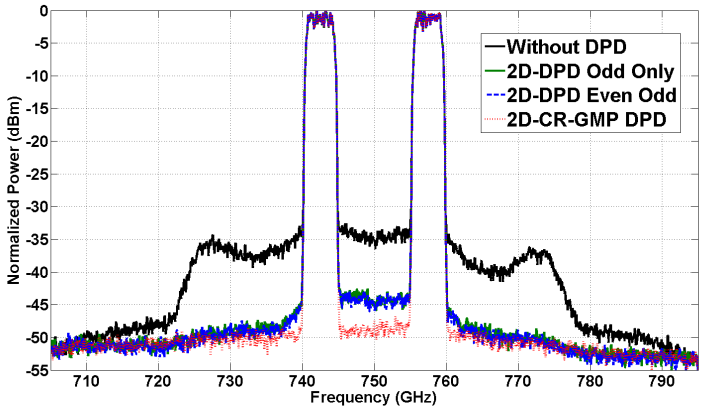

Figure 6.14: 750 MHz band power spectrum density of the GaN Doherty PA driven with a dual-band 20 MHz WCDMA signal, without DPD and when linearized with the 2D-CR-GMP and 2D-DPD linearizers

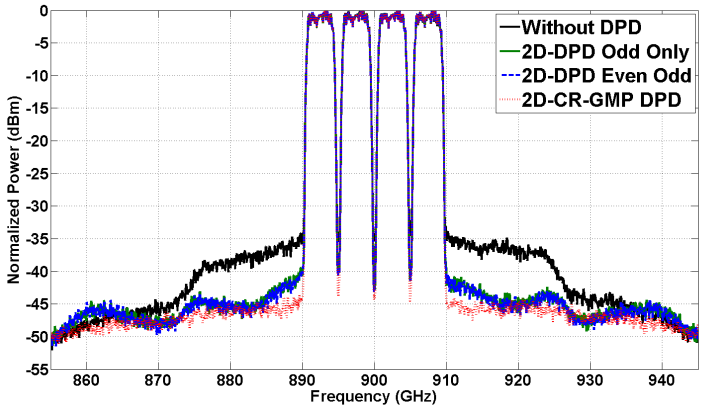

Figure 6.15: 900 MHz band power spectrum density of the GaN Doherty PA driven with a dual-band 20 MHz WCDMA signal, without DPD and when linearized with the 2D-CR-GMP and 2D-DPD linearizers

6.7 Conclusion

This chapter proposes a closed-form deterministic approach to identifying CR-GMP models for multi-band nonlinear PAs modeling/linearization. First, block diagrams that describe the PA behavior under multi-band signals are proposed. Second, derivations of dual- and tri-band models (2D-CR-GMP and 3D-CR-GMP) are detailed. A general form multi-band PA behavioral model/DPD is then proposed. Extensive validations of the 2D-CR-GMP and 3D-CR-GMP models show good linearization capabilities for two different PAs driven with wide-/multi-band signals. The proposed models are compared to contemporary 2D-DPD and 3D-DPD models; and, it is shown that the 2D-CR-GMP is robust to time delay

misalignment between the different input and output signals. The proposed models are inherently low complexity and allow for high nonlinearity and large memory span, while keeping a manageable number of coefficients.

Chapter 7

Conclusion

Emerging wireless networks call for ubiquitous and multi-standard wireless access that satisfies the demands of wireless users by increasing data rates, broadband Internet and video transmission. To ensure such wireless access, radio systems require the deployment of wide-/multi-band signals. These signals impose linearity and efficiency challenges on radio frequency (RF) transmitters, particularly for the power amplifier (PA) stage, due to their high peak-to-average power ratio (PAPR) and large frequency range. PAs are, however, inherently nonlinear in their efficient region. Consequently, to ensure linear operation, PAs need to operate in the back-off region far from their peak power capability, leading to an increase in the operational cost of base stations. Digital predistortion (DPD) has been introduced in the literature to linearize the PA output signal in its nonlinear region, thereby enhancing the trade-off between linearity and efficiency. DPD consists of incorporating an extra nonlinear function before the PA, to preprocess the input signal, so that the cascaded system (DPD+PA) behaves linearly.

In this thesis, it is shown that conventional PA behavioral models/DPD schemes have been limited to single- and narrow-band signals. The direct application of these conventional models to wide-/multi-band signals creates challenging requirements that have not been adequately addressed in the literature. These challenges include unrealistic sampling rates in the observation path of PA input/output signals. In addition, the complexity of existent PA behavioral models/DPD schemes increases exponentially with the PA nonlinearity order and memory taps, leading to an unmanageably high number of coefficients. To that end, newer modeling schemes are needed to tackle these problems, and new approaches are essential to reduce their complexity.

7.1 Summary of Contributions

The main goal of this thesis has been the development of new modeling schemes that can accurately model and linearize wide-/multi-band PAs with reasonable complexity. Physical inspiration and a study of the sources of memory effects have been investigated to derive simple and powerful modeling schemes. First, an innovative pruning technique for the Volterra series based on G-Functionals was proposed. This technique allowed the selection of dominant kernels without the estimation of the nonlinearity order or the memory depth. This pruning results in a reduced-complexity behavioral model/DPD that allows high values for the nonlinearity order and memory depth while keeping a manageable number of coefficients. Therefore, this pruning technique is suitable for wideband signals. In addition, it allows for better numerical stability by achieving a smaller conditioning number and coefficient range than the widely used dynamic deviation reduction (DDR) approach ones.

For wideband PAs, the conventional DPD identification procedure requires the deployment of a transmitter observation receiver (TOR) with a bandwidth that is equal to at least five times the input signal bandwidth. This requirement restricts the effectiveness of the DPD to only narrow-band signals. In fact, the high complexity and the high cost to achieve a wideband TOR bandwidth degrade the effectiveness of the DPD in trading off linearity and efficiency. In this context, a novel procedure was proposed to allow the construction of a DPD while using a TOR bandwidth that is equal to only twice the input signal bandwidth. This $^2/_5$ reduction allowed the application of DPD to wideband signals without extensively increasing the cost of the DPD implementation and complexity.

For multi-band PAs, a new approach for generating behavioral model schemes is proposed. It starts with simplified block diagrams that characterize the behavior of a typical PA driven by a multi-band signal. Second complexity-reduced generalized memory polynomial (CR-GMP) models were derived from the feedback PA block diagram for the case of single-, dual-, tri-band cases. Then a general formulation for the multi-band case was proposed. The resulting CR-GMP models were inherently low complexity and allowed good linearization while deploying fewer coefficients. In addition, the proposed models were proven to be robust against time delay misalignment.

The thesis has developed of reduced complexity behavioral models/DPD schemes that are capable of predicting and linearizing different PA topologies and technologies output signals. These PAs were driven with wideband (up to 60 MHz) and multi-band signals with widely spaced carrier frequencies (higher than 100 MHz).

7.2 Future Work

In Chapter 4, it was demonstrated that the TOR bandwidth can be reduced without affecting the DPD linearization performances. However, one critical condition for that reduction was the deployment of a wideband PA input signal, as shown in Figure 7.1. This wideband PA input signal can be synthesized digitally, i.e., as a DPD output signal, then converted to analog and up-converted to RF. This process calls for the deployment of high speed digital-to-analog converters (DACs) and consequently limits the effectiveness of the DPD when applied for wideband signals. Similar to the TOR bandwidth reduction, the transmitter path sampling speed can be restricted to only twice the input Nyquist rate (INR), i.e., a reduction of the DAC speed, as shown in Figure 7.2. To ensure a wideband PA input signal, a nonlinear analog circuit can be cascaded between the DPD and the PA. Doing so allows the expansion of the DPD output signal bandwidth to five times the original input signal bandwidth. Consequently, the DAC speed is reduced and the TOR bandwidth reduction condition is satisfied, simultaneously. One way to implement this nonlinear system is analog predistortion. In this way, DPD and analog predistortion are combined, thus reducing DAC speeds, thanks to the analog predistortion function, and reducing the analog-to-digital (ADC) speed while achieving excellent linearization capabilities, thanks to the DPD. However, combining DPD and analog predistortion calls for new identification algorithms to take into account the presence of the two linearizers.

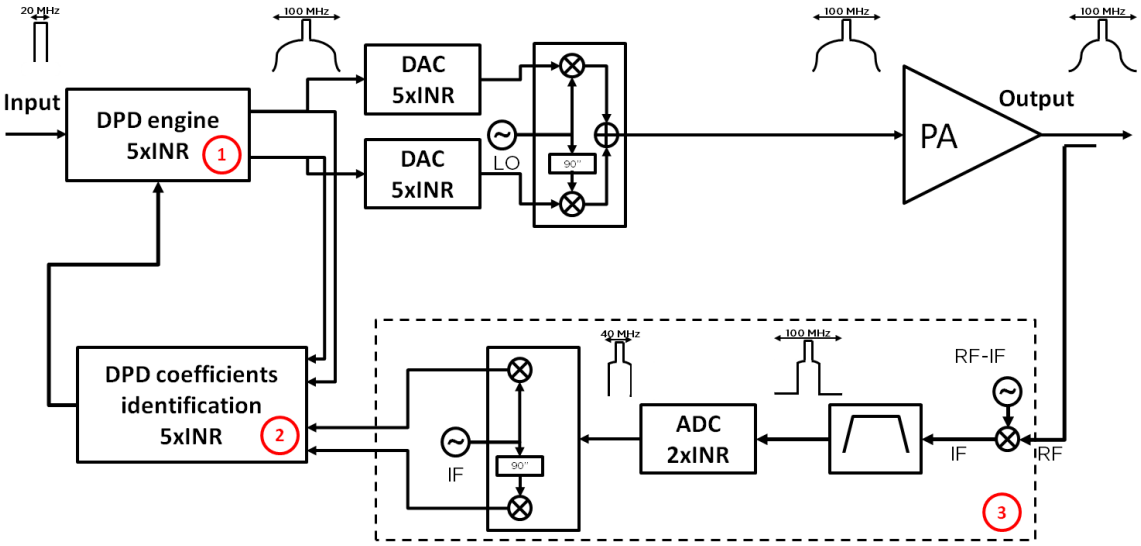

Figure 7.1: Linearization block diagram with reduced speed analog-to-digital converter

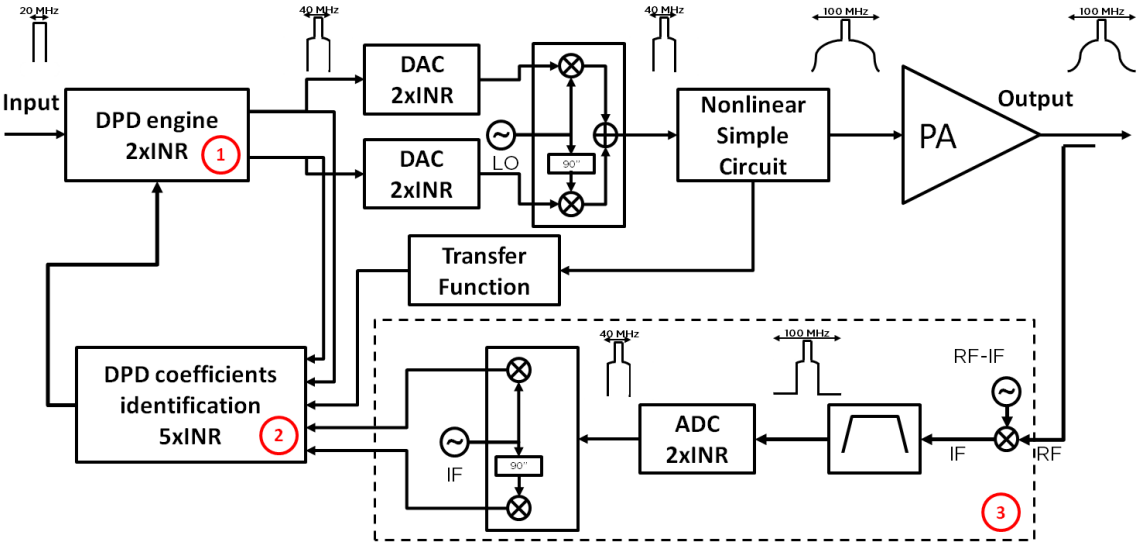

Figure 7.2: Linearization block diagram with reduced speed analog-to-digital and digital-to-analog converters

7.3 List of Publication

Patent Application

1. F. Mkadem, and S. Boumaiza, "Digital Predistortion of Wideband Power Amplifier with Reduced Observation Bandwidth," *United States Letters Patent*, 1557-101U (P40849_US2)

Peer-Reviewed Journal Papers

1. F. Mkadem, and S. Boumaiza, "Multi-Band Complexity-Reduced Generalized-Memory-Polynomial Power Amplifier Digital Predistortion," *Submitted to IEEE Transaction on Microwave Theory and Techniques*, July 2014

2. F. Mkadem, and S. Boumaiza, "Digital Predistortion of Wideband Power Amplifiers with Reduced Transmitter Observation Receiver Bandwidth," *Submitted to IEEE Transaction on Microwave Theory and Techniques*, May 2014

3. F. Mkadem, M. C. Fares, S. Boumaiza, and J. Wood, "Complexity-Reduced Volterra Series Model for Power Amplifier Digital Predistortion," *Analog Integrated Circuits and Signal Processing*, vol. 79, issue 2, pp 331-343, May 2014

Magazine

1. N. Mrabet, F. Mkadem, O. Petelin, H. Medini, and S. Boumaiza, "Mixed-Signal Multimode Radio Software/Hardware Development Platform," *IEEE Microwave Magazine*, vol. 14, no. 1, pp. 128-133, January 2013

Papers in Refereed Conference Proceedings

1. F. Mkadem, and S. Boumaiza, "Nonlinear System Behavioral Modeling Using Reduced Transmitter Observation Receiver Bandwidth," 83^{rd} *ARFTG*, Tampa, Florida, June 2014

2. F. Mkadem, D. Y. Wu, and S. Boumaiza, "G-Functionals for Nonlinear Power Amplifier Digital Predistortion," *IEEE MTT International Microwave Symposium*, Montreal, QC, Canada, 2012

3. N. Mrabet, I. Mohammad, F. Mkadem, C. Rebai, and S. Boumaiza, "Optimized Hardware for Polynomial Digital Predistortion System Implementation," *Radio Wireless Week*, Santa Clara, CA, 2012

4. F. Mkadem, S. Boumaiza, J. Staudinger, and J. Wood, "Systematic Pruning of Volterra Series using Wiener G-Functionals for Power Amplifier and Predistorter Modeling," *European Microwave Week*, Manchester, UK, 2011

5. S. Boumaiza, F. Mkadem, and M. Ben Ayed, "Digital Predistortion Challenges in the Context of Software Defined Transmitters, *XXX Symposium of the URSI*, Istanbul, Turkey, 2011

Appendix A

Three-tone Characterization

This appendix illustrates that modeling a memoryless PA, with an order of nonlinearity equal to 3, and driven by a three-tone signal, using only the inner tones, is not practical, as suggested by [89]. For that, let us consider a memoryless PA model given in (A.1); the goal is the estimation of a_1 and a_3.

$$\begin{aligned} y(n) &= a_1 x(n) + a_3 x(n)^3 \\ a_1 &\approx 0.977 + 0.0591i \\ a_3 &\approx \frac{2.197 - 3.907i}{1000} \end{aligned} \tag{A.1}$$

Figure A.1 shows the PA input and output spectrums; i.e., blue is for PA input spectrum and red is for output one. It can be observed that the 3^{rd} order inter-modulation (IMD3) is quite high, only 20 dBc from the in-band tones. However, it is also observed that the in-band amplification is barely noticed. In fact, zooming into the three in-band tones frequency range, Figure A.2, shows a difference of less than 0.01 dB. Consequently, it is deduced that: theoretically, only the in-band tones are necessary to estimate a_1 and a_3 (by measuring two of the in-band output tones and deducing the amplification difference). However, in practice, measurement errors can affect the estimation of this difference and consequently the accuracy of a_1 and a_3 modeling. For example, the R&S®FSW signal and spectrum analyzer used in the measurement of Chapter 4 has a measurement uncertainty of less than 0.4 dB. This measurement uncertainty is higher than the difference that should be captured to accurately estimate a_1 and a_3. Hence, it was concluded in Chapter 4 that IMD3 measurement is necessary to mitigate the a for-mentioned problem.

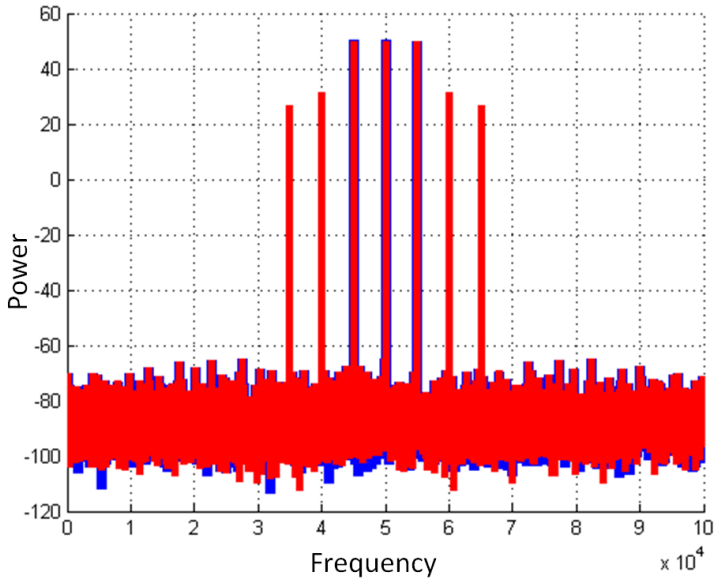

Figure A.1: Input and output spectrums of a memoryless nonlinear power amplifier driven by a three-tone signal

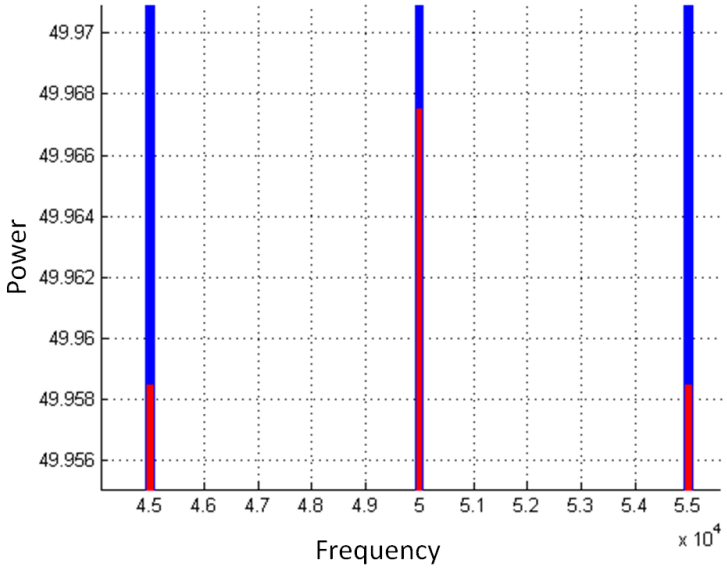

Figure A.2: Zoom into the in-band of the input and output spectrums of a memoryless nonlinear power amplifier driven by a three-tone signal

Appendix B

Over Modeling with Reduced Transmitter Observation Receiver

This appendix use an example to illustrate that the sampling rate reduction depends on the model structure. In other words, when the model structure is comparable to the real structure of the system being modeled, reducing the sampling rate is more likely.

Let us consider a system being represented by a polynomial given in (B.1), where $x \in [-1, 1]$. Figures B.1 and B.2 illustrate the curves of y as a function of x (the blue curve is for the original polynomial). The goal here is the estimation of this polynomial.

$$y(n) = -1.0462\, x + 0.2403\, x^4 - 1.0864\, x^5 + 1.3511\, x^7 + 0.0303\, x^9 + 0.9680\, x^{13} \quad \text{(B.1)}$$

A first test consists of sampling only four points from x and y, i.e.,$\{\ (-1, 0.0234), (-0.34, 0.3631), (0.32, -0.3354), (0.98, 0.157)\ \}$, and finding a polynomial model that best fits these points. In this test, the nonlinearity of the polynomial is selected to be exactly the same as that of the polynomial being modeled as given in (B.2); i.e., the model structure is identical to the original polynomial (correct structure). MATLAB identification results in a polynomial (represented by the red curve in Figure B.1) that perfectly fits the original polynomial for the sampled points and perfectly extrapolates elsewhere.

$$y(n) = a_1 x^1 + a_4 x^4 + a_5 x^5 + a_7 x^7 + a_9 x^9 + a_{13} x^{13} \quad \text{(B.2)}$$

A second test consists of sampling the same four points from x and y and finding a polynomial model that best fits these points. In this test, the nonlinearity of the polynomial

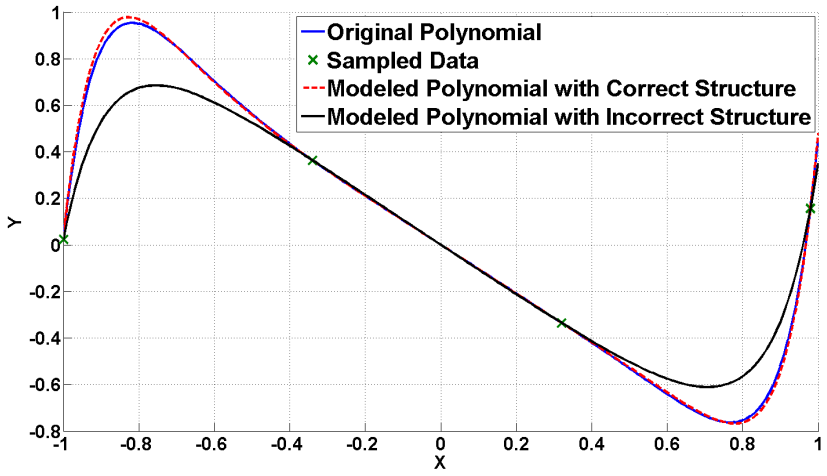

Figure B.1: Modeling using four sampling points

is selected to be 13, as given in (B.3); i.e., the model structure is unlike the original polynomial (incorrect structure). In fact, in this case, the polynomial in (B.3) contains more terms than needed to represent exactly the structure of (B.1). MATLAB identification results in a polynomial (represented by the black curve in Figure B.1) that perfectly fits the sampled points but fails to extrapolate elsewhere. To solve this issue, if the sampled points are augmented to eight instead of four, MATLAB identification now results in a model that perfectly fits the sampled points and perfectly extrapolate elsewhere.

$$y(n) = b_1 x^1 + b_2 x^2 + \cdots + b_{13} x^{13} \tag{B.3}$$

It is important to note that sampling eight or four points is a non-deterministic problem as the number of coefficients is higher that the sampled points. This problem is equivalent to sampling the PA output signal using a TOR bandwidth that is less than the PA output bandwidth. However, when the model structure is close to the system being modeled, fewer samples are needed. However, augmenting the number of samples allows more flexibility in the structure choice. This is equivalent to the case in Chapter 4 where; if the DPD structure is correct, the sampling rate is reduced to as low as the input Nyquist rate (INR). However, when the DPD structure is incorrect, the sampling rate is reduced to only $2 \times INR$.

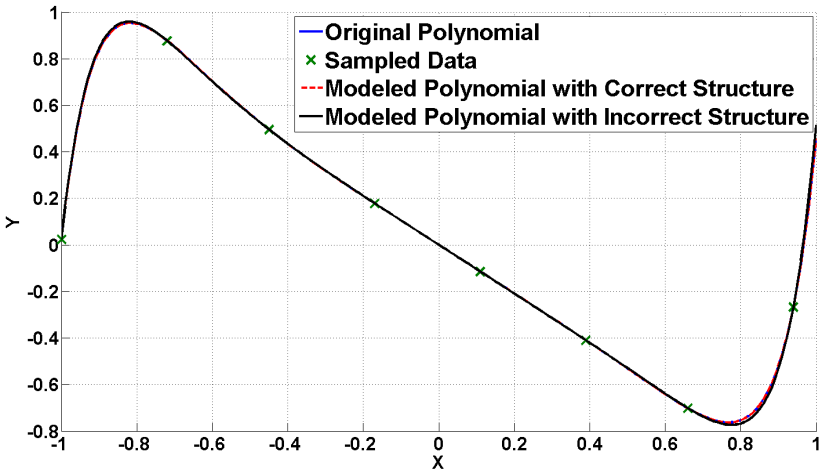

Figure B.2: Modeling using eight sampling points

References

[1] Universal mobile telecommunications system (UMTS); base station (BS) radio transmission and reception (FDD) (3GPP TS 25.104 version 6.7.0 Release 6). Etsi ts 125 104 v6.7.0, European Telecommunications Standards Institute, CH-1211 Geneva, Switzerland, September 2004. 4, 17, 72, 76, 118

[2] HSPA to LTE-Advanced: 3GPP broadband evolution to IMT-Advanced (4G). Technical report, Raysavy Research, LLC, September 2009. 2

[3] Connecting America: the national broadband plan. Technical report, Federal Communications Commission, 2011. 1

[4] P. H. Aaen, J. A. Plá, and J. Wood. *Modeling and characterization of RF and microwave power FETs*. Cambridge press, 2007. 23, 24, 84

[5] A. Ahmed, E.R. Srinidhi, and G. Kompa. Efficient PA modeling using neural network and measurement setup for memory effect characterization in the power device. *IEEE MTT-S International Microwave Symposium Digest*, Jun. 12 – Jun. 17 2005. 25

[6] O. Andersen, N. Bjorsell, and N. Keskitalo. A test-bed designed to utilize Zhu's general sampling theorem to characterize power amplifiers. In *IEEE Instrumentation and Measurement Technology Conference, I2MTC*, pages 201–204, May 2009. 32

[7] L. Anttila, P. Handel, and M. Valkama. Joint mitigation of power amplifier and I/Q modulator impairments in broadband direct-conversion transmitters. *IEEE Transactions on Microwave Theory and Techniques*, 58(4):730–739, April 2010. 7

[8] 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*, 59(10):2547–2553, October 2011. xiv, 6, 40, 42, 43, 102

[9] V. A. Bohara, M . Abi Hussein, and O. Venard. A parameter identification algorithm for multi-stage digital predistorter. In *European Microwave Conference (EuMC)*, pages 416–419, October 2013. 7

[10] N. Boulejfen, A. Harguem, and F. M. Ghannouchi. New closed-form expressions for the prediction of multitone intermodulation distortion in fifth-order nonlinear RF circuits/systems. *IEEE Transactions on Microwave Theory and Techniques*, 52(1):121–132, January 2004. 16

[11] S. Boumaiza. Advanced memory polynomial linearization techniques. In *Workshop on Power Amplifier Linearization: From Advanced Analog and Digital Techniques to Practical Real-Time Implementations*, June 2009. 27

[12] S. Boumaiza and F. Mkadem. Wideband RF power amplifier predistortion using real-valued time-delay neural networks. In *European Microwave Conference*, pages 1449–1452, October 2009. 25

[13] S. Boumaiza, F. Mkadem, and M. B. Ayed. Digital predistortion challenges in the context of software defined transmitters. In *XXXth URSI General Assembly and Scientific Symposium*, pages 1–4, August 2011. 3

[14] S. Boyd, L. O. Chua, and C. A. Desoer. Ima journal of mathematical control and information. *Oxford University Press*, 1(3):243–282, 1984. 25

[15] 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*, pages 1517–1520, Jun. 2008. 33

[16] M. Cabarkapa, N. Neskovic, and D. Budimir. 2-D nonlinearity compensation technique for concurrent dual-band wireless transmitters. In *IEEE MTT-S International Microwave Symposium Digest*, pages 1–3, June 2013. 42

[17] H. Cao, A. S. Tahrani, C. Fager, T. Eriksson, and H. Zirath. I/Q imbalance compensation using a nonlinear modeling approach. *IEEE Transactions on Microwave Theory and Techniques*, 57(3):513–518, March 2009. 7

[18] H. H. Chen, C. S. Maa, Y. C. Wang, and J. T. Chen. Dynamically optimum lookup-table spacing for power amplifier predistortion linearization. *IEEE Transactions on Microwave Theory and Techniques*, 54(5):2118–2127, May 2006. 5, 24

[19] S. Chen, X. Hong, Y. Gong, and C. J. Harris. Digital predistorter design using B-spline neural network and inverse of De Boor algorithm. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 60(6):1584–1594, June 2013. 7

[20] K. J. Cho, J. H. Kim, and S. P. Stapleton. A highly efficient Doherty feedforward linear power amplifier for WCDMA base-station applications. *IEEE Transactions on Microwave Theory and Techniques*, 53(1):292–300, January 2005. 3

[21] A. S. Ciminski. Recurrent neural networks usefulness in digital predistortion of power amplifiers. *15th International Conference on. Microwaves, Radar and Wireless Communications, MIKON2004*, 1:249–252, May 17 – May 19 2004. 25

[22] M. Cooper. The myth, the law, and the spectrum. *IEEE Spectrum*, 38(1):62–63, January 2006. 2

[23] G. Coviello, F. Cannone, and G. Avitabile. Robust behavioral non uniform look-up table spacing in adaptive digital baseband predistortion technique for RF power amplifier. In *AFRICON*, pages 1–5, September 2013. 5, 24

[24] S. C. Cripps. *RF Power Amplifiers for Wireless Communications*. Artech House, 2^{nd} edition, 2006. 4, 12

[25] T. R. Cunha, J. C. Pedro, and E. G. Lima. Low-pass equivalent feedback topology for Power Amplifier modeling. In *IEEE MTT-S International Microwave Symposium Digest*, pages 1445–1448, Jun. 2008. 84

[26] G. Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4):303–314, 1989. 25

[27] M. Djamai, S. Bachir, and C. Duvanaud. Kalman filtering algorithm for on-line memory polynomial predistortion. In *European Microwave Conference (EuMC)*, pages 575–578, October 2008. 7

[28] Y. H. Fang, M.C.E. Yagoub, F. Wang, and Q.J. Zhang. A new macromodeling approach for nonlinear microwave circuits based on recurrent neural networks. *IEEE Transaction on Microwave Theory and Techniques*, 48(12):2335–2344, Dec. 2000. 25

[29] M. Faulkner. Amplifier linearization using RF feedback and feedforward techniques. *IEEE Transaction on Vehicular Techniques*, 47(1):209–215, February 1998. 4, 19

[30] B. Fehri and S. Boumaiza. Baseband equivalent Volterra series for digital predistortion of dual-band power amplifiers. *IEEE Transactions on Microwave Theory and Techniques*, 62(3):700–714, March 2014. 6, 40, 41, 43

[31] F. Filicori and G. Vannini. Mathematical approach to large-signal modelling of electron devices. *Electronics Letters*, 27(4):357–359, February 1991. 5

[32] W. A. Frank. Sampling requirements for Volterra system identification. *IEEE Signal Processing Letters*, 3(9):266–268, September 1996. 32

[33] R. Friedrich, K. Sabbagh, B. El-Darwiche, and M. Singh. Trends in telecommunication reform 2010/11, enabling tomorrow's digital world. 11^{th} edition, International Telecommunication Union, CH-1211 Geneva, Switzerland, March 2011. 1, 3

[34] K. I. Funahashi. On the approximate realization of continuous mappings by neural networks. *Neural Networks*, 2(3):183–192, 1989. 25

[35] F. M. Ghannouchi, S. A. Bassam, M. Helaoui, and A. Kwan. Digital multi-band predistortion linearizer with nonlinear subsampling algorithm in the feedback loop, April 2013. US Patent App. 13/274,290. 41

[36] F. M. Ghannouchi and O. Hammi. Behavioral modeling and predistortion. *IEEE Microwave Magazine*, 10(7):52–64, December 2009. 4, 19, 27, 41

[37] P. Gilabert, G. Montoro, and E. Bertran. On the Wiener and Hammerstein models for power amplifier predistortion. In *Asia-Pacific Conference Proceedings*, volume 2, December 2005. 26

[38] L. Guan, C. Yu, and A. Zhu. Bandwidth-constrained least squares-based model extraction for band-limited digital predistortion of RF power amplifiers. In *Workshop on Integrated Nonlinear Microwave and Millimetre-Wave Circuits (INMMIC)*, 2012. 33

[39] L. Guan and A. Zhu. Dual-loop model extraction for digital predistortion of wideband RF power amplifiers. *IEEE Microwave and Wireless Components Letters*, 21(9):501–503, September 2011. 7

[40] L. Guan and A. Zhu. Simplified dynamic deviation reduction-based Volterra model for Doherty power amplifiers. In *Workshop on Integrated Nonlinear Microwave and Millimetre-Wave Circuits*, pages 1–4, April 2011. 5, 28, 29

[41] O. Hammi, S. Boumaiza, M. Jaidane-Saidane, and F. M. Ghannouchi. Digital subband filtering predistorter architecture for wireless transmitters. *IEEE Transactions on Microwave Theory and Techniques*, 53(5):1643–1652, May 2005. 33

[42] F. P. Hart and M. B. Steer. Modeling the nonlinear response of multitones with uncorrelated phase. *IEEE Transactions on Microwave Theory and Techniques*, 55(10):2147–2156, October 2007. 16

[43] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366, 1989. 25

[44] M. Isaksson, D. Wisell, and D. Ronnow. Wide-band dynamic modeling of power amplifiers using radial-basis function neural networks. *IEEE Transaction on Microwave Theory and Techniques*, 53(11):3422–3428, Nov. 2005. 25

[45] M. Isaksson, D. Wisell, and D. Ronnow. A comparative analysis of behavioral models for RF power amplifiers. *IEEE Transactions on Microwave Theory and Techniques*, 54(1):348–359, January 2006. 4, 19

[46] H. Jiang and P. A. Wilford. Digital predistortion for power amplifiers using separable functions. *IEEE Transactions on Signal Processing*, 58(8):4121–4130, August 2010. 22

[47] H. Jiang, X. Yu, and P. A. Wilford. Digital predistortion using stochastic conjugate gradient method. *IEEE Transactions on Broadcasting*, 58(1):114–124, March 2012. 7

[48] J. Kang, J. Yoon, K. Min, D. Yu, J. Nam, Y. Yang, and B. Kim. A highly linear and efficient differential CMOS power amplifier with harmonic control. *IEEE Journal of Solid-State Circuits*, 41(6):1314–1322, June 2006. 3

[49] P.B. Kenington and L. Astier. Power consumption of A/D converters for software radio applications. *IEEE Transactions on Vehicular Technology*, 49(2):643–650, March 2000. 7, 32

[50] W. Kester. Undersampling applications. xiv, 32, 35

[51] J. Kim and K. Konstantinou. Digital predistortion of wideband signals based on power amplifier model with memory. *Electronics Letters*, 37(23):1417–1418, Nov. 2001. 27

[52] J. Kim, F. Mkadem, and S. Boumaiza. A high efficiency and multi-band/multi-mode power amplifier using a distributed second harmonic termination. In *European Microwave Integrated Circuits Conference*, pages 420–423, September 2010. 2

[53] K. Konstantinou and D. K. Paul. Analysis and design of broadband, high efficiency feedforward amplifiers. In *IEEE MTT-S International Microwave Symposium Digest*, volume 2, pages 867–870, June 1996. 19

[54] H. Ku and J. S. Kenney. Behavioral modeling of nonlinear RF power amplifiers considering memory effects. *IEEE Transactions on Microwave Theory and Techniques*, 51(12):2495–2504, December 2003. 26

[55] H. Ku, M. D. Mckinley, and J. S. Kenney. Extraction of accurate behavioral models for power amplifiers with memory effects using two-tone measurements. In *IEEE MTT-S International Microwave Symposium Digest*, volume 1, pages 139–142, June 2002. 16

[56] A. K. Kwan, S. A. Bassam, M. Helaoui, and F. M. Ghannouchi. Concurrent dual band digital predistortion using look up tables with variable depths. In *IEEE Topical Conference on Power Amplifiers for Wireless and Radio Applications (PAWR)*, pages 25–27, January 2013. 5, 24

[57] C. F. Lanzani, G. Kardaras, and D. Boppana. Remote radio heads and the evolution towards 4G networks. *MTI Radiocomp & Altera*, 2009. 3

[58] C. F. Lanzani, G. Kardaras, and D. Boppana. Remote radio heads and the evolution towards 4G networks. White paper, MTI Radiocomp & Altera, February 2009. 7

[59] J. Liszewski, B. Schubert, W. Keusgen, and A. Kortke. Low-complexity FPGA implementation of Volterra predistorters for power amplifiers. In *IEEE Topical Conference on Power Amplifiers for Wireless and Radio Applications (PAWR)*, pages 41–44, January 2011. 5

[60] F. Liu, K. Zheng, W. Xiang, and H. Zhao. Design and performance analysis of an energy-efficient uplink carrier aggregation scheme. *IEEE Journal on Selected Areas in Communications*, 32(2):197–207, February 2014. 3

[61] T. Liu, S. Boumaiza, and F. M. Ghannouchi. Augmented hammerstein predistorter for linearization of broad-band wireless transmitters. *IEEE Transactions on Microwave Theory and Techniques*, 54(4):1340–1349, June 2006. 26

[62] K. Madani. Reducing the intermodulation distortion in multi-carrier microwave power amplifiers. In *Symposium on High Performance Electron Devices for Microwave and Optoelectronic Applications*, pages 153–157, November 1999. 4

[63] N. Messaoudi, M. C. Fares, S. Boumaiza, and J. Wood. Complexity reduced odd-order only memory polynomial pre-distorter for multi-carrier doherty power amplifier linearization. In *IEEE MTT-S International Microwave Symposium Digest*, pages 419–422, June 2008. 16, 27

[64] D. Mirri, G. Luculano, F. Filicori, G. Pasini, G. Vannini, and G. P. Gabriella. A modified Volterra series approach for nonlinear dynamic systems modeling. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 49(8):1118–1128, August 2002. 5

[65] F. Mkadem. Behavioural modeling and linearization of RF power amplifier using artificial neural networks. Master's thesis, University of Waterloo, 200 University Ave. W., Waterloo, ON, N2L 3G1, Canada, 2010. 25

[66] F. Mkadem, M. B. Ayed, S. Boumaiza, J. Wood, and P. Aaen. Behavioral modeling and digital predistortion of power amplifiers with memory using two hidden layers artificial neural networks. In *IEEE MTT-S International Microwave Symposium Digest*, pages 656–659, May 2010. 25

[67] F. Mkadem and S. Boumaiza. Digital predistortion of wideband power amplifiers with reduced transmitter observation receiver bandwidth. *IEEE Transactions on Microwave Theory and Techniques*. 8

[68] F. Mkadem and S. Boumaiza. Multi-band generalized-memory-polynomial digital predistortion for multi-band power amplifier. *IEEE Transactions on Microwave Theory and Techniques*. 9

[69] F. Mkadem and S. Boumaiza. Extended Hammerstein behavioral model using artificial neural networks. *IEEE Transactions on Microwave Theory and Techniques*, 57(4):745–751, April 2009. 25, 26, 27

[70] F. Mkadem and S. Boumaiza. Extended hammerstein model for rf power amplifier behavior modeling. In *European Microwave Conference, EuMC*, October 2009. 25, 26, 27

[71] 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*, 59(4):913–923, April 2011. 7, 25, 28

[72] F. Mkadem and S. Boumaiza. Nonlinear system behavioral modeling using reduced transmitter observation receiver bandwidth. In 83^{rd} *ARFTG*, June 2014. 81

[73] F. Mkadem, S. Boumaiza, J. Staudinger, and J. Wood. Systematic pruning of Volterra series using Wiener G-functionals for power amplifier and predistorter modeling. In *European Microwave Integrated Circuits Conference (EuMIC)*, pages 482–485, October 2011. 8, 46

[74] F. Mkadem, M. C. Fares, S. Boumaiza, and J. Wood. Complexity-reduced Volterra series model for power amplifier digital predistortion. *Analog Integrated Circuits and Signal Processing*, 79(2):331–343, May 2014. 9

[75] F. Mkadem, D. Y. T. Wu, and S. Boumaiza. Wiener G-functionals for nonlinear power amplifier digital predistortion. In *IEEE MTT-S International Microwave Symposium Digest (MTT)*, pages 1–3, June 2012. 8, 46

[76] D. R. Morgan, Z. Ma, J. Kim, M. G. Zierdt, and J. Pastalan. A generalized memory polynomial model for digital predistortion of RF power amplifiers. *IEEE Transactions on Microwave Theory and Techniques*, 54(10):3852–3860, Oct. 2006. 84, 87, 88, 89, 104, 105

[77] N. Mrabet, F. Mkadem, O. Petelin, H. Medini, and S. Boumaiza. Mixed-signal multimode radio software/hardware development platform. *IEEE Microwave Magazine*, 14(1):128–133, January 2013. 28

[78] N. Mrabet, I. Mohammad, F. Mkadem, C. Rebai, and S. Boumaiza. Optimized hardware for polynomial digital predistortion system implementation. In *Radio Wireless Week*, January 2012. 27, 28

[79] A. K. Mrunal, M. Shirasgaonkar, and R. M. Patrikar. Power amplifier linearization using a diode. In *IEEE Mediterranean Electrotechnical Conference*, pages 173–176, May 2006. 24

[80] N. Naraharisetti, P. Roblin, C. Quindroit, M. Rawat, and S. Gheitanchi. 2D Quasi exact inverse of PA model in digital predistorter for concurrent dual-band system. In *IEEE Wirless and Microwave Technology Conference, WAMICON*, June 2014. 41

[81] O. Nelles. *Nonlinear system identification*. Springer, 2001. 26

[82] B. O'Brien, J. Dooley, and T.J. Brazil. RF power amplifier behavioral modeling using a globally recurrent neural network. *IEEE MTT-S International Microwave Symposium Digest*, pages 1089–1092, Jun. 11 – Jun. 16 2006. 25

[83] H. Ozbay. *Introduction to feedback control theory*. CRC Press, 1999. 19, 20

[84] H. Paaso and A. Mammela. Comparison of direct learning and indirect learning predistortion architectures. In *IEEE International Symposium on Wireless Communication Systems, ISWCS '08*, pages 309–313, October 2008. 4, 7

[85] J. C. Pedro, N. B. Carvalho, and P. M. Lavrador. Modeling nonlinear behavior of band-pass memoryless and dynamic systems. In *IEEE MTT-S International Microwave Symposium Digest*, volume 3, pages 2133–2136, Jun. 2003. 84

[86] J. C. Pedro and N. B. de Carvalho. On the use of multitone techniques for assessing rf components' intermodulation distortion. *IEEE Transactions on Microwave Theory and Techniques*, 47(12):2393–2402, December 1999. 16

[87] J. C. Pedro and S. A. Maas. A comparative overview of microwave and wireless power-amplifier behavioral modeling approaches. *IEEE Transactions on Microwave Theory and Techniques*. 4, 19

[88] C. Quindroit, N. Naraharisetti, P. Roblin, S. Gheitanchi, V. Mauer, and M. Fitton. 2D forward twin nonlinear two-box model for concurrent dual-band digital predistortion. In *IEEE Radio and Wireless Symposium*, pages 1–4, January 2014. 6, 40, 41

[89] D. Ronnow, D. Wisell, and M. Isaksson. Three-tone characterization of nonlinear memory effects in radio-frequency power amplifiers. *IEEE Transactions on Instrumentation and Measurement*, 56(6):2646–2657, December 2007. 62, 132

[90] A. A. M. Saleh. Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers. *IEEE Transactions on Communications*, 29(11):1715–1720, November 1981. 24

[91] M. Schetzen. Theory of pth-order inverses of nonlinear systems. *IEEE Transactions on Circuits and Systems*, 23(5):285–291, May 1976. 7, 22, 59, 61

[92] M. Schetzen. *The mathematical theory of communication*. John Wiley & Sons, 1980. 35, 59

[93] M. Schetzen. *Volterra and Wiener theories of nonlinear systems*. John Wiley & Sons, 1980. x, 6, 25, 46, 47, 48, 49, 50

[94] M. Seo, K. Kim, M. Kim, H. Kim, J. Jeon, M. K. Park, H. Lim, and Y. Yang. Ultrabroadband linear power amplifier using a frequency-selective analog predistorter. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 58(5):264–268, May 2011. 24

[95] A. Shukla, B. Willamson, J. Burns, E. Burbidge, A. Taylor, and D. Robinson. A study for the provision of aggregation of frequency to provide wider bandwidth services,

QinetiQ proprietary. Technical report, QinetiQ Proprietary, QINETIQ/06/01773, 2006. 2

[96] H. Songbai, Y. Xiaohuan, and B. Jingfu. Applications of feed-forward neural networks to WCDMA power amplifier model. *Asia-Pacific Conference Proceedings Microwave Conference Proceedings, APMC2005*, 5, Dec 4 – Dec. 7 2005. 25

[97] J. Staudinger, J. C. Nanan, and J. Wood. Memory fading Volterra series model for high power infrastructure amplifiers. In *IEEE Radio and Wireless Symposium*, pages 184–187, January 2010. 6, 18, 87, 104

[98] A. G. Stegmayer and O. Chiotti. Volterra NN-based behavioral model for new wireless communications devices. *Neural Computing and Applications*, 18(3):283–291, Apr. 2009. 25

[99] J. T. Stonick, V. L. Stonick, J. M. F. Moura, and R. S. Zborowski. Memoryless polynomial adaptive predistortion [tv transmitters]. In *International Conference on Acoustics, Speech, and Signal Processing*, volume 2, pages 981–984, May 1995. 24

[100] J. Tsimbinos and K. V. Lever. Input Nyquist sampling suffices to identify and compensate nonlinear systems. *IEEE Transactions on Signal Processing*, 46(10):2833–2837, October 1998. 32

[101] J. Vuolevi and T .Rahkonen. *Distortion in RF power amplifiers*. Artech House, 2003. 25

[102] D. Wisell and P. *Händel*. Simultaneous measurement of transmitter and receiver amplitude and phase ripple. Report trita-ee, KTH EE, January 2009. 7

[103] D. H. Wisell. Exploring the sampling rate requirements for behavioural amplifier, September 2006. 32

[104] J. Wood, M. LeFevre, D. Runton, J.C. Nanan, B. H. Noori, and P. H. Aaen. Envelope-domain time series (ET) behavioral model of a doherty RF power amplifier for system design. *IEEE Transaction on Microwave Theory and Techniques*, 54(8):3163–3172, Aug. 2006. 25

[105] D. Y. T. Wu, F. Mkadem, and S. Boumaiza. Design of a broadband and highly efficient 45W GaN power amplifier via simplified real frequency technique. In *IEEE MTT-S International Microwave Symposium Digest*, pages 1090–1093, May 2010. 53, 74, 91

[106] Y. T. D. Wu. *A doherty power amplifier with extended bandwidth and reconfigurable back-off level*. PhD thesis, University of Waterloo, 200 University Ave. W., Waterloo, ON, N2L 3G1, Canada, 2013. 74, 91

[107] H. Xiang, C. Yu, J. Gao, S. Li, Y. Wu, M. Su, and Y. Liu. Dynamic deviation reduction-based concurrent dual-band digital predistortion. *International Journal of RF and Microwave Computer-Aided Engineering*, 24:401–411, May 2013. 6, 40, 41

[108] D. Xie and Z. Zeng. A new adaptive method for digital predistortion using NLMS and RLS. In M .Zhao and J. Sha, editors, *Communications and Information Processing*, volume 288 of *Communications in Computer and Information Science*, pages 261–268. Springer Berlin Heidelberg, 2012. 7

[109] G. Xu, M. Li, and Y. Xi. Radial basis function neural network models for power-amplifier design. *International Conference on Communications, Circuits and Systems, ICCCAS2004*, 2:1066–1070, Jun. 27 – Jun. 29 2004. 25

[110] Y. Ye, T. Liu, X. Zeng, and J. He. Generalized Hammerstein-based dynamic nonlinear behavior models for wideband RF transmitters. In *International Conference on Wireless Communications, Networking and Mobile Computing, WiCom*, pages 684–687, September 2007. 26

[111] M. Younes and F. M. Ghannouchi. On the modeling and linearization of a concurrent dual-band transmitter exhibiting nonlinear distortion and hardware impairments. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 60(11):3055–3068, November 2013. 42

[112] M. Younes, A. Kwan, M. Rawat, and F. M. Ghannouchi. Linearization of concurrent tri-band transmitters using 3-D phase-aligned pruned Volterra model. *IEEE Transactions on Microwave Theory and Techniques*, 61(12):4569–4578, December 2013. 6, 44

[113] M. Younes, A. Kwan, M. Rawat, and F. M. Ghannouchi. Three-dimensional digital predistorter for concurrent tri-band power amplifier linearization. In *IEEE MTT-S International Microwave Symposium Digest (IMS)*, pages 1–4, June 2013. 6, 44

[114] 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*, 60(12):4198–4208, December 2012. 33, 36, 58, 64

[115] Chao Yu, Lei Guan, and Anding Zhu. Band-limited Volterra series-based behavioral modeling of RF power amplifiers. In *Microwave Symposium Digest (MTT), 2012 IEEE MTT-S International*, pages 1–3, June 2012. 36

[116] Z. Zeng and D. Xie. A new adaptive algorithm for digital predistortion using GNGD with momentum factor. In M. Zhao and J . Sha, editors, *Communications and Information Processing*, volume 288 of *Communications in Computer and Information Science*, pages 446–453. Springer Berlin Heidelberg, 2012. 7

[117] D. Zhou and V.E. DeBrunner. Novel adaptive nonlinear predistorters based on the direct learning algorithm. *IEEE Transactions on Signal Processing*, 55(1):120–133, January 2007. 7

[118] A. Zhu and T. J. Brazil. RF power amplifier behavioral modeling using Volterra expansion with Laguerre functions. In *IEEE MTT-S International Microwave Symposium Digest*, June 2005. 5

[119] A. Zhu, P. J. Draxler, J. J. Yan, T. J. Brazil, D. F. Kimball, and P. M. Asbeck. Open-loop digital predistorter for RF power amplifiers using dynamic deviation reduction-based Volterra series. *IEEE Transactions on Microwave Theory and Techniques*, 56(7):1524–1534, July 2008. 32

[120] A. Zhu and J. C. Pedro. Distortion evaluation of RF power amplifiers using dynamic deviation reduction based Volterra series. In *IEEE/MTT-S International Microwave Symposium*, pages 965–968, June 2007. 5, 28, 29

[121] 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*, 54(12):4323–4332, December 2006. 5, 28, 29

[122] A. Zhu, J. C. Pedro, and T. R. Cunha. Pruning the Volterra series for behavioral modeling of power amplifiers using physical knowledge. *IEEE Transactions on Microwave Theory and Techniques*, 55(5):813–821, May 2007. 5, 28, 88

[123] Y. M. Zhu. Generalized sampling theorem. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 39(8):587–588, Aug. 1992. 8, 32, 59