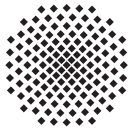


University of Stuttgart
Germany

Institute of
Robust Power
Semiconductor Systems

Master's Thesis

Digital Pre-Distortion of Broadband Communication Links using Open-Loop Architecture

Florian Wiewel

Supervisor: M. Sc. Benjamin Schoch
Prof. Dr.-Ing. Ingmar Kallfass

Period: 15.04.2021 – 05.10.2021

Stuttgart, 05.10.2021

Postanschrift: Institut für Robuste Leistungshalbleitersysteme
Pfaffenwaldring 47

D-70569 Stuttgart

Tel.: +49 (0) 711 685 68700
Fax.: +49 (0) 711 685 58700
E-Mail: sekretariat@ilh.uni-stuttgart.de
Web: www.ilh.uni-stuttgart.de

Declaration

I hereby declare that this thesis is my own work and effort and follows the regulations related to good scientific practice of the University of Stuttgart in its latest form. All sources cited or quoted are indicated and acknowledged by means of a comprehensive list of references.

Stuttgart, 05.10.2021

Florian Wiewel

Executive Abstract

The ever-increasing demand for higher data rates and lower latency in wireless communications ultimately forces the developers of the underlying systems to use broadband links with higher order modulation formats to meet this demand. The use of these modulation formats results in strict linearity requirements on the system. On top of that power efficiency is also an important aspect, since it has an impact on the operational costs and ecological compatibility. In a typical macrocell for modern wireless communication systems the power amplifier (PA) of the base station consumes about 60% of the overall power and it is also the PA, which typically exhibits the strongest nonlinear transfer characteristic in the system. Unfortunately, power efficiency and linearity represent conflicting requirements in PA design. As a result a compromise between these two requirements has to be made. Usually the PA is designed for high power efficiency and with moderate nonlinear transfer characteristics. To compensate for the nonlinearity in the PA a technique called digital pre-distortion (DPD) is applied, which estimates the nonlinearity in the PAs transfer characteristic and applies the corresponding inverse transfer function to the complex baseband input signal of the PA in the digital domain. In contrast to many of the DPD experiments found in literature, which are applied to signals with bandwidths in the range of tens of megahertz, the targeted linearization bandwidth in this work is 5 GHz. For this purpose an open-loop DPD architecture based on the Volterra theory of nonlinear systems specifically the p-th order inverse has been implemented in software and applied to different amplifiers including a waveguide E band transmitter operating around 73 GHz. Up to a signal bandwidth of around 1 GHz significant improvements in terms of signal quality could be observed in the conducted experiments. For signals with higher bandwidths problems with signal integrity caused the DPD to fail. Finally the various problems are analyzed and potential improvements for increasing the DPD performance for wideband signals are suggested.

Zusammenfassung

Die ständig steigenden Anforderungen bezüglich Datenrate und Geschwindigkeit an drahtlose Kommunikationssysteme zwingen die Entwickler dieser System zur Nutzung von immer breitbandigeren Übertragungskanälen und höherer wertigen Modulationsverfahren. Die Nutzung dieser Modulationsverfahren setzt ein lineares Übertragungsverhalten des Systems voraus. Nur so kann die notwendige Signalqualität erreicht werden. Zugleich wird auch die Energieeffizienz der Systeme wichtiger, da sie einen entscheidenden Anteil zu den Betriebskosten beträgt und die ökologische Verträglichkeit mitbestimmt. In einer typischen Makrozelle eines drahtlosen Kommunikationssystems verbraucht der Leistungsverstärker der Basisstation ca. 60% der Gesamtleitung und es ist auch dieser Verstärker, der typischerweise die größten Probleme in Bezug auf Linearität für das Gesamtsystem darstellt. Unglücklicherweise stellen hohe Energieeffizienz und lineares Übertragungsverhalten gegensätzliche Anforderungen bei der Entwicklung eines Leistungsverstärkers dar. Daher wird ein gewisses Maß an Nichtlinearität im Übertragungsverhalten des Leistungsverstärker akzeptiert aber dafür eine hohe Energieeffizienz angestrebt. Um die Nichtlinearität zu kompensieren, wird eine digitale Vorverzerrung verwendet. Diese Vorverzerrung basiert auf einer Schätzung der vorhandenen Nichtlinearität im Übertragungsverhalten des Verstärkers und wendet die entsprechende inverse Übertragungsfunktion auf das komplexe Eingangssignal des Verstärkers in der digitalen Domäne an. Im Gegensatz zu vielen in der Literatur zu findenden Experimente zur digitalen Vorverzerrung, deren Linearisierungsbandbreite häufig im unteren Megahertz-Bereich liegt, beträgt die angestrebte Linearisierungsbandbreite in dieser Arbeit 5 GHz. Zu diesem Zweck wurde eine digitale Vorverzerrung mit offener Regelschleife in Software implementiert, die auf der Volterra Theorie für nichtlineare Systeme beziehungsweise der Inversen p-ter Ordnung basiert. Der entsprechende Algorithmus wurde auf verschiedene Verstärker und Signale mit unterschiedlicher Bandbreite angewendet. Unter anderem wurde ein E-Band Verstärker, der im Frequenzbereich um 73 GHz arbeitet, verwendet. Dabei konnten signifikante Verbesserungen der Signalqualität bis zu einer Linearisierungsbandbreite von ca. 1 GHz erreicht werden. Bei Signalen mit noch größerer Bandbreite verhinderten Probleme mit der Signalintegrität eine erfolgreiche Anwendung der digitalen Vorverzerrung. Abschließend wurden die verschiedenen Probleme analysiert und potentielle Verbesserungsmöglichkeiten vorgeschlagen, die eine erfolgreiche Anwendung der digitalen Vorverzerrung bei großen Signalbandbreiten ermöglichen könnten.

Contents

List of Figures

List of Tables

1. Introduction

The rapid digitalization of more and more aspects of life such as entertainment, healthcare, internet of things, industry 4.0, autonomous driving, etc. generates a demand for more sophisticated communication systems, which can keep up with the ever increasing requirements on bandwidth and latency. As a result the fifth generation of technology standard for broadband cellular networks (5G) is being developed and widely deployed in the near future. It will provide ultra low latency and up to 10 times higher data rates compared to the previous generation (4G) [10]. These improvements of course come at a cost. Developers must find new ways to transmit the massive amount of information while meeting regulatory requirements resulting from a densely packed frequency allocation plan and also keeping power consumption and equipment cost as low as possible.

1.1. Motivation

A simplified block diagram of a typical wireless transmitter is depicted in Figure 1.1. The payload data is outputted by a baseband processing block, which generates a complex signal containing I and Q data. Each of these signals is then converted from the digital domain into the analog domain by a digital to analog converter (DAC). Lowpass filters are an integral component of the digital to analog conversion process, which filter out unwanted mirror images of the baseband signal appearing at integer multiples of the DACs sampling frequency. An IQ modulator is used to map the complex baseband signal to the corresponding modulation symbols. This process additionally includes shifting the modulated signal to a low intermediate frequency (IF). A mixer takes the modulated signal and a carrier signal generated by a local oscillator (LO) and shifts the modulated signal to the desired radio frequency (RF). The mixing process also generates unwanted spectral components, which are filtered out by a subsequent bandpass filter. The RF signal is then amplified by a PA. The amplification is necessary in order to achieve the signal power required to overcome the losses (mainly free space path loss and atmospheric absorption) of the wireless link and allow the receiver to pick up enough signal energy for error free demodulation/decoding of the received signal. Since the PA is a non-ideal component, it also generates unwanted spectral components during the amplification process. Before the RF signal is passed on to the antenna for radiation it is again passed through a bandpass filter to make sure that only signal components inside the desired frequency band are present.
Based on the same principle of operation the presented typical wireless transmitter architecture can be converted into a corresponding receiver architecture by reversing the direction of the signal flow. However, in this thesis only the wireless transmitter will be considered, since the distortions

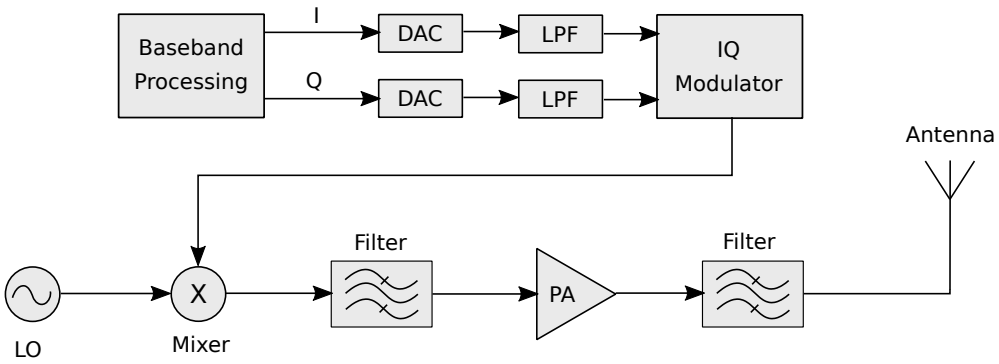


Figure 1.1.: Block diagram - typical wireless transmitter.

introduced by the PA in the wireless transmitter are usually more severe than in a wireless receiver. A wide variety of different wireless transmitter architectures (e.g. heterodyne conversion, direct (zero-IF) conversion, etc.) exist today. They utilize different modulation and up-/downconversion schemes as well as different digital/analog domain boundaries. Each of these architectures exhibits a variety of advantages and disadvantages. Depending on the desired application a suitable architecture is selected and implemented. But in the end all of these architectures contain a PA, which is a major factor in terms of power consumption and consequently also cost of operation. In Figure 1.2 the distribution of power consumption in different base station types for modern wireless communication systems is shown. Depending on the type of base station the contribution of the PA to the total power consumption ranges from 59% in a macrocell base station to 22% in a femtocell base station. On the other hand the contribution of the base band processing ranges

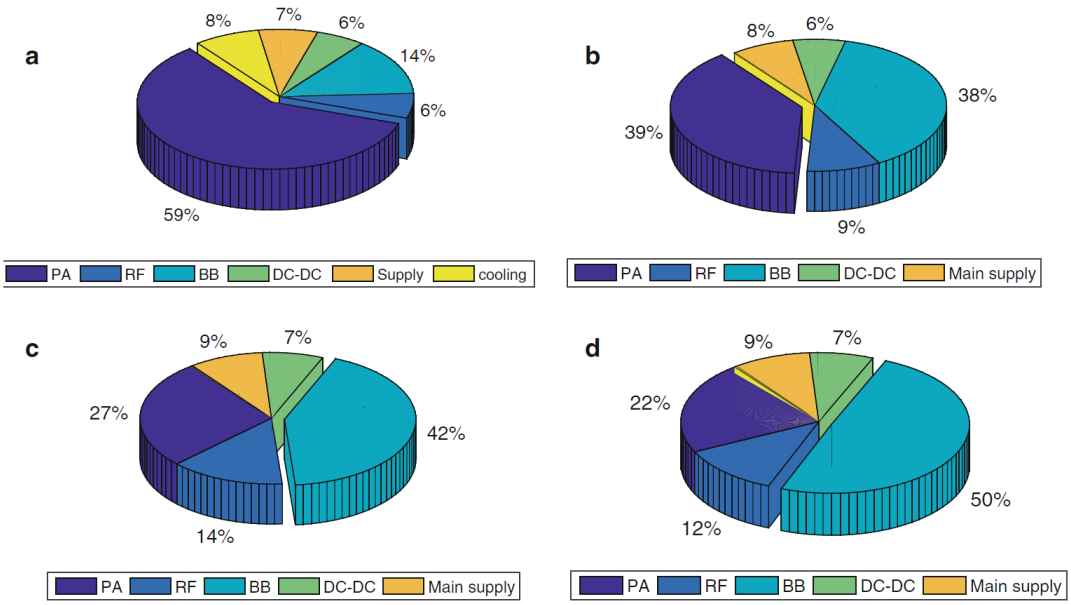


Figure 1.2.: Overview - base station power consumption: a) macrocell b) microcell c) picocell d) femtocell [8].

from 14% in a macrocell base station to 50% in a femtocell base station. So depending on the type of base station a trade-off between optimizing the power consumption of the PA and the baseband processing has to be made in order to reduce the total power consumption.
In addition to the reduction in power consumption an increased data rate is also a requirement imposed on next generation communication systems. One way of increasing the data rate of a communication link is to use a modulation scheme with a higher spectral efficiency. This results in more information bits being contained in one transmitted symbol. But usually modulation schemes with higher spectral efficiency also impose higher requirements on the linearity of the transmitters PA to maintain an acceptable bit error rate and limit spectral regrowth into neighboring channels at the same average output power level. To meet the linearity requirement, the power level of the input signal to the PA has to be reduced so that the PA is still operated in its linear region. This process is called power backoff. Unfortunately the efficiency of a typical PA is highest when operating near saturation. Therefore the efficiency of the PA is reduced when power backoff is performed.
To overcome the dilemma created by requiring high linearity while also maintaining high power efficiency in a PA, a technique called DPD has been created. The basic concept of DPD is illustrated in Figure 1.3. A signal processing block (DPD) is placed in front of the PA. It is located in the

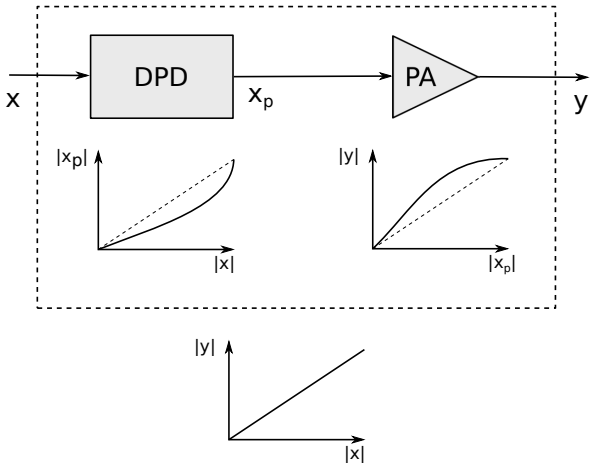


Figure 1.3.: Overview - digital pre-distortion.

digital domain and contains a nonlinear transfer function, which represents an approximation to the inverse of the PAs intrinsic nonlinear transfer function. By passing the input signal x through the DPD it is pre-distorted (x_p) in such a way that the output signal y of the PA is a linearly scaled and delayed version of the input signal x. By applying DPD to a PA, the overall transfer function of the overall system then becomes linear again. Depending on the quality of the DPD, it can significantly reduce the distortions introduced by the PA and therefore allow operation of the PA much closer to its saturation point, which increases its power efficiency.

1.2. Problem Formulation

In order to estimate the inverse of the PAs nonlinear transfer function for the DPD system, the input and the distorted output signal of the PA needs to be sampled. Depending on the order of nonlinearity shown by the PA, the bandwidth of the distorted output signal can be several times that of the input signal. As a consequence, the distorted output signal needs to be sampled at a rate several times higher than that of the input signal. A factor of five is typically found in literature for this increase in sampling rate [7,11]. Achieving these increased sampling rates does not represent a major problem for 3G or 4G signals with 5 MHz respectively 20 MHz of channel bandwith. But for 5G or 802.11ad communication systems with bandwidths of up to 2 GHz the increased sampling rate certainly is a major problem. Furthermore, with increasing bandwidth other problems such as LO leackage or IQ imbalance emerge and need to be addressed in order to successfully perform DPD on a system.
The institute of robust power semiconductor systems has founded or is taking part in several research projects for investigating the potential use of the waveguide E-band (70 to 90 GHz) frequency range for high data rate communication systems. Identifying an appropriate DPD algorithm and applying it to a wideband E-band communication link is the objective of this thesis. A linearization bandwidth of up to 5 GHz is targeted. But since the resulting requirements in terms of sampling rate and baseband processing complexity for such a high bandwidth exceed the performance of most commercially available data converters (ADC/DAC) and digital signal processing devices (FPGA/DSP), the DPD algorithm will utilize an open-loop architecture. This approach allows for an offline estimation process of the inverse transfer function of the PA in software and the use of standard measurement equipment (oscilloscope) for sampling the output signal of the PA.

2. Theory and Background

In this chapter the theory needed for understanding the operating principle of DPD is presented. First, a brief overview of wireless communications systems including common transmitter architectures, modulation schemes as well as spectral and power efficiency is given. This section is followed by a more detailed description of the distortions introduced by a PA. Here the different types of distortions are examined and figures of merit for quantification of the effects of these distortions are introduced. Finally an introduction into the Volterra theory of nonlinear systems is given, which represents the foundation of the open-loop DPD algorithm utilized in this thesis.

2.1. Wireless Communication Systems

A typical wireless communications system can be divided into several blocks as shown in Figure 2.1. The most general blocks are the transmitter and the receiver. Either both of these blocks are implemented in the same physical device, which is capable of duplex communication and is called a transceiver or they are split up into a dedicated transmitter and receiver devices. Further, the transmitter and receiver block each can be divided into a digital modem and a RF front end. Since transmitter and receiver are operating with the same basic principle but inverse signal flow direction, only the blocks contained in the transmitter will be described in more detail.

The information to be transmitted originates from a source and shall be transmitted over a RF communication channel to the sink. Typically the source outputs data containing redundancy. The redundant part of the data represents no gain in information for the sink and is removed during the source encoding process. The source encoding is typically not integrated into the digital modem but as a preprocessing step. After that the information is passed on to the channel encoding process. Here redundancy is added to the information again but in a controlled way. This redundant information is produced by applying error control codes on the information and can later be used by the receiver for error detection and correction. The modulator takes the encoded information bits and maps them to modulation symbols. Depending on the modulation scheme (e.g. BPSK, QPSK, 64-QAM, OFDM...) multiple bits of information might be contained in one modulation symbol. The modulator may also include modern channel access methods like code division multiple access (CDMA). In order to limit the frequency band occupied by the modulated signal, pulse shaping is applied. This is typically done by filtering the modulated signal with a raised-cosine or root-raised-cosine filter, which smoothens the hard edges of the modulated signal in the time domain and therefore reduces its high frequency contents. Afterwards the filtered digital signal is translated into a corresponding analog representation by a DAC and a suitable

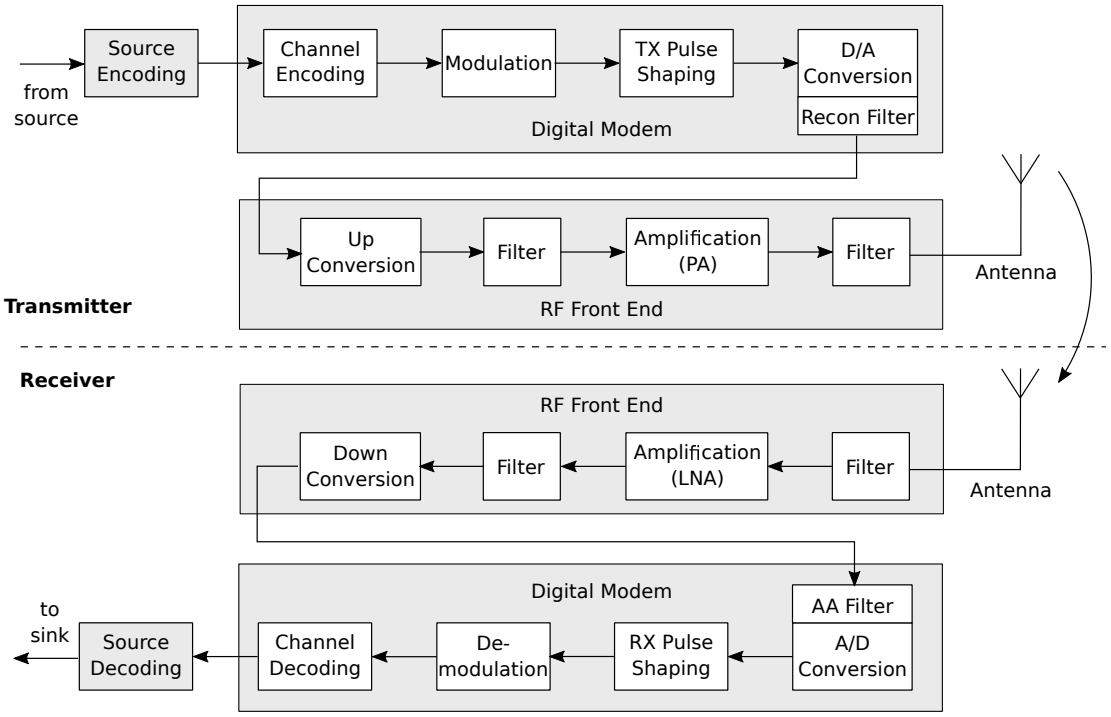


Figure 2.1.: Overview - communications system.

reconstruction filter, which filters out the unwanted spectral mirror images of the signal appearing at integer multiples of the DACs sampling frequency.
A mixer is used to shift the baseband signal to the desired RF band used for transmission. This process is called upconversion and involves a nonlinear device, which is fed with the baseband signal and a LO signal. The nonlinearity in the mixer creates new signals at the sums and differences of the frequencies of the input signals. A filter is then used to filter out the unwanted mixing products and keep only the signal at the RF band. In order to overcome the significant loses of the RF communication link (free space path loss, atmospheric absorption) and provide enough signal energy for the receiver to properly decode the signal, a PA is used. As already mentioned, the PA is most efficient when operated near its saturation point but unfortunately the nonlinearity of the PAs transfer function increases when operating near saturation. The nonlinearity must be kept as low as possible in order to utilize spectrally efficient modulation schemes. These contradicting requirements represent the biggest challenge in the PA design process. Usually a rather efficient PA with moderate nonlinear characteristics is designed and external techniques for linearization are applied to it on system level. Before radiating the RF signal using the antenna, a band limiting filter is applied to ensure that only signals falling into desired RF band are emitted. [7]

2.1.1. Transmitter Architectures

The actual implementation of the system blocks mentioned in section 2.1 may vary depending on the intended application of the transceiver and the available system resources. Two of the most common architectures, the superheterodyne and direct conversion architecture, will be presented

in the following subsections. Only the transmitters will be considered in more detail, since the corresponding receiver uses the same basic principle but with inverted signal flow direction and is not relevant in terms of DPD.

2.1.1.1. Superheterodyne Conversion Transmitter

The most common architecture for wireless transmitters is the so called superheterodyne architecture. Its main characteristic is the use of one or more intermediate frequencies in the upconversion process. A block diagram of the superheterodyne architecture is shown in Figure 2.2. The archi-

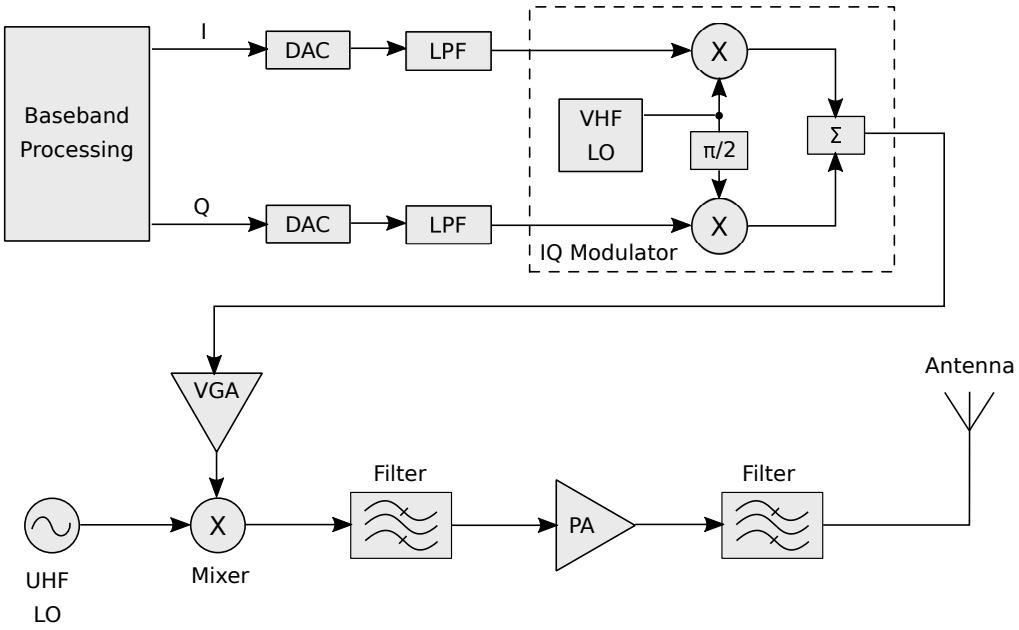


Figure 2.2.: Architecture - superheterodyne transmitter.

tecture can be divided in terms of signal frequency: the so called baseband section contains the baseband processing, DAC, reconstruction filters and the IQ modulator. All these components operate at a frequency offset of 0 Hz and are designed for processing a complex signal (respectively I and Q signals). The IQ modulator combines the complex baseband signal with the LO signal and a 90° phase shifted version of the LO signal in order to generate an IF representation of the baseband signal. In the IF section the signal is amplified by a variable gain amplifier (VGA), which is usually implemented as a multistage amplifier. In the following RF section, the IF signal is upconverted to the final RF frequency by a mixer. Unwanted spectral components generated by the nonlinearity of the mixer are suppressed by a corresponding bandpass filter. A PA is used to lift the RF signals power to the desired level. Again, a bandpass filter is used to filter out unwanted spectral components generated by the nonlinearity of the PA ensuring that only signal components inside the desired frequency band are present. Finally, the RF signal is emitted using a suitable antenna. Note that depending on the sampling rate of the DAC the boundary between digital and analog domain may also be located after the IQ modulator.
The superheterodyne architecture provides a number of advantages, which make it a popular

choice for many wireless transmitters. Most of the gain is produced in the IF section, which makes it easier to achieve high and stable gains compared to performing this operation in the RF section. Furthermore, components cost and power consumption in the IF section are typically less than in the RF section. But the most important advantage is the elimination of IQ imbalance and DC offset/LO leakage problems when the digital to analog conversion is performed after the IQ modulator. Consider a superheterodyne transmitter with the architecture depicted in Figure 2.2 and an amplitude imbalance g and phase imbalance ϕ on the Q signal path, which is in reality distributed over the complete Q signal path but will be concentrated in the IQ modulator for modeling purposes. Let

$$s_{\mathrm{BB}}(t) = I(t) + jQ(t) \tag{2.1}$$

be the complex baseband signal and

$$\begin{aligned} x_{\mathrm{LO}}(t) &= cos(\omega_{\mathrm{LO}}t) - jgsin(\omega_{\mathrm{LO}}t + \phi) \\ &= K_1 e^{-j\omega_{\mathrm{LO}}t} + K_2 e^{j\omega_{\mathrm{LO}}t} \end{aligned} \tag{2.2}$$

with

$$\begin{aligned} K_1 &= (1 + ge^{-j\phi})/2, \\ K_2 &= (1 - ge^{j\phi})/2 \end{aligned} \tag{2.3}$$

the LO signal containing the IQ imbalance [16]. Performing the up conversion process in the frequency domain

$$\begin{aligned} X_{\mathrm{IF}}(\omega) = S_{\mathrm{BB}}(\omega) * X_{\mathrm{LO}}(\omega) &= S_{\mathrm{BB}}(\omega) * (K_1 2\pi\delta(\omega + \omega_{\mathrm{LO}}) + K_2 2\pi\delta(\omega - \omega_{\mathrm{LO}})) \\ &= 2\pi(K_1 S_{\mathrm{BB}}(\omega + \omega_{\mathrm{LO}}) + K_2 S_{\mathrm{BB}}(\omega - \omega_{\mathrm{LO}})) \end{aligned} \tag{2.4}$$

reveals that an unwanted image of the baseband signal spectrum appears at ω_{LO} with complex amplitude K_2. Therefore $X_{\mathrm{IF}}(\omega)$ is not the spectrum of a purely analytical signal with only positive

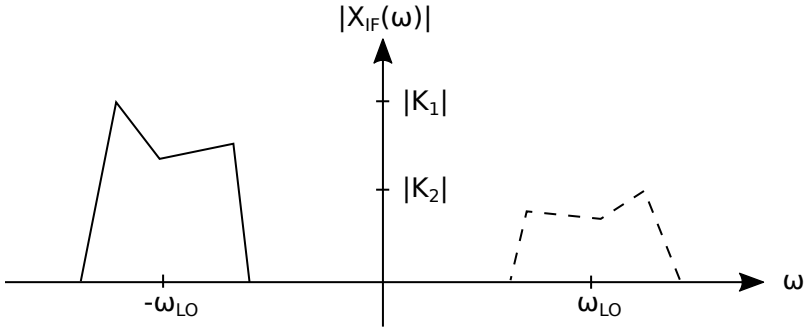


Figure 2.3.: Amplitude spectrum of $X_{\mathrm{IF}}(\omega)$ with IQ imbalance.

or negative frequency components. The amplitude of $X_{\mathrm{IF}}(\omega)$ is shown in Figure 2.3. The spectral image at ω_{LO} will cause distortions, since it is not filtered out or compensated when the real-valued

IF signal

$$\begin{aligned} x_{\mathrm{IF,Re}}(t) &= Re\{\mathcal{F}^{-1}\{X_{\mathrm{IF}}(\omega)\}\} \\ &= I(t)cos(\omega_{\mathrm{LO}}t) + Q(t)gsin(\omega_{\mathrm{LO}}t + \phi) \end{aligned} \tag{2.5}$$

is outputted by the mixer. When the IQ modulation is performed in the digital domain, perfect matching ($g = 1$, $\phi = 0$) can be achieved. In that case becomes $K_1 = 2$, $K_2 = 0$ and only the desired signal spectrum located at $-\omega_{\mathrm{LO}}$ is present in $X_{\mathrm{IF}}(\omega)$ resulting in an undistorted IF signal $x_{\mathrm{IF}}(t)$ at the output of the mixer. Furthermore, digital IQ modulation does not suffer from a finite isolation between the VHF LO and output signal ports of the mixers used for IQ modulation. As a result, there is no LO leakage into the IF signal band. Only during the upconversion process does the UHF LO signal leak into the output signal of the mixer. But since the leaked LO signal is offset by the IF frequency from the RF signal band, it is possible to filter out this unwanted spurious signal. However, a solid frequency planning has to be performed in order to ensure that the out of band emissions are not interfering with adjacent RF channels. [9]

2.1.1.2. Direct Conversion Transmitter

The direct conversion transmitter can be seen as a special case of the superheterodyne transmitter, which is operating at an IF of 0 Hz. For that reason the direct conversion architecture is also know as zero IF architecture. The block diagram is shown in Figure 2.4. The direct conversion architecture

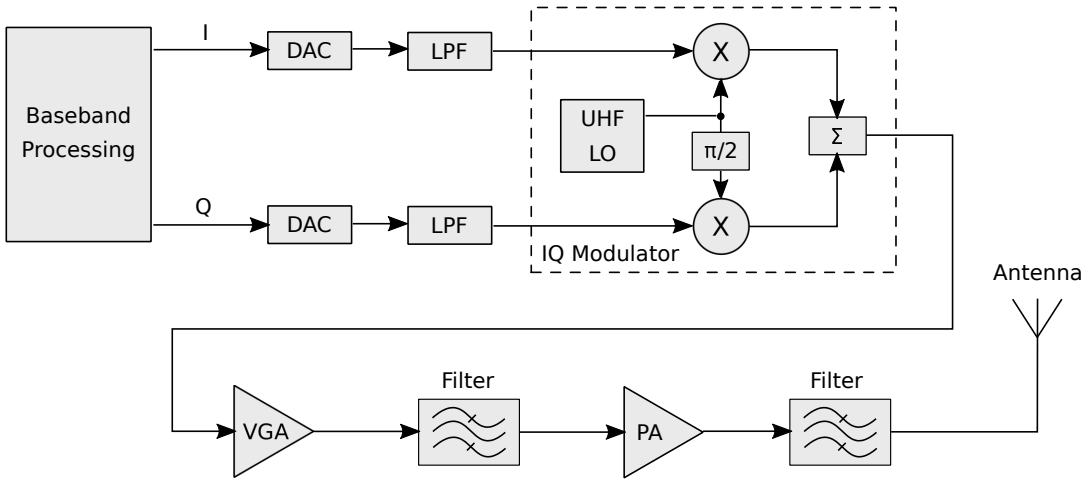


Figure 2.4.: Architecture - direct conversion transmitter.

usually contains less components than the superheterodyne architecture since there is no IF section. The complex baseband signal is generated by a pair of DACs and corresponding reconstruction filters. The baseband signal is then directly upconverted to the desired RF band using a UHF LO signal. The output power level is controlled by a VGA, which ensures that the PA is not overdriven. Unwanted spectral components, which may have emerged during the amplification processes are filtered out by bandpass filters. Finally, the RF signal is amplified by a PA, filtered again and

radiated by the antenna.
Compared to the superheterodyne architecture (with digital-analog boundary after the IQ modulator) the achievable baseband bandwidth can be higher, since the upconversion process is not performed in the digital domain and therefore the complete DAC bandwidth can be used for baseband signal generation. Alternatively the sampling rate of the digital baseband components can be lower, when no increased bandwidth is necessary. Due to the nonexistent IF section there are less spurious signals present, which makes frequency planing less complex and also less components are required resulting in less cost and less required space. However, the direct conversion architecture contains technical challenges, which make it a less popular choice in RF transceiver design. Since the complex baseband signal is directly upconverted to the desired RF frequency band and this process can not be performed in the digital domain, it is not possible to avoid the IQ imbalance issues as described for the superheterodyne architecture. Further, LO leakage represents a problem especially for large bandwidth transmissions. This is due to the fact that the leaked LO signal is located in the middle of the RF frequency band and cannot be filtered out without also removing parts of the desired RF signal. Therefore the leaked LO signal may show up as a strong interferer as depicted in Figure 2.5. LO leakage results from a finite isolation between the LO and RF signal

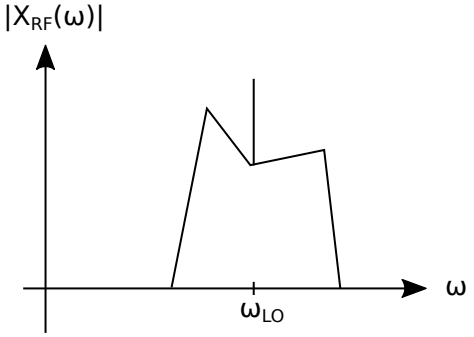


Figure 2.5.: Amplitude spectrum of $X_{\mathrm{RF}}(\omega)$ with LO leakage.

ports of the IQ modulator. But it can also result from DC offsets on the complex baseband signal. Consider the complex baseband signal

$$s_{\mathrm{BB}}(t) = s(t) + s_{\mathrm{off}} \tag{2.6}$$

consisting of the actual data carrying signal $s(t)$ and the complex DC offset s_{off}. The upconversion process performed in the frequency domain

$$\begin{aligned} X_{\mathrm{RF,Re}}(\omega) &= Re\{\mathcal{F}\{(s(t) + s_{\mathrm{off}})e^{j\omega_{\mathrm{LO}}t}\}\} \\ &= Re\{\mathcal{F}\{s(t)e^{j\omega_{\mathrm{LO}}t} + s_{off}e^{j\omega_{\mathrm{LO}}t}\}\} \\ &= Re\{S(\omega - \omega_{\mathrm{LO}}) + s_{\mathrm{off}}2\pi\delta(\omega - \omega_{\mathrm{LO}})\} \end{aligned} \tag{2.7}$$

and shows that the DC component is shifted to the center of the RF frequency band. This is problematic, since there is no way to filter out the leaked LO signal without also removing parts

of the data carrying signal. As a result, extremely narrow band filters are required in order to minimize the impact on the data carrying signal. When the leaked LO signal is not filtered out, it can saturate the receiver circuits and degrade its sensitivity. [9]

2.1.2. Modulation Schemes

In order to efficiently transmit information over a communication channel, it is required to encode the information onto a baseband signal. The information may be encoded into the amplitude $R(t)$ and/or the phase $\Phi(t)$ of the baseband signal. The defining characteristic of a baseband signal is the presence of spectral magnitude in the vicinity of $f = 0$ Hz and zero magnitude elsewhere. Depending on the properties of the communication channel, it is also necessary to convert the baseband signal into a bandpass signal by shifting the baseband signal spectrum to a carrier frequency $f = \pm f_c$. The combined process of information encoding and frequency shifting is called modulation. For any physical bandpass signal exists a complex envelope representation

$$\begin{aligned} x(t) &= Re\{s_{BB}(t)e^{j\omega_c t}\} \\ &= R(t)cos(\omega_c t + \Phi(t)) \\ &= I(t)cos(\omega_c t) - Q(t)sin(\omega_c t) \end{aligned} \tag{2.8}$$

where

$$s_{BB}(t) = I(t) + jQ(t) = R(t)e^{j\Phi(t)} \tag{2.9}$$

denotes the complex envelope of $x(t)$, $\omega_c = 2\pi f_c$ represents the associated carrier frequency and $I(t)$, $Q(t)$ represent the in-phase and quadrature modulation signals. The complex envelope $s_{BB}(t)$ is also called the baseband signal and results from applying a mapping

$$s_{BB}(t) = g[m(t)]. \tag{2.10}$$

onto the modulating signal $m(t)$, which represents the information to be encoded. With the cartesian form respectively the quadrature modulation ($I(t)$, $Q(t)$ signals) it is possible to represent all other modulation types [2]. A selection of modulation types and their corresponding quadrature modulation is given in Table 2.1.

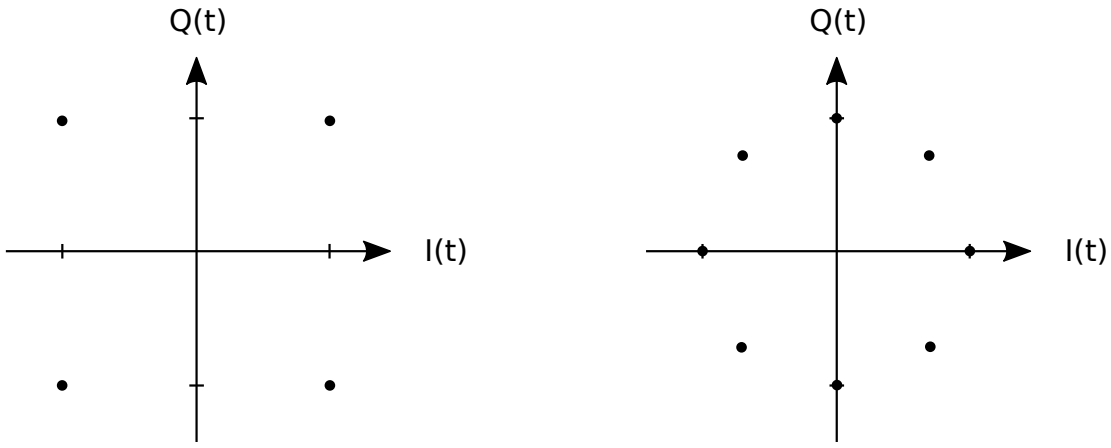
Quadrature modulation (QM) is a common modulation type in modern communication systems, since it is easy to implement in the digital baseband domain and provides flexibility in terms of adopting the modulation format to the characteristics of the channel (i.e. signal-to-noise ratio (SNR)). QM can be used to implement several different modulation formats.

Type of Modulation	**Mapping Function** $g(m)$	$I(t)$	$Q(t)$
AM	$A_c[1+m(t)]$	$A_c[1+m(t)]$	0
DSB-SC	$A_c m(t)$	$A_c m(t)$	0
PM	$A_c e^{jD_p m(t)}$	$A_c cos[D_p m(t)]$	$A_c sin[D_p m(t)]$
FM	$A_c e^{jD_f \int_{-\infty}^{t} m(\sigma)\,d\sigma}$	$A_c cos[D_f \int_{-\infty}^{t} m(\sigma)\,d\sigma]$	$A_c sin[D_f \int_{-\infty}^{t} m(\sigma)\,d\sigma]$
QM	$A_c[m_1(t)+jm_2(t)]$	$A_c m_1(t)$	$A_c m_2(t)$

Table 2.1.: Complex envelope functions for various types of modulation. [2]

M-PSK

One such modulation format is called M-ary Phase-Shift Keying (M-PSK). As the name implies, it utilizes M different phase states of the complex baseband signal in order to encode the information. In its most basic form M is equal to 4 and that particular modulation format is called quadrature phase-shift keying (QPSK). The constellation diagram for QPSK is shown in Figure 2.6a. For higher orders of M more phase states are added equally spaced onto a circle with constant radius in the constellation diagram. In Figure 2.6b the constellation diagram for 8-PSK is shown. The baseband



(a) QPSK constellation diagram. **(b)** 8-PSK constellation diagram.

Figure 2.6.: M-PSK constellation diagrams.

signal $s_{BB}(t)$ can be constructed using QM where

$$I(t) = \sum_{n} A_c cos(\phi_n) h(t - nT_s) \tag{2.11}$$

and

$$Q(t) = \sum_{n} A_c sin(\phi_n) h(t - nT_s) \tag{2.12}$$

represent the modulating functions for mapping the M discrete phase states ϕ_i, which are assigned to $log_2(M)$ information bits, onto the baseband signal. $h(t - nT_s)$ represents the pulse shaping function, which is shifted by n times the symbol duration T_s. Note that the amplitude of the baseband signal is always A_c, which means that M-PSK is a constant envelope modulation format. However, this is only true for rectangular-shaped data pulses. [2]

M-QAM

M-ary quadrature amplitude modulation (M-QAM) uses a similar approach as M-PSK but with the difference that it is not restricted to using constellation points located on a circle with constant amplitude. Instead, the information is encoded in the amplitude and phase of the baseband signal. The most basic form for M equal 4 is exactly the same as QPSK and has the same constellation diagram as shown in Figure 2.6a. The constellation diagram for 16-QAM is shown in Figure 2.7. For higher order M-QAM formats the constellation diagram contains more equally spaced points. The construction of the M-QAM baseband signal can be performed using QM with the modulating

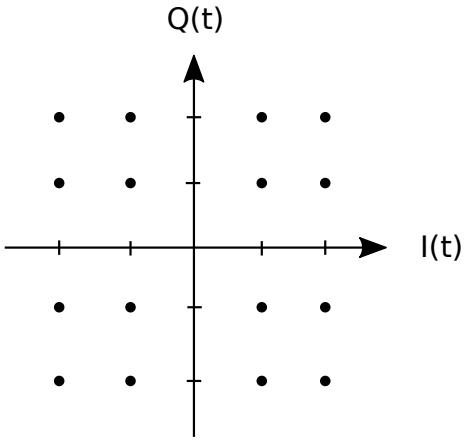


Figure 2.7.: 16-QAM constellation diagram.

functions

$$I(t) = \sum_n x_i h(t - nT_s) \tag{2.13}$$

and

$$Q(t) = \sum_n y_i h(t - nT_s) \tag{2.14}$$

where (x_i, y_i) represents the allowed constellation points associated with $log_2(M)$ information bits and $h(t - nT_s)$ is the pulse shaping function, which is shifted by n times the symbol duration T_s. [2]

2.1.3. Pulse Shaping

Since wireless communication systems are sharing the frequency spectrum with other users, the division of the frequency spectrum into dedicated subchannels is necessary in order to avoid interference between users. As a result, only bandlimited signals are permitted by regulatory authorities for use on each of the subchannels. Rectangular shaped pulses have an absolute bandwidth of infinity and are therefore not suitable for use on bandlimited channels. However, improper filtering of the rectangular shaped pulses can lead to intersymbol interference (ISI) as depicted in Figure 2.8. In order to prevent ISI from happening, transmitting $H_T(f)$ and receiver

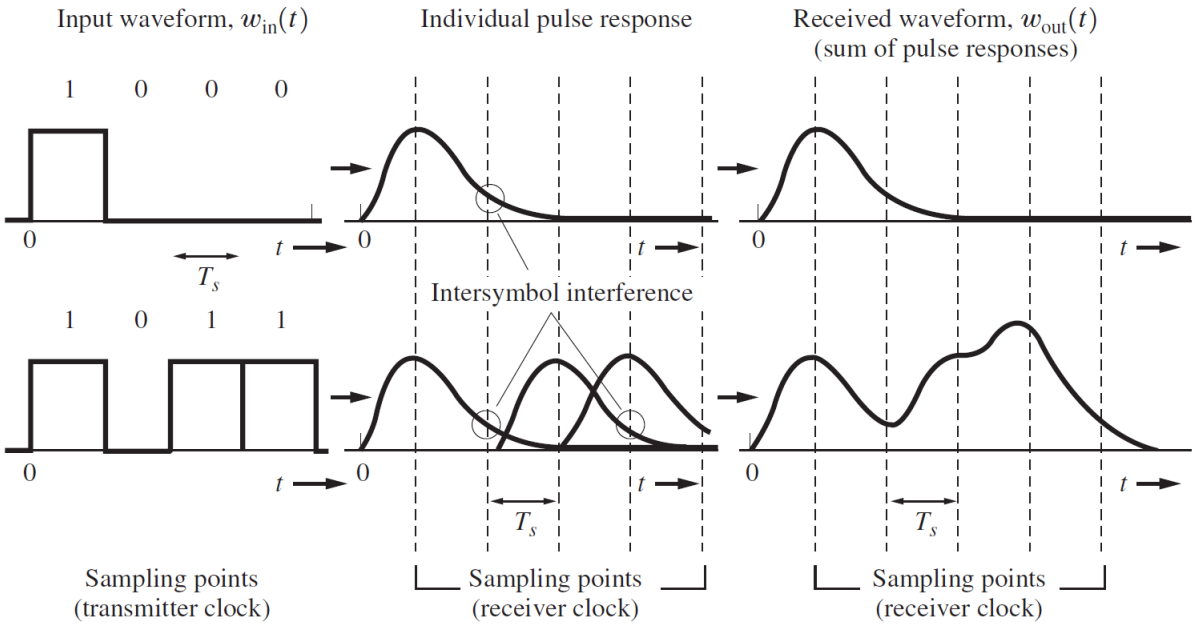


Figure 2.8.: ISI for not properly filtered rectangular pulses. [2]

filters $H_R(f)$ have to be added into the communications system as depicted in Figure 2.9. The

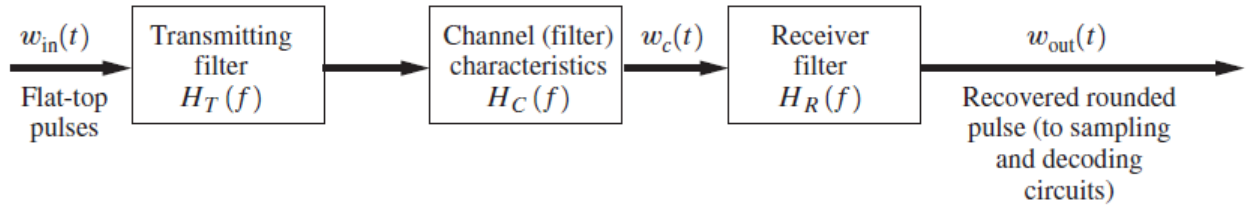


Figure 2.9.: Baseband representation of communications system (simplified). [2]

transmitter signal is

$$w_{in}(t) = \left[\sum_n a_n \delta(t - nT_s)\right] * h(t), \tag{2.15}$$

where a_n represents any of the allowed symbol amplitudes, T_s the symbol duration and $h(t)$ the rectangular pulse shape function. Using the equivalent impulse response of the communications system model in Figure 2.9

$$h_e(t) = h(t) * h_T(t) * h_C(t) * h_R(t) \tag{2.16}$$

the corresponding output signal at the receiver can be written as

$$w_{out}(t) = \left[\sum_n a_n \delta(t - nT_s)\right] * h_e(t) \tag{2.17}$$

Note that $h_e(t)$ represents the output signal waveform at the receiver for rectangular shaped transmit pulses. The channel characteristic $h_C(t)$ is usually fixed and only compensated for when the transmitting and receive filters are implemented as adaptive filters. For a non-adaptive approach the channel characteristics are ignored ($h_C(t) = \delta(t)$). But the transmitting $h_T(t)$ and and receiver filters $h_R(t)$ can be chosen in such a way that the ISI is minimized. This can be achieved by constraining the equivalent impulse response to fulfill the following constraint

$$h_e(kT_S + \tau) = \begin{cases} C, & k = 0 \\ 0, & k \neq 0 \end{cases} \tag{2.18}$$

with the discrete sampling time index k, a non-zero constant C and the receiver clock offset τ. An obvious choice fulfilling this constraint would be the $sin(x)/x$ function and with a clock offset $\tau = 0$

$$h_e(t) = \frac{sin(\pi f_s t)}{\pi f_s t} \tag{2.19}$$

where $f_s = 1/T_s$ denotes the sampling frequency required in the receiver. However, this function puts very strong requirements on the receiver clock synchronization and is therefore rarely used in practice. A more commonly used equivalent impulse response is the so called raised cosine-rolloff Nyquist filter with the transfer function

$$H_e(f) = \begin{cases} 1, & |f| < f_1 \\ \frac{1}{2}\left(1 + cos\left[\frac{\pi(|f| - f_1)}{2f_\Delta}\right]\right), & f_1 < |f| < B \\ 0, & |f| > B \end{cases} \tag{2.20}$$

and the corresponding impulse response

$$h_e(t) = 2f_0\left(\frac{sin(2\pi f_0 t)}{2\pi f_0 t}\right)\left[\frac{cos(2\pi f_\Delta t)}{1 - (4f_\Delta t)^2}\right] \tag{2.21}$$

where B is the absolute bandwidth, f_0 is the 6-dB bandwidth of the filter and the parameters

$$f_\Delta = B - f_0 \tag{2.22}$$

and

$$f_1 = f_0 - f_\Delta. \tag{2.23}$$

The rolloff factor is defined as

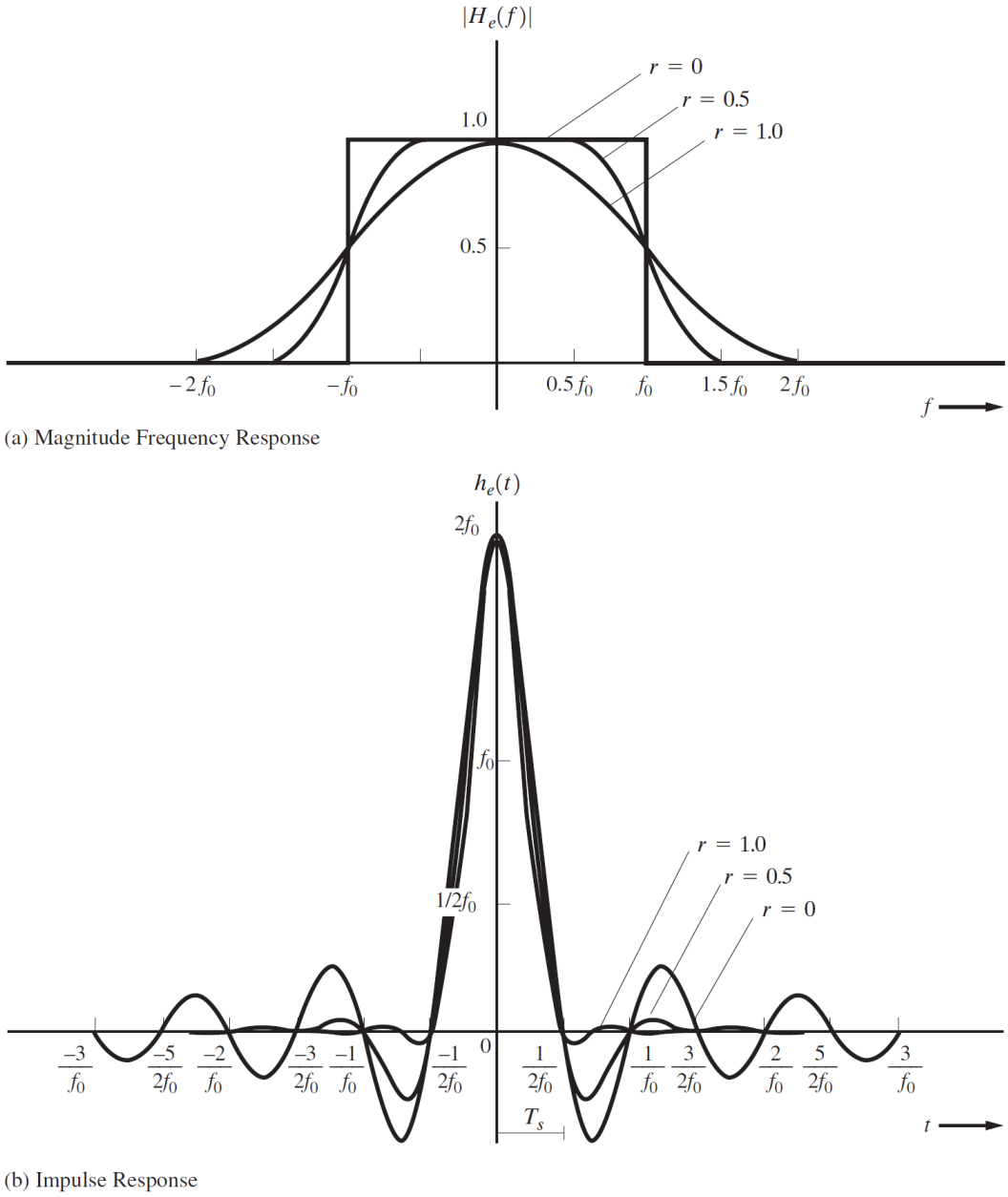


Figure 2.10.: Raised cosine-rolloff Nyquist filter characteristics. [2]

$$r = \frac{f_\Delta}{f_0} \tag{2.24}$$

and basically controls the steepness of the slopes of the transfer function and also the amount of decay of the impulse response. This becomes obvious when the graphical representations in Figure 2.10 are considered. Note that the equivalent impulse response can be split up onto the transmitting $h_T(t)$ and receiver filters $h_R(t)$. The split filters are called root raised cosine filters and are commonly used in practice. [2]

2.1.4. Spectral Efficiency

Due to the band limitations imposed on the channel of wireless communication systems by regulatory authorities, it becomes important to use the available bandwidth efficiently. This ensures that a maximum amount of information can be transferred in a given time interval. Since M-QAM is one of the most popular modulation formats, only its spectral efficiency will be covered. The absolute bandwidth of a M-QAM baseband signal with raised cosine filtering is

$$B = \frac{1}{2}(1+r)D \tag{2.25}$$

where $D = R/l$, R is the data rate, $l = log_2(M)$ is the number of bits per symbol and r is the rolloff factor of the raised cosine filter. For the actual transmitted bandpass signal the bandwidth is $B_T = 2B$, which leads to a spectral efficiency of

$$\eta = \frac{R}{B_T} = \frac{log_2(M)}{1+r} \tag{2.26}$$

The spectral efficiency of a selection of M-QAM signals with different rolloff factors is given in Table 2.2. Note that the order M of the QAM signal cannot be increased infinitely. For the same

	Spectral efficiency $\eta\left[\frac{bit/s}{Hz}\right]$			
QAM order M	$r = 0.00$	$r = 0.35$	$r = 0.50$	$r = 1.0$
2	1.00	0.74	0.67	0.50
4	2.00	1.48	1.33	1.00
8	3.00	2.22	2.00	1.50
16	4.00	2.96	2.67	2.00
32	5.00	3.70	3.33	2.50

Table 2.2.: Spectral efficiency for M-QAM signals with raised cosine filtering. [2]

output power, higher order modulation formats place constellation points closer to each other. With the presence of noise on the channel, this will lead to errors during the demodulation in the

receiver, because the noise will cause some received symbols to fall onto neighboring constellation points. For an AWGN channel, the maximum achievable spectral efficiency is determined by the channel capacity C as

$$\eta_{\max} = \frac{C}{B} = log_2\left(1 + \frac{S}{N}\right) \tag{2.27}$$

where S/N denotes the SNR. This leads to the upper bound for the spectral efficiency of the M-QAM signal on an AWGN channel of [2]

$$\eta < \eta_{\max}. \tag{2.28}$$

2.2. RF Power Amplifier Characteristics

2.2.1. RF Power Amplifier Efficiency

Depending on the area coverage of a communications system, the RF PA can make up a major part of the overall power consumption (see Figure 1.2). Therefore it is important to design the RF PA for high power efficiency. Among other techniques, this can be achieved by setting the quiescent bias point of the transistor properly. Figure 2.11 and Figure 2.12 shows the simplified schematic and waveforms for a field-effect transistor (FET) based RF PA, which is biased for class AB operation.

In the waveform for the gate voltage it can be seen that the quiescent bias voltage V_q is located

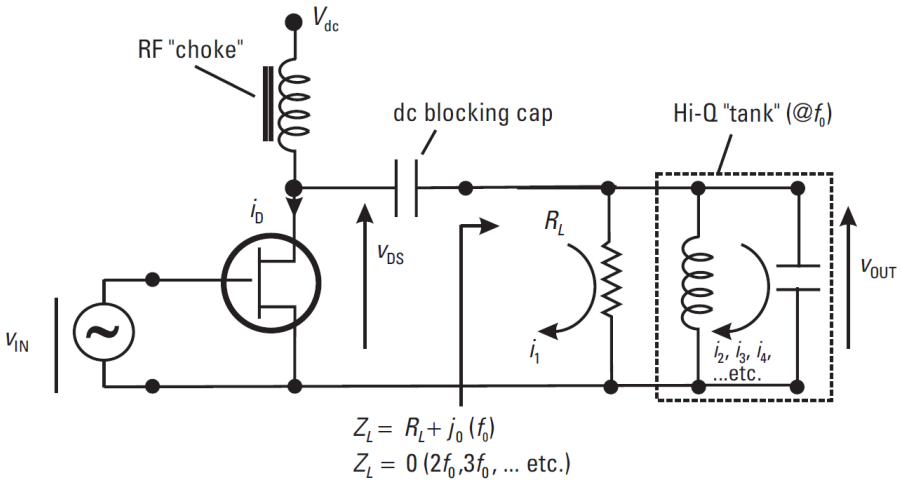


Figure 2.11.: RF PA (class AB) - schematic. [3]

slightly above the threshold voltage V_t at which the FET becomes conductive. In consequence the superimposed input signal leads to the gate voltage dropping below the threshold voltage for a certain time and ultimately clipping in the drain current. The clipping produces odd order

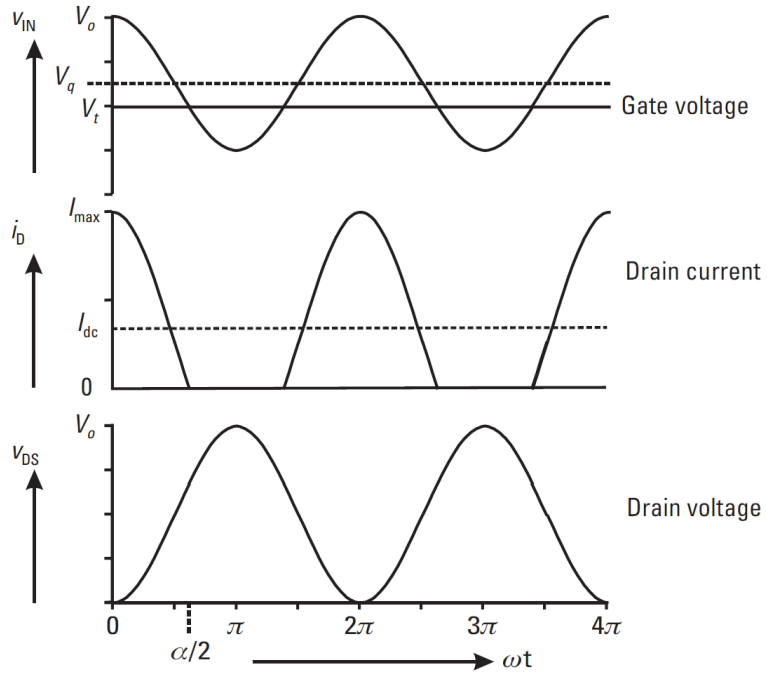


Figure 2.12.: RF PA (class AB) - waveforms. [3]

harmonic distortions at the output of the FET. However, the key part of the RF PA in class AB operation is the shunt resonator at the output. It has a resonant frequency at the fundamental of the input signal f_0 and shorts any higher order harmonic distortions. This way the output signal waveform approximates the sine wave at the input with an amplitude, which is dependent on the drive level and the load resistor R_L.

RF PA operation is classified depending on the proportion of the conduction phase in one period of the input sine wave. This proportion is measured in degree or radians. A class A biased RF PA is conductive for the full period of 2π, wheres a class B biased RF PA only conducts for half the period (π). The class AB RF PA bias point is located inbetween the bias points of class A and class B. The efficiency and also the RF output power of a RF PA heavily depends on the conduction angle as can be seen in Figure 2.13.

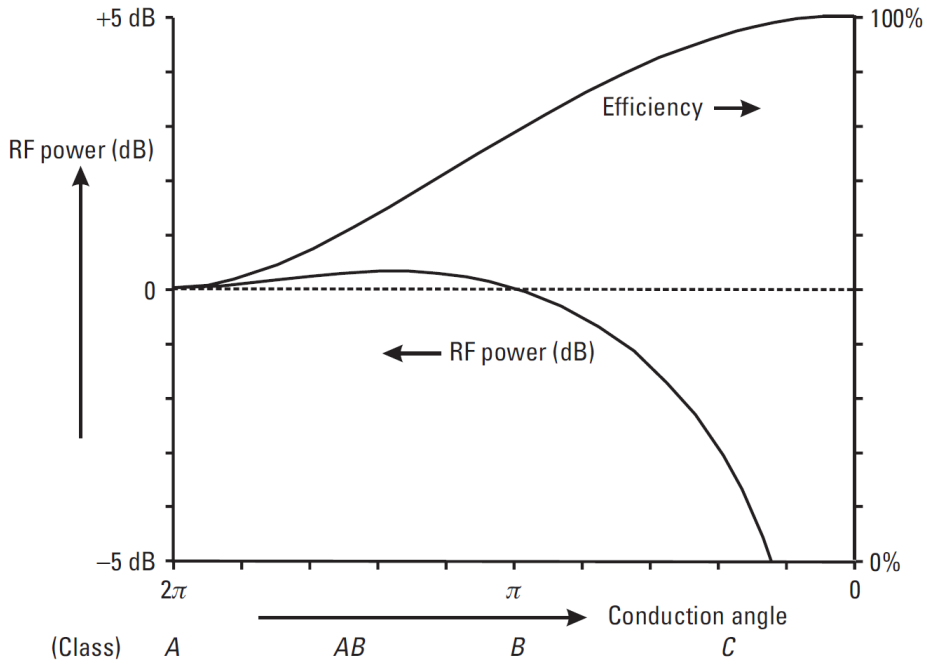


Figure 2.13.: RF PA classes and corresponding efficiency and RF power. [3]

Note that this plot is only valid for maximum input drive level. But the advantages in terms of efficiency for operation with smaller conduction angles is also true for lower input drive levels as depicted in Figure 2.14. This is an important fact, since depending on the modulation format

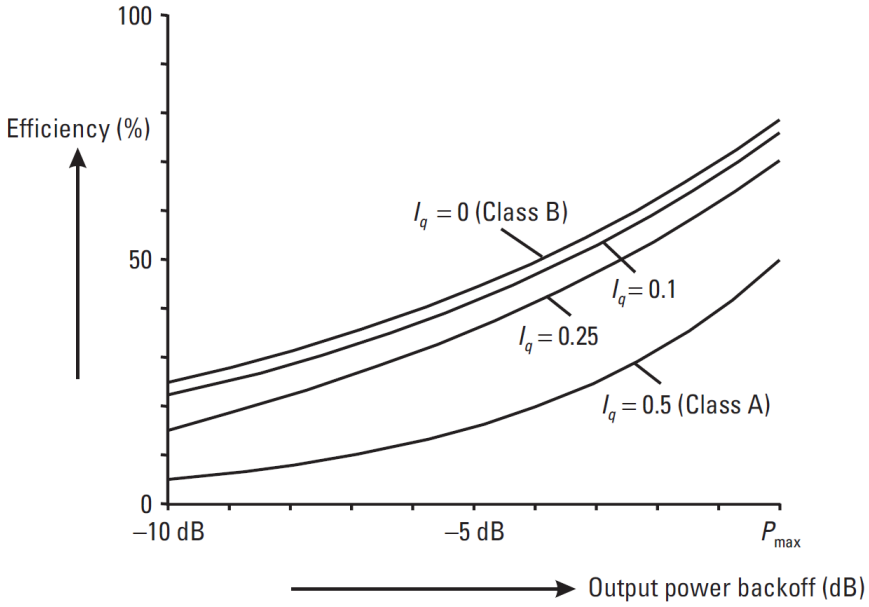


Figure 2.14.: RF PA efficiency vs power backoff. [3]

used it might not be possible to operate the RF PA near its saturation point. In contrast to constant

envelope modulation formats, such as QPSK, higher order modulation formats like 16-QAM, which also encode information in the signal amplitude, have stronger requirements on linearity of the RF PAs transfer characteristic. To satisfy these requirements, it is necessary to reduce the input drive level respectively output power so that the RF PA mainly operates in its linear region. The process of reducing the input drive level respectively output power is called power backoff and represents a simple and inefficient form of linearization for the RF PA. [3] served as a productive source for this subsection.

2.2.2. Intermodulation Distortion

A RF PA, which is operated near its saturation point, produces distortions in the output signal. These distortions can be classified into frequency dependent and frequency independent parts. The frequency independent parts of these distortions are called intermodulation (IM) and harmonic (HM) distortions. The best way to illustrate the mechanism behind IM and HM distortions is to stimulate a memoryless nonlinear system model with a two-tone signal. The nonlinear system considered in this example is a polynomial of order $P = 3$ and is defined by

$$y(t) = ax(t) + bx^2(t) + cx^3(t), \tag{2.29}$$

where $y(t)$ denotes the output signal, $x(t)$ the input signal and a, b, c the coefficients of the linear, quadratic and cubic terms in the model. The two-tone signal is defined as

$$x(t) = A_1 cos(\omega_1 t) + A_2 cos(\omega_2 t) \tag{2.30}$$

with the magnitudes A_1 and A_2, and the angular frequencies ω_1 and ω_2 ($\omega_2 > \omega_1$). After the two-tone signal has passed through the nonlinear model the following output signal emerges

$$\begin{aligned}
y(t) =& a[A_1 cos(\omega_1 t) + A_2 cos(\omega_2 t)] \\
&+ \left[\left(\frac{3}{4}cA_1^2 + \frac{3}{4}cA_2^2\right)A_1 cos(\omega_1 t)\right] + \left[\left(\frac{3}{4}cA_2^2 + \frac{3}{4}cA_1^2\right)A_2 cos(\omega_2 t)\right] \\
&+ \left[\frac{3}{4}cA_1^2 A_2 cos((2\omega_1 - \omega_2)t)\right] + \left[\frac{3}{4}cA_1 A_2^2 cos((2\omega_2 - \omega_1)t)\right] \\
&+ \left[\frac{1}{2}bA_1^2 + \frac{1}{2}bA_2^2\right] + [bA_1 A_2 cos((\omega_2 - \omega_1)t)] \\
&+ \frac{1}{2}b\left[A_1^2 cos(2\omega_1 t) + A_2^2 cos(2\omega_2 t)\right] + [bA_1 A_2 cos((\omega_2 + \omega_1)t)] \\
&+ \frac{1}{4}c\left[A_1^3 cos(3\omega_1 t) + A_2^3 cos(3\omega_2 t)\right] \\
&+ \frac{3}{4}cA_1 A_2 [A_1 cos((2\omega_1 + \omega_2)t) + A_2 cos((2\omega_2 + \omega_1)t)].
\end{aligned} \tag{2.31}$$

The first line of this rather complex expression represents the desired linearly scaled version of the two-tone signal whereas the remaining lines contain all the HM and IM distortions. In the second line the fundamental distortions are listed, which are located at the same frequencies as

the input signal (ω_1, ω_2). The third line contains the in-band third order IM distortions, which are located at frequencies close to the band of interest ($2\omega_1 - \omega_2$, $2\omega_2 - \omega_1$). Line four contains the lower second order IM distortions ($\omega_1 - \omega_2$, $\omega_2 - \omega_1$) and DC components. Line five contains second order HM distortions ($2\omega_1$, $2\omega_2$) and the upper second order IM distortions ($\omega_1 + \omega_2$, $\omega_2 + \omega_1$). The last two lines include the third order HM ($3\omega_1$, $3\omega_2$) and out-band third order IM distortions ($2\omega_1 + \omega_2$, $2\omega_2 + \omega_1$). The distortions, which are located at frequencies far away from the frequency band of interest can be easily filtered out before the signal is forwarded onto the channel but the in-band third order IM and fundamental HM distortions cannot be filtered out, because they are located directly in or too close to the signal band. They appear as spectral regrowth around the signal band and deformation in the spectrum of the signal itself. Figure 2.15 shows the amplitude spectrum of the distorted output signal $y(t)$ generated by the third order nonlinear system.

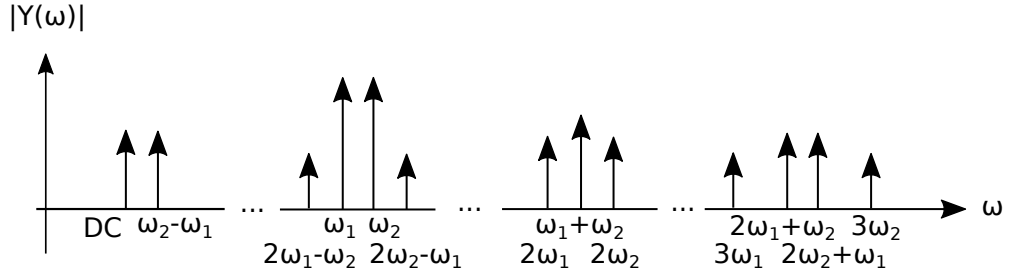


Figure 2.15.: Spectrum including intermodulation and harmonic distortions caused by memoryless third order nonlinear system. [7]

The HM/IM distortion mechanism presented in this example can be generalized to Nth order nonlinear systems, which produce mixing products upto Nth order resulting in much more complex analytic expressions for the distorted output signals and denser packed frequency spectra. [7]

2.2.3. Memory Effects

Besides the static IM and HM distortions RF PAs also suffer from frequency-dependent distortions, which are called memory effects. Memory effects are caused by energy storing circuit or chip elements and can be classified by the type of energy (electrical or thermal), which is stored in the corresponding element. Although memory effects may not necessarily cause significant deterioration of the RF PAs linearity, they can have a negative impact on the performance of linearization techniques and therefore complicate the trade-off between RF PA efficiency and linearity. [12]

2.2.3.1. Electrical Memory Effects

The main sources for electrical memory effects are the transistors terminations, which include internal and external parasitic elements as well as matching networks. In order to understand this type of memory effects, it is therefore important to understand the transistor terminations. A schematic for a metal-semiconductor field-effect transistor (MeSFET) based amplifier in common source configuration is shown in Figure 2.16.

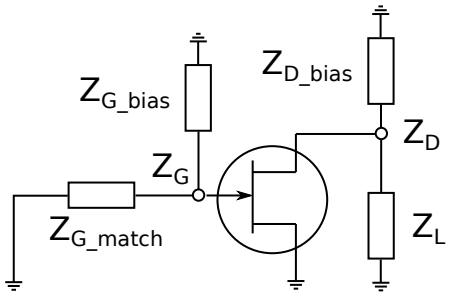


Figure 2.16.: MeSFET with matching and bias networks [7]

It includes the transistor itself, the gate and drain biasing networks represented by the impedances $Z_{\text{G\_bias}}$ and $Z_{\text{D\_bias}}$, the input matching network $Z_{\text{G\_match}}$ and the load impedance Z_{L}. The gate node impedance is given by

$$Z_{\text{G}} = Z_{\text{G\_bias}} \parallel Z_{\text{G\_match}} \parallel Z_{\text{G\_in}} \tag{2.32}$$

and the drain node impedance by

$$Z_{\text{D}} = Z_{\text{D\_bias}} \parallel Z_{\text{L}} \parallel Z_{\text{D\_in}}, \tag{2.33}$$

where $Z_{\text{G\_in}}$ and $Z_{\text{D\_in}}$ are the input impedances looking into the gate respectively the drain of the MeSFET. According to [12] the nonlinear circuit elements can be modeled using nonlinear current sources, which can influence the voltage waveform of the transistors output signal over the node impedances. A simplified model for the transistors nonlinear transfer characteristic consists of a cascade of two nonlinear systems H and F each of order three consisting of three Volterra kernels (H_1, H_2, H_3 and F_1, F_2, F_3), where the system H represents the gate voltage V_G as a function of the transistor input voltage V_{in} and F the drain voltage V_D as a function of the base voltage. For more details on the Volterra series representation for nonlinear systems see section 2.3. A block diagram of the nonlinear model is shown in Figure 2.17.

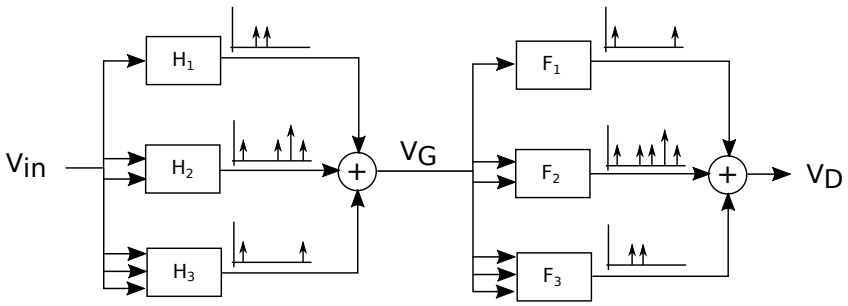


Figure 2.17.: MeSFET model consisting cascade of two nonlinear models. [12]

The third order in-band IM distortions at the output of the model (V_{D}) are generated by multiple signals passing through different Volterra kernels of the two cascaded nonlinear systems. The first system H generates a multi-tone signal with the components shown in Figure 2.15, which are the

input for the second nonlinear system F. Keep in mind, that this mixing process is performed for real-valued input signals, which also include negative frequencies: The third order IM distortions of the first system generated by kernel H_3 pass through the kernel F_1, which linearly amplifies this signal at the fundamental frequency. The linearly amplified signal from H_1 passes through the cubic kernel F_3 and also adds to the in-band IM distortions at the output. The envelope signal generated by H_2 mixes in the quadratic kernel F_2 with the fundamental signal coming from H_1. Finally the second harmonic signal from H_2 mixes in F_2 with the fundamental signal from H_1. As a result the third order IMs at the output of the model includes signals, which passed through different order nonlinearities including even orders at different frequencies. As a consequence electrical memory effects are caused by frequency-dependent gate and drain node impedances at envelope (around DC), fundamental and different order harmonic frequencies. The distortions around the envelope frequency have the same bandwidth as the input signal, whereas the bandwidth of interest is three times that of the input signal for the third order IM distortions at the fundamental frequency and two times around the second HM frequency. The node impedances of the RF PA should be kept constant over these frequency bands in order to minimize the memory effects. For the fundamental and harmonic frequency bands it is usually no problem to control the node impedances for narrow band single-carrier signals, because the modulation frequency range is small compared to the center frequency of the bands of interest. However, for broadband or multi-carrier signal keeping the node impedances constant is more complicated. An interesting fact, is that the envelope frequency band is always located around DC and does not change with the center operating frequency of the input signal. This is not the case for the fundamental and harmonic frequency bands. So if the memory effects change significantly when the center frequency of the input signal is shifted, this means that the memory effects are caused mainly by frequency-dependent node impedances in the fundamental and harmonic frequency bands. Further details on electrical memory effects can be found in[7, 12]. The information presented in this subsection has been extracted and summarized from the above mentioned references.

2.2.3.2. Electrothermal Memory Effects

A transistor is not only effected by electrical memory effects but also by electrothermal memory effects, which are caused by chip elements like the silicon or lead frame that can store thermal energy and influence electrical parameters of the transistor. The dissipated power in the MeSFET is given by

$$P_{\mathrm{DISS}} = v_{\mathrm{DS}}(t) i_{\mathrm{DS}}(t), \tag{2.34}$$

where $v_{\mathrm{DS}}(t)$ is the drain-source voltage and $i_{\mathrm{DS}}(t)$ is the drain-source current. Since the dissipated power is the product of two signals at the fundamental frequency, the resulting spectrum contains DC, envelope, sum and harmonic frequencies. The temperature variations of the chip components are proportionally related to the heat flow by the thermal impedance Z_{TH}. This thermal impedance is not purely resistive because the chip components have a thermal mass, which is capable of

storing energy. As a result the temperature of the components does not change instantly but with a frequency-dependent phase-shift. This effect creates a distributed lowpass filter, which means that usually only low frequency components of the dissipated power signal like the DC and envelope components are significant for electrothermal memory effects. The bandwidth of these thermal effects can in the range of 100 kHz upto 1 MHz. As consequence the temperature of the chip can be expressed as

$$T = T_{\mathrm{AMB}} + R_{\mathrm{TH}} P_{\mathrm{DISS}}(0\ Hz) + Z_{\mathrm{TH}}(\omega_1 - \omega_2) P_{\mathrm{DISS}}(\omega_1 - \omega_2) \tag{2.35}$$

with the ambient temperature T_{AMB}. It can be seen that the electrothermal memory also depend on the signals bandwidth $(\omega_1 - \omega_2)$ as the temperature in addition to the product of the DC thermal resistance and dissipated power also contains the same product but for the envelope signal. The heat in a chip flows mostly vertically so that self-heating of the chip is the dominant factor for this kind of memory effects and not surrounding heat sources. [12]

2.2.4. Figures of Merit

In order to quantify the distortions caused by the nonlinear transfer characteristic of the RF PA, several different metrics can be used. These metrics consider certain signal properties in the time or frequency domain of the distorted output signal.

2.2.4.1. AM/AM and AM/PM Characteristics

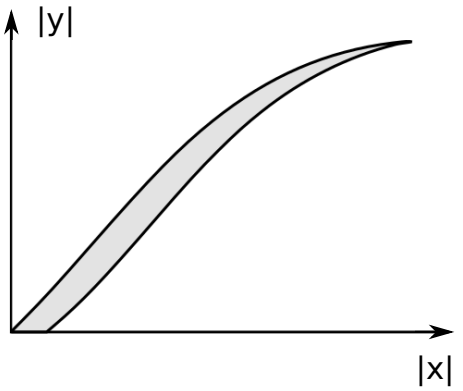
The nonlinear transfer characteristic of a dynamic nonlinear RF PA is completely described a a set of four characteristics: the amplitude modulation to amplitude modulation (AM/AM), the amplitude modulation to phase modulation (AM/PM), the phase modulation to phase modulation (PM/PM) and the phase modulation to amplitude modulation (PM/AM). Since the distortions exhibited by the RF PA are amplitude dependent and phase modulated signals have a constant amplitude, which is not affected by the distortions, RF PAs are usually characterized by AM/AM and AM/PM characteristics. Also the PM/PM and PM/AM distortions can be minimized by careful RF PA design and do not have a significant impact on the DPD performance [7].
Considering the complex baseband input $x(n)$ and output $y(n)$ signals

$$\begin{aligned} x(n) &= I_{\mathrm{in}}(n) + jQ_{\mathrm{in}}(n), \\ y(n) &= I_{\mathrm{out}}(n) + jQ_{\mathrm{out}}(n) \end{aligned} \tag{2.36}$$

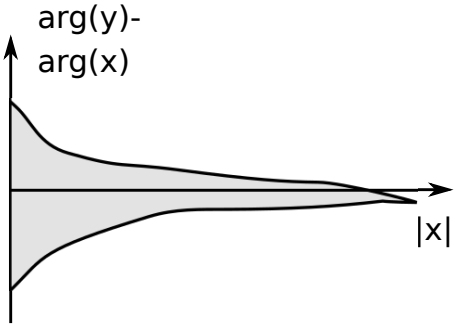
the instantaneous complex gain of the RF PA can be written as

$$\begin{aligned} G(|x(n)|) &= |G(|x(n)|)| e^{j\ arg[G(|x(n)|)]} \\ &= \frac{\sqrt{I_{\mathrm{out}}(n)^2 + Q_{\mathrm{out}}(n)^2}}{\sqrt{I_{\mathrm{in}}(n)^2 + Q_{\mathrm{in}}(n)^2}}\ e^{\left[tan^{-1}(\frac{Q_{\mathrm{out}}(n)}{I_{\mathrm{out}}(n)}) - tan^{-1}(\frac{Q_{\mathrm{in}}(n)}{I_{\mathrm{in}}(n)})\right]}. \end{aligned} \tag{2.37}$$

Note, that the gain $G(|x(n)|)$ is a function only of the input signals amplitude. This is only true, when PM/PM and PM/AM distortions are not exhibited by the RF PA. The AM/AM characteristic consists of the normalized output signal magnitude plotted as a function of the normalized input signal magnitude. The slope of this curve is directly related to the magnitude of the instantaneous gain $G(|x(n)|)$. Similarly the AM/PM characteristic contains the output signals phase variation plotted as a function of the normalized input signal magnitude, which is related to the phase of the instantaneous gain $G(|x(n)|)$. Examples for AM/AM and AM/PM characteristics are plotted in Figure 2.18. In the context of this thesis only modulated signals are used as test signals for the



(a) AM/AM example characteristic.



(b) AM/PM example characteristic.

Figure 2.18.: Examples of AM/AM and AM/PM characteristics.

RF PA respectively DPD. Since the AM/AM and AM/PM curves are generated using a sequence of samples from the input and output signals with different histories, they can also be used to get an impression of the memory effects exhibited by the RF PA. [7]

2.2.4.2. Error Vector Magnitude

An important metric for modulated signals is the so called error vector magnitude (EVM). It is also a time domain metric as the AM/AM and AM/PM characteristics and is defined as the magnitude of the error vector, which is calculated from the complex baseband signal as the difference between the distorted constellation vector and the ideal constellation vector of the reference constellation in the complex plane. The EVM is defined as

$$EVM = \sqrt{\frac{\frac{1}{N}\sum_{i=1}^{N}|S_i - S_{r,i}|^2}{\frac{1}{N}\sum_{i=1}^{N}|S_{r,i}|^2}}, \tag{2.38}$$

where S_i represents the received distorted constellation vector, $S_{\mathrm{r,i}}$ the reference constellation vector for the next closest constellation point and N is the total number of evaluated constellation vectors. An illustration for the EVM is shown in Figure 2.19.

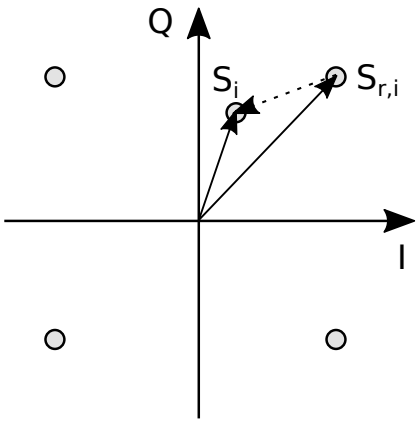


Figure 2.19.: EVM example. [7]

The EVM includes amplitude and phase errors in the received signal. Depending on the used transmission standard (e.g. UMTS) EVM thresholds are given, which should not be exceeded for reliable communication. [7]

2.2.4.3. Adjacent Channel Power Ratio

The adjacent channel power ratio (ACPR) is a frequency domain metric for modulated signals, which represents a critical linearity parameter for RF PAs respectively complete transmitters. As described in subsection 2.2.2, odd-order IM distortions can fall so close to the communication channel around the fundamental frequency that it is not possible to filter them out, since the transmitter may not only be used for transmitting on one but multiple adjacent channels. As a result, these distortions represent interference for neighboring channels and must be minimized. The transmission standard in use most probably includes a spectrum emission mask, which defines threshold values for the maximum power of the interference caused by the transmitter. The ratio between the power in the main communication channel and its neighboring channels is called the ACPR. It is calculated by integrating over a defined channel bandwidth in the power spectral density (PSD) of the transmitter for the main channel and its adjacent channels, which are defined by the same channel bandwidth and an offset frequency from the main channels center frequency. Figure 2.20 shows an example for the ACPR measurement parameters.

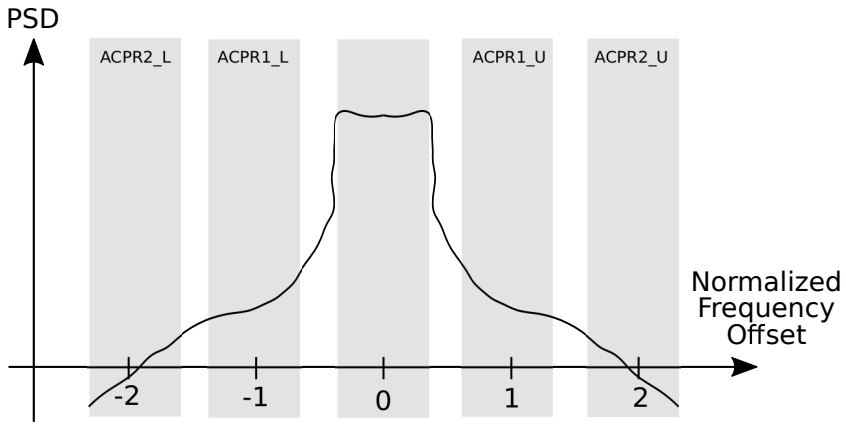


Figure 2.20.: ACPR example.

Depending on the transmission standard multiple adjacent channels can be defined. In this example there are two lower (ACPR1/2_L) and two upper (ACPR1/2_U) adjacent channels. [7]

2.3. Volterra Theory of Nonlinear Systems

The Volterra theory provides a mathematical framework for modeling nonlinear time invariant systems and forms the basis for the DPD algorithm used in this work. It was developed by the mathematician Vito Volterra and first applied for modeling nonlinear systems by Norbert Wiener [13]. It is based on the so called Volterra series

$$y(t) = \int_{-\infty}^{\infty} h_1(\tau_1)x(t-\tau_1)\,d\tau_1 + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_2(\tau_1,\tau_2)x(t-\tau_1)x(t-\tau_2)\,d\tau_1 d\tau_2 + \int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} h_n(\tau_1,\ldots,\tau_n)x(t-\tau_1)\ldots x(t-\tau_n)\,d\tau_1\ldots d\tau_n \tag{2.39}$$

in which for $n = 1,2,\ldots,$

$$h_n(\tau_1,\ldots,\tau_n) = 0 \qquad \text{for any } \tau_j < 0, \qquad j = 1,2,\ldots,n. \tag{2.40}$$

This is an infinite series of multi-dimensional convolutions where $h_n(\tau_1,\ldots,\tau_n)$ is called the Volterra kernel and can be interpreted as a n-dimensional impulse response. The Volterra series represents a non-parametric model for mapping the input of the system $x(t)$ to the output $y(t)$. In order to have a compact notation the so called Volterra operator of order n is defined as

$$H_n[x(t)] = \int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} h_n(\tau_1,\ldots,\tau_n)x(t-\tau_1)\ldots x(t-\tau_n)\,d\tau_1\ldots d\tau_n. \tag{2.41}$$

2.3.1. First-Order Volterra Systems

According to Equation 2.39 the input-output relationship of a first-order Volterra system is given by

$$y(t) = H_1[x(t)] = \int_{-\infty}^{\infty} h_1(\tau_1)x(t-\tau_1)\,d\tau_1. \tag{2.42}$$

It can be easily seen, that the first-order Volterra system is a linear time-invariant (LTI) system for which the superposition principle

$$y(t) = H_1[x(t)] = H_1\left[\sum_{n=1}^{N} c_\mathrm{n} x_\mathrm{n}(t)\right] = \sum_{n=1}^{N} c_\mathrm{n} H_1[x_\mathrm{n}(t)] \tag{2.43}$$

holds. Here, c_n represents the factors for the linear combination of input signals $x_\mathrm{n}(t)$. Therefore, LTI systems described by the linear operator $H_1[\cdot]$ can be viewed as a special case of the more general Volterra systems. [13]

2.3.2. Higher-Order Volterra Systems

In order to understand the characteristics of higher-order Volterra systems, first a second order system consisting of a LTI system containing all of the memory followed by a squarer is considered. Its block diagram is shown in Figure 2.21.

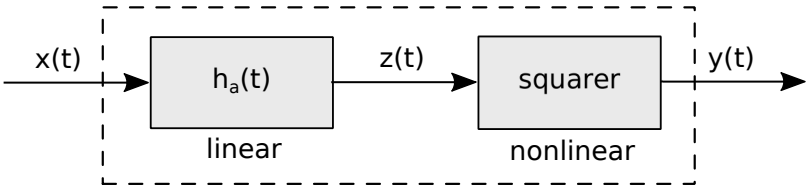


Figure 2.21.: Second order Volterra system - block diagram. [13]

The output of this system is given by

$$\begin{aligned} y(t) = [z(t)]^2 &= \left[\int_{-\infty}^{\infty} h_\mathrm{a}(\tau)x(t-\tau)\,d\tau\right]^2 \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_\mathrm{a}(\tau_1)h_\mathrm{a}(\tau_2)x(t-\tau_1)x(t-\tau_2)\,d\tau_1 d\tau_2 \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_2(\tau_1,\tau_2)x(t-\tau_1)x(t-\tau_2)\,d\tau_1 d\tau_2, \end{aligned} \tag{2.44}$$

which is equal to the definition of a Volterra system of second order $y(t) = H_2[x(t)]$. For an input signal $x(t)$ consisting of a linear combination of $x_1(t)$ and $x_2(t)$ the output signal of the system is

given by

$$y(t) = H_2[x(t)] = H_2[x_1(t) + x_2(t)] = H_2[x_1(t)] + 2H_2\{x_1(t), x_2(t)\} + H_2[x(t)] \tag{2.45}$$

and reveals that the nonlinearity of this system results from the bilinear Volterra operator

$$H_2\{x_1(t), x_2(t)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_2(\tau_1, \tau_2) x_1(t-\tau_1) x_2(t-\tau_2)\, d\tau_1 d\tau_2, \tag{2.46}$$

which is only linear in each of its input arguments when the other input argument is constant. According to the binomial identities for higher-order Volterra systems more and more p-linear Volterra operators are forming the nonlinearity. Therefore, all Voterra systems with order higher than one are nonlinear time-invariant systems. A real nonlinear system may not only consist of a single N-order Volterra operator but of a sum of multiple Volterra operators of different order. The corresponding input-output relation is given by

$$y(t) = H[x(t)] = \sum_{n=1}^{N} H_{\mathrm{n}}[x(t)] \tag{2.47}$$

where $H[\cdot]$ represents the combined Volterra operator of the system and N the total number of included Volterra operators. [13]

2.3.3. Cascade of Nonlinear Systems

A simple cascade of two nonlinear systems defined by the Volterra operators $H[\cdot]$ and $K_{(\mathrm{p})}[\cdot]$ will be evaluated in this section. The cascade as a whole can be described by the Volterra operator $Q[\cdot]$ and the corresponding block diagram is shown in Figure 2.22.

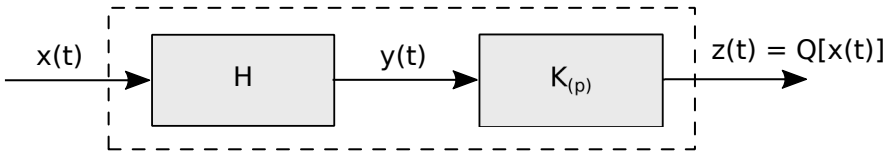


Figure 2.22.: Cascade of nonlinear systems - block diagram. [13]

In preparation for the pth-order inverse the first nonlinear system, whose input-output relation is given by

$$y(t) = H[x(t)] = \sum_{n=1}^{\infty} H_{\mathrm{n}}[x(t)] \tag{2.48}$$

and which consists of an infinite Volterra series, and the second nonlinear system defined by the

finite Volterra series

$$z(t) = K_{(\mathrm{p})}[y(t)] = \sum_{n=1}^{p} K_{\mathrm{n}}[y(t)] \tag{2.49}$$

are considered. In order to determine the output signal $z(t)$ of the cascade, the input to the system will be defined as $x_{\mathrm{in}}(t) = cx(t)$. This leads to the following power series representing the output signal

$$z(t) = Q[cx(t)] = \sum_{n=1}^{\infty} c^n Q_{\mathrm{n}}[x(t)]. \tag{2.50}$$

Further, the output of the cascade can also be represented by

$$z(t) = \sum_{m=1}^{p} K_m \left[\sum_{n=1}^{\infty} c^n y_{\mathrm{n}}(t) \right] \tag{2.51}$$

with $y_{\mathrm{n}}(t) = H_{\mathrm{n}}[x(t)]$ resulting in the power series

$$z(t) = \sum_{m=1}^{p} \sum_{n_1=1}^{\infty} \dots \sum_{n_m=1}^{\infty} c^{n_1+\dots+n_m} K_{\mathrm{m}}\{y_{\mathrm{n}_1}, \dots, y_{\mathrm{n}_m}\} \tag{2.52}$$

A comparison between the powers for the coefficients c reveals the structure of each of the individual Volterra systems $Q_{\mathrm{n}}[x(t)]$ in terms of the fundamental Volterra operators $H[\cdot]$ and $K_{(\mathrm{p})}[\cdot]$. Since the smallest possible value for $n_1 + \dots + n_m$ in Equation 2.52 is m, it follows that the first power of c has to be related to $m = 1$ and $n_1 = 1$. Using this result for Equation 2.51 it becomes clear that the structure of $Q_1[x(t)]$ is given by

$$Q_1[x(t)] = K_1[y_1(t)] = K_1[H_1[x(t)]]. \tag{2.53}$$

The block diagram for $Q_1[x(t)]$ is shown in Figure 2.23.

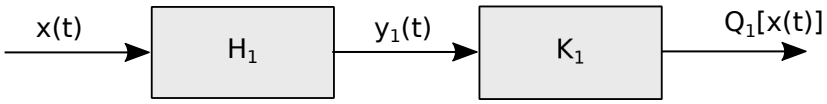


Figure 2.23.: Block diagram of the Volterra subsystem $Q_1[\cdot]$. [13]

The same procedure is used to obtain the structure of $Q_2[x(t)]$. Comparing the power series in terms of c^2 shows that this coefficient is generated for $m = 1$ with $n_1 = 2$ and $m = 2$ and $n_1 = n_2 = 1$. This gives the structure

$$Q_2[x(t)] = K_1[y_2(t)] + K_2[y_1(t)] = K_1[H_2[x(t)]] + K_2[H_1[x(t)]] \tag{2.54}$$

for the second order Volterra system. The corresponding block diagram is shown in Figure 2.24.

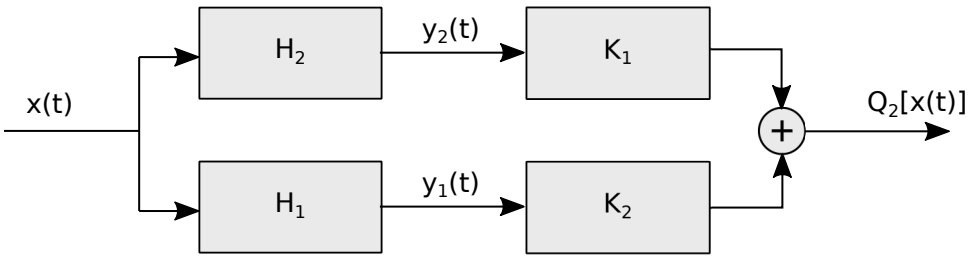


Figure 2.24.: Block diagram of the Volterra subsystem $Q_2[\cdot]$. [13]

This process of comparing the power series coefficients can be continued for determining all the remaining Volterra subsystems $Q_n[x(t)]$. The structure of these Volterra subsystems tends to be more and more complex with higher order. For the approximation of real nonlinear systems the contribution of higher order Volterra operators can often be neglected and the infinite Volterra series given in Equation 2.50 can be truncated. Note that in contrast to LTI systems, the order of nonlinear systems connected in cascade does have an influence of the overall system output. [13]

2.3.4. pth-Order Inverse

The description for a cascade of nonlinear systems can be used for determining the pth-order inverse of the system defined by the Volterra series $H[x(t)]$. The pth-order inverse represents the basis for the DPD algorithm used in this thesis and is used for compensating the nonlinear transfer characteristic of the RF PA. Considering the system depicted in Figure 2.22, pth-order means that only the first p Volterra operators of $H[x(t)]$ are being inverted. So the input-output relation of the overall system can be represented as

$$Q[x(t)] = x(t) + \sum_{n=p+1}^{\infty} Q_n[x(t)]. \tag{2.55}$$

From this it follows that the first order Volterra operator of the cascaded system needs to fulfill the following constraint

$$Q_1[x(t)] = K_1[H_1[x(t)]] = x(t). \tag{2.56}$$

Since this first order Volterra operator is linear, this leads to $K_1[\cdot] = H_1^{-1}[\cdot]$. Where $K_1[\cdot]$ is the inverse of the linear operator $H_1[\cdot]$. Note that a pth-order inverse only exists for a stable and causal operator $K_1[\cdot]$. For the second order Volterra operator of the cascaded system the output needs to be zero

$$Q_2[x(t)] = K_1[H_2[x(t)]] + K_2[H_1[x(t)]] = 0. \tag{2.57}$$

It can be seen that this is expression is true for $K_2[\cdot] = -K_1[H_2[H_1^{-1}[\cdot]]]$. The same procedure can be applied for determining the remaining Volterra operators $K_n[\cdot]$ of the cascaded system up to order p. Note that in Figure 2.22 the inverting Volterra operator $K_{(p)}$ is placed after the Volterra operator

$H[\cdot]$, which is being inverted. As a result, $K_{(p)}$ is called the post-inverse of $H[\cdot]$. Although in general the positioning of nonlinear systems in a cascade has an influence on the output of the overall system, it was shown in [13] that the post-inverse operator $K_{(p)}$ is equivalent to the pre-inverse operator, which is located before $H[\cdot]$ as shown in Figure 2.25.

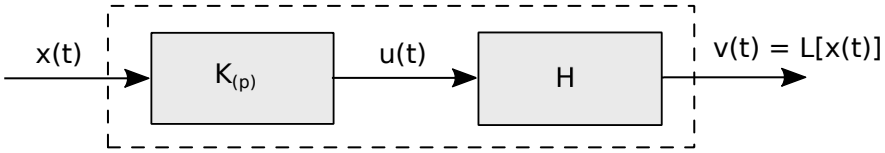


Figure 2.25.: pth-order pre-inverse of the Volterra system $H[x(t)]$.

However, it should be noted that only the first p terms of pre- and post-inverse are identical but not the remaining terms as described by the following equation

$$L[x(t)] = x(t) + \sum_{n=p+1}^{\infty} L_n[x(t)]. \tag{2.58}$$

3. Digital Pre-Distortion

In this chapter first an overview of different approaches for the linearization of nonlinear systems using DPD is given. Further, the specific DPD algorithm used in the context of this thesis is presented and the important aspect of normalization gain is discussed.

3.1. Overview

The basic concept of DPD is the linearization of a nonlinear system by the use of a digital signal processing block, which is placed in front of the nonlinear system and distorts the input signal in such a way that the combined transfer characteristic of the DPD and nonlinear system is linear again. The corresponding block diagram is shown in Figure 1.3. The nonlinear system is represented by the PA to be linearized.
Several different approaches are available for DPD. These approaches can be classified according to the position of the DPD in the estimation loop for the DPD coefficients. When the DPD is located inside the estimation loop, the system is called closed-loop. The block diagram for a closed-loop DPD system is shown in Figure 3.1. The error signal consisting of the difference between the

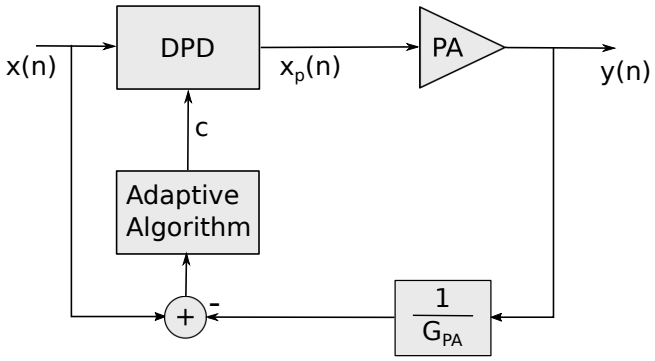


Figure 3.1.: Closed-loop DPD system. [7]

input signal of the system x and the output signal y of the PA divided by the expected gain of the linearized PA G_{PA} is fed back into an adaptive algorithm. The adaptive algorithm minimizes the error signal by updating the DPD coefficients c of the DPD. Closed-loop DPD systems usually have a slow convergence speed and high computational complexity, since there is no direct relation between the error signal and the DPD coefficients. Furthermore, they also exhibit issues with divergence when the PA is driven into saturation and the adaptive algorithm tries to compensate for the missing output power by increasing the power of the pre-distorted input signal x_{p} [7].

In an open-loop DPD system the DPD is located outside of the estimation loop. Here the input signal x_p and output signals y of the PA divided by the expected gain of the linearized PA G_{PA} are fed back into the adaptive algorithm. The corresponding block diagram is shown in Figure 3.2. The

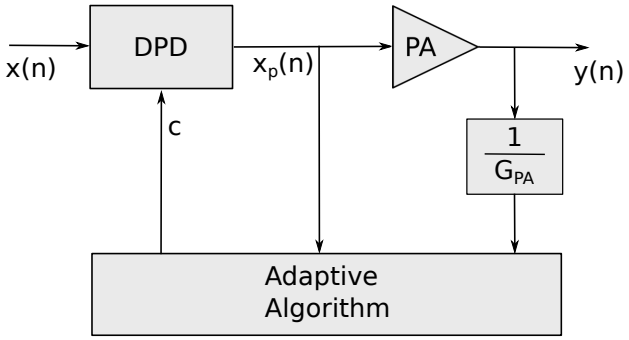


Figure 3.2.: Open-loop DPD system. [7]

fact that the input and output signals of the PA are directly observed by the adaptive algorithms allows the open-loop DPD system to be implemented in two distinct variants. In the first variant a model of the PA is extracted from the input and output signals and inverted, which provides the coefficients for the DPD. This variant is known as the direct learning architecture. Its block diagram is shown in Figure 3.3. The PA model estimates the linearly scaled version of the output

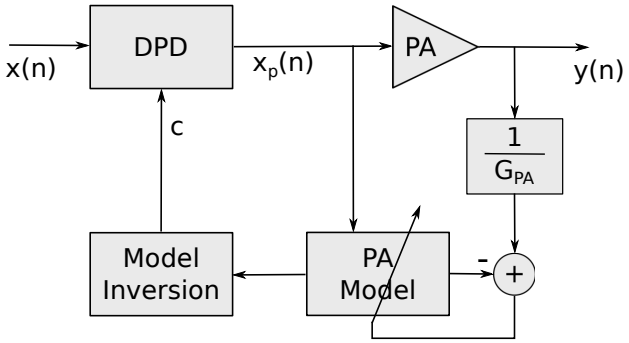


Figure 3.3.: Open loop DPD system with direct learning architecture. [7]

signal y/G_{PA} by adopting the PA model parameters in order to minimize the error signal, which is represented by the difference of the estimated and actual output signal of the PA. The PA model is then inverted and the corresponding coefficients are used for the DPD. Memory effects exhibited by the PA can cause fitting errors in the model inversion process and degrade DPD performance. The second variant of open-loop DPD is based on the pth-order inverse theory and is called indirect learning architecture. Here the input and linearly scaled output signals of the PA are used for estimating the pth-order post-inverse of the PA. As mentioned in subsection 2.3.4, the pth-order post-inverse is equivalent to the pth-order pre-inverse and therefore the estimated post-inverse parameters are directly copied into the DPD. The block diagram is shown in Figure 3.4. An important advantage of the open-loop approach for DPD is that it does not require a real-time implementation of the adaptive algorithm as it is the case for the closed-loop approach. This is due to the fact, that in an open-loop DPD system it is not necessary to know the coefficients used in the

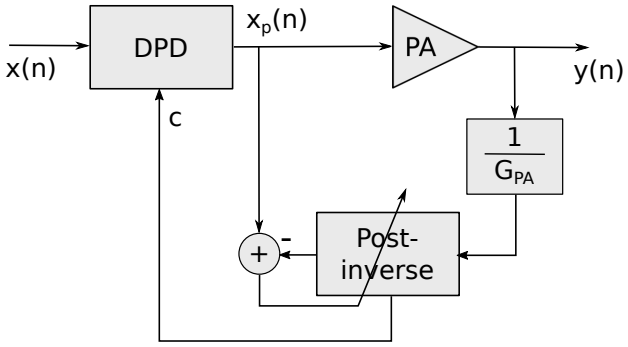


Figure 3.4.: Open-loop DPD system with indirect learning architecture. [7]

DPD for the estimation process. As a result, the estimation process in the open-loop approach can be implemented offline in software.
An another aspect of DPD system, which greatly influences the performance and complexity, is the nonlinear model used for implementing the DPD or PA modeling. An overview of the complexity and performance for selected models is given in Figure 3.5. Since an extensive description of all of

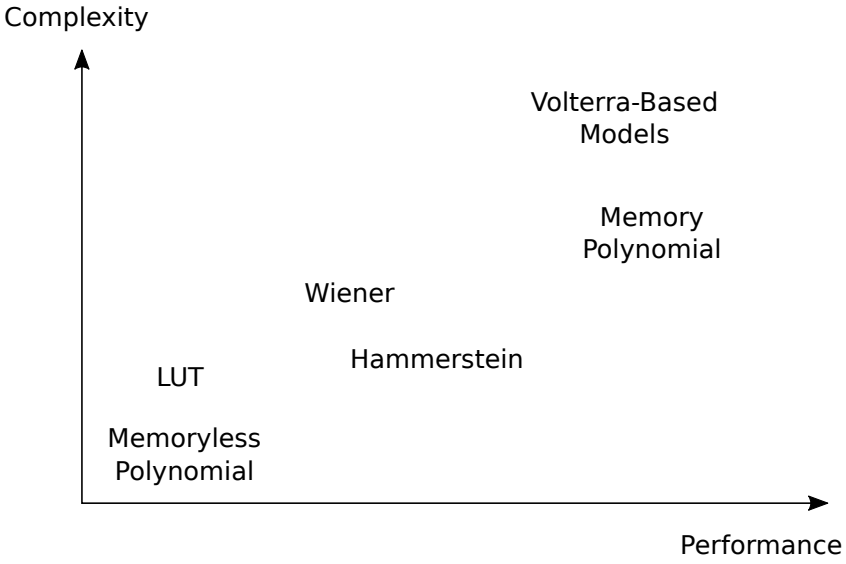


Figure 3.5.: Complexity and Performance of different DPD models. [6]

the models mentioned in Figure 3.5 would be beyond the scope of this thesis, only the basic concept of the different models is described. The most basic and least powerful model is the memoryless polynomial model, which uses only the current sample as an input $x(n)$ for a polynomial function. For discrete systems with nonlinearity order P it is defined by

$$y_{\mathrm{P}}(n) = \sum_{p=1}^{P} c_{\mathrm{p}} x(n)^p, \tag{3.1}$$

where $y_{\mathrm{P}}(n)$ represents the output sample and c_{p} is the complex factor for the pth nonlinearity. The look-up table (LUT) model is comparable in complexity and performance to the memoryless

polynomial model. It is also a memoryless model, which consists of two LUTs containing the phase and amplitude factors for individual power levels of the input signal. The Wiener model consists of a cascade of a FIR filter and memoryless nonlinear function such as the LUT model. Its block diagram is given in Figure 3.6. The input output relation is given by

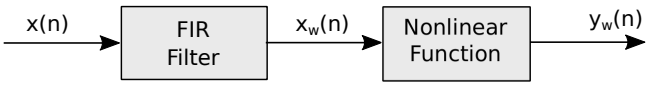


Figure 3.6.: Wiener model - block diagram. [7]

$$y_{\mathrm{w}}(n) = G_{\mathrm{w}}[|x_{\mathrm{w}}(n)|]x_{\mathrm{w}}(n), \tag{3.2}$$

where $G_{\mathrm{w}}[\cdot]$ is the complex gain implemented by memoryless nonlinear function and

$$x_{\mathrm{w}}(n) = \sum_{m=0}^{M} a_{\mathrm{m}} x(n-m) \tag{3.3}$$

is the output of the FIR filter with coefficients a_{m} and memory depth M. The Hammerstein model exhibits a very similar structure as the Wiener model. It consists of a memoryless nonlinear function followed by a FIR filter. The block diagram is given in Figure 3.7. As a consequence the input

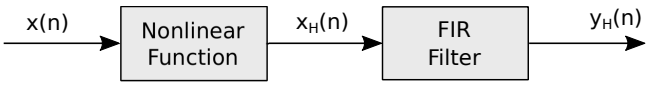


Figure 3.7.: Hammerstein model - block diagram. [7]

output relation is given by

$$y_{\mathrm{H}}(n) = \sum_{m=0}^{M} a_m x_{\mathrm{H}}(n-m), \tag{3.4}$$

where M is the memory depth, a_{m} are the FIR filter coefficients and

$$x_{\mathrm{H}}(n) = G_{\mathrm{H}}[|x(n)|]x(n) \tag{3.5}$$

is the output of the memoryless nonlinear function with the complex gain $G_{\mathrm{H}}[\cdot]$. Due to the similar structure of the Wiener and Hammerstein model the complexity and performance of these models are similar. The memory polynomial model is a simplification of the Volterra series based model, where all cross elements of the Volterra kernel respectively n-dimensional impulse response $h_{\mathrm{n}}(\tau_1,...,\tau_{\mathrm{n}})$ in Equation 2.39 are set to zero. This leads to the following input output relation for a discrete system

$$y_{\mathrm{MP}}(n) = \sum_{m=0}^{M} \sum_{k=1}^{K} a_{\mathrm{mk}} x(n-m)|x(n-m)|^{k-1}, \tag{3.6}$$

where $y_{\mathrm{MP}}(n)$ is the output sample, M is the memory depth, K is the order of nonlinearity and a_{mk} are the model parameters. Due to this simplification, the memory polynomial model is less complex but also less powerful than a Volterra series based model. In the context of this work a Volterra series based model has been used. A detailed description of it is given in the next section. [7]

3.2. Open Loop Digital Pre-Distorter

For this work an open loop DPD system with indirect learning architecture and a Volterra series based nonlinear model has been selected, because this combination provides good performance and the estimation process can be implemented offline using software. The implementation in this work is based on the DPD system presented in [19] but does not include the proposed undersampling and parameter interpolation steps in the model extraction respectively DPD process.

3.2.1. Algorithm and Setup

The open loop DPD system used in this work can be divided into three main sections: the stimulus, the device under test (DUT) and the monitor. The stimulus section contains a signal generator, the DPD, a switch and a upconversion module. The DUT is the RF PA, which will be linearized and the monitor section contains a downconversion module and the signal analysis. The corresponding block diagram is shown in Figure 3.8. The signal generation is implemented completely in software

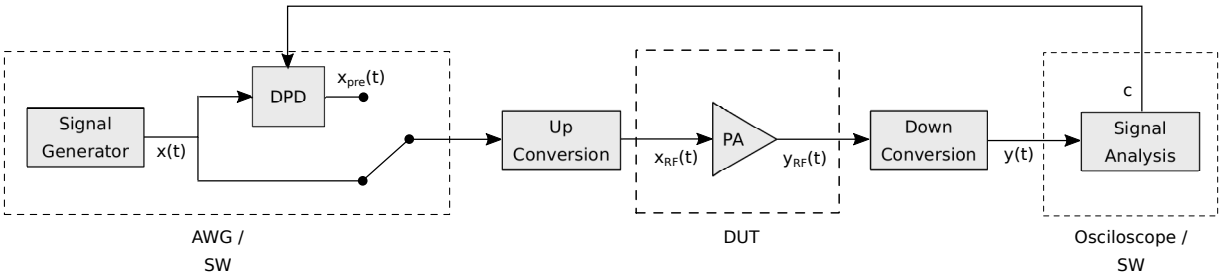


Figure 3.8.: Open loop DPD system used in the context of this thesis.

(Matlab). The resulting sample sequence is then converted into an analog complex baseband signal by an arbitrary waveform generator (AWG). Depending on the DUT and its operating frequency range, the upconversion process is either implemented in the analog domain by a dedicated IQ mixer for high operating frequencies or in the digital domain for lower operating frequencies. The same is true for the downconversion process. The signal analysis is implemented by a high-performance oscilloscope and software.
The DPD algorithm consists of five consecutively executed steps:

Signal Generation A uniformly distributed sequence of pseudorandom integers in the range of $[0,M]$ is generated and mapped to their corresponding modulation symbols by an M-QAM modulator. The resulting complex baseband signal is then passed through a root raised-

cosine filter in order to limit its bandwidth so it fits into the available channel bandwidth. Usually a rolloff factor of 0.35 is used. The output of the signal generation step is the complex baseband signal $x(n)$ respectively $x(t)$ when it has been converted into an analog signal by the AWG.

Data Acquisition The switch in Figure 3.8 is set to the lower position so the signal $x(t)$ is up converted into the desired frequency band and it is passed to the DUT. The signal power of $x_{\mathrm{RF}}(t)$ is set in such a way that the RF PA will be driven into its saturation region where it exhibits a nonlinear transfer characteristic. The distorted output signal of the RF PA $y_{\mathrm{RF}}(t)$ is then down converted and recorded by the oscilloscope. Note that due to the bandwidth expansion caused by the nonlinearity of the RF PA the Nyquist frequency and analog bandwidth of the oscilloscope need to be multiple times that of the undistorted signal. The actual values depend on the order of nonlinearity exhibited by the RF PA.

Signal Alignment and Parameter Extraction The distorted discrete complex baseband signal $y(n)$, which has been recorded by the oscilloscope, needs to be properly aligned to the input $x(n)$ in preparation for the extraction of the model parameters of the post-inverse respectively DPD. Depending on the hardware setup the signal alignment process may include carrier frequency offset compensation, sampling rate adaption and time alignment. Carrier frequency offset compensation is necessary when the up- and downconversion modules are not operating in a coherent manner. Sampling rate adaption is necessary when AWG and oscilloscope are not using the same sampling rates. It can be performed in a simple way by interpolation and decimation when the AWG and oscilloscope sampling clocks are phase locked. The time alignment can be performed by using cross-correlation, when the signals have the same sampling frequency and carrier frequency offset. After the signals have been properly aligned the signals are scaled to their respective maximum values. The consequences of this scaling operation are described in subsection 3.2.4. Finally, the parameters of the post-inverse respectively DPD model c can be extracted. More details on the parameter extraction are given in subsection 3.2.3.

Pre-distortion With the model parameters c being available, the pre-distortion is performed by changing the switch to its upper position in Figure 3.8. As a result the exact same input signal $x(t)$ from the signal generation step is passed through the DPD model and the rest of the signal chain. The linearized output signal of the RF PA is then being recorded by the oscilloscope.

Evaluation In order to properly evaluate the DPD performance, it is necessary to repeat the data acquisition step with a reduced power level of the input signal. This is due to the fact, that the pre-distorted output signal has a reduced average power level compared to the distorted output signal and therefore a comparison between these two signals would not be fair. The distorted and pre-distorted output signals need to have the same average power levels. The reason for the reduction in output power for pre-distorted signals is explained in subsection 3.2.4.

3.2.2. Nonlinear Model

The nonlinear model used in the DPD is based on the Volterra series. However, the classical Volterra series as given in Equation 2.39 is impractical, since the number of parameters grows exponentially with the order of nonlinearity and memory length. In order to reduce the number of parameters to be estimated, a new modified Volterra series with dynamic deviation was proposed in [18]. This approach is based on the following modifications to the classical Volterra series: The input-output relation for discrete nonlinear systems with finite memory can be written as a discrete form of the Volterra series

$$y(n) = \sum_{p=1}^{P} \left[\sum_{i_1=0}^{M-1} \dots \sum_{i_p=0}^{M-1} h_{\mathrm{p}}(i_1, \dots, i_{\mathrm{p}}) \prod_{r=1}^{P} x(n - i_{\mathrm{r}}) \right], \tag{3.7}$$

where P is the order of nonlinearity, M is the memory depth and $h_{\mathrm{p}}(i_1, \dots, i_{\mathrm{p}})$ represents the Volterra kernel. By introducing the dynamic deviation function

$$e(n, i) = x(n - i) - x(n), \tag{3.8}$$

which represents the deviation of the delayed $x(n - i)$ and current input sample $x(n)$, it is possible to split into a static $y_{\mathrm{s}}(n)$ and dynamic $y_{\mathrm{d}}(n)$ part:

$$y(n) = y_{\mathrm{s}}(n) + y_{\mathrm{d}}(n). \tag{3.9}$$

The static part is given by

$$y_{\mathrm{s}}(n) = \sum_{p=1}^{P} a_{\mathrm{p}} x^p(n) \tag{3.10}$$

and the dynamic part by

$$y_{\mathrm{d}}(n) = \sum_{p=1}^{P} \sum_{r=1}^{P} x^{p-r}(n) \cdot \sum_{i_1=0}^{M-1} \dots \sum_{i_r=0}^{M-1} g_{\mathrm{p_r}}(i_1, \dots i_{\mathrm{r}}) \prod_{j=1}^{r} e(n, i_j). \tag{3.11}$$

After resubstitution of Equation 3.8 in Equation 3.11 and rearrangement the discrete Volterra series in Equation 3.7 can be expressed as

$$y(n) = \sum_{p=1}^{P} h_{\mathrm{p}} x^p(n) + \sum_{p=1}^{P} \left(\sum_{p=1}^{P} \left[x^{p-r}(n) \sum_{i_1=1}^{M-1} \dots \sum_{i_r=i_{\mathrm{r}-1}}^{M-1} h_{\mathrm{p_r}}(i_1, \dots, i_r) \prod_{j=1}^{r} x(n - i_j) \right] \right). \tag{3.12}$$

In this form the parameter r is directly related to the number of delayed input samples $x(n - i_j)$ to be included in the model. This allows for controlling the degree of dynamics in the model independently of the order of nonlinearity P, which is not possible in the classical Volterra series given in Equation 3.7. As a consequence the modified Volterra series with dynamic deviation

reduction be used for modeling highly nonlinear systems with reduced model complexity due to the removal of high-order dynamics. In most communications systems the available channel bandwidth is limited and it makes sense to use a corresponding complex baseband representation of the nonlinear model used for DPD. Considering the dynamic deviation recduction Volterra model with only first order dynamics ($r = 1$) and the fact, that only odd-order nonlinearities contribute distortions, which fall in or close to the signal band, the equivalent complex baseband model can be written as [19]

$$\tilde{y}(n) = \sum_{p=0}^{\frac{P-1}{2}} \sum_{i=0}^{M} \tilde{g}_{2\mathrm{p}+1,1}(i)|\tilde{x}(n)|^{2p}\tilde{x}(n-i) + \sum_{p=1}^{\frac{P-1}{2}} \sum_{i=1}^{M} \tilde{g}_{2\mathrm{p}+1,2}(i)|\tilde{x}(n)|^{2(p-1)}\tilde{x}^2(n)\tilde{x}^*(n-i). \tag{3.13}$$

As described in [18], a truncation of Equation 3.12 after the first order dynamics can be used for modeling highly nonlinear systems but its effectiveness is limited to systems with short nonlinear memory duration compared to the bandwidth of the baseband signal. However, it is also stated that memory effects in RF PAs decrease with time so that input samples with longer time delay have a reduced impact on the output signal of the model. Also nonlinear dynamics of higher order tend to have less influence on the output signal. As a result it is reasonable to limit the the dynamic deviation reduction Volterra model to first- ($r = 1$) or second order ($r = 2$) dynamics.

3.2.3. Parameter Extraction

Note that the output signal $y(n)$ of the dynamic deviation reduction Volterra model given in Equation 3.12 is linear with respect to its coefficients $h_{\mathrm{p_r}}(i_1,...,i_{\mathrm{r}})$. This means that classic linear system identification algorithms can be used for model parameter extraction using the input and output samples of the nonlinear system. In [19] the use of the least squares (LS) estimation algorithm is suggested. Since there is no reasonable assumption for the noise distribution of the output signal available, maximum likelihood estimation cannot be used for parameter extraction.
The post-inverse used in the indirect learning architecture is a nonlinear system and can therefore also be represented in the form of Equation 3.13 by swapping the input $\tilde{x}(n)$ and output samples $\tilde{y}(n)$. The dynamic deviation reduction Volterra model for the post-inverse can then be written as

$$\tilde{x}(n) = \sum_{p=0}^{\frac{P-1}{2}} \sum_{i=0}^{M} \tilde{g}_{2\mathrm{p}+1,1}(i)|\tilde{y}(n)|^{2p}\tilde{y}(n-i) + \sum_{p=1}^{\frac{P-1}{2}} \sum_{i=1}^{M} \tilde{g}_{2\mathrm{p}+1,2}(i)|\tilde{y}(n)|^{2(p-1)}\tilde{y}^2(n)\tilde{y}^*(n-i), \tag{3.14}$$

where $\tilde{x}(n)$ represents the output and $\tilde{y}(n)$ the input samples of the post-inverse. In preparation for the LS estimation the unknown model parameters $\tilde{g}_{2\mathrm{p}+1,\mathrm{j}}(i)$ in Equation 3.2.3 are rearranged into a parameter vector $\mathbf{\Theta}$ and a matrix $\mathbf{\Upsilon}$ containing all the product terms $\tilde{y}(n), \tilde{y}(n-i), ..., |\tilde{y}(n)|^2\tilde{y}(n), ...$ is constructed. The output of the model is $\mathbf{z} = [\tilde{z}(M+1), ..., \tilde{z}(L])^T$ and the unmodeled error is $\mathbf{e} = [\tilde{e}(M+1), ..., \tilde{e}(L])^T$, where $e(n) = \tilde{x}(n) - \tilde{z}(n)$ and L is the length of the sample sequences recorded in the data acquisition step. With this new notation the signal model for the post-inverse can be

rewritten as

$$\mathbf{x} = \mathbf{Y}\boldsymbol{\Theta} + \mathbf{e}. \tag{3.15}$$

This problem formulation allows the use of the well-known LS estimator, which minimizes the cost function

$$J(\boldsymbol{\Theta}) = \sum_{n=M+1}^{L} |e(n)|^2 = \mathbf{e}^H\mathbf{e} \tag{3.16}$$

and has the solution

$$\hat{\boldsymbol{\Theta}} = arg\ \min_{\boldsymbol{\Theta}} J(\boldsymbol{\Theta}) = (\mathbf{Y}^H\mathbf{Y})^{-1}\mathbf{Y}^H\mathbf{x}. \tag{3.17}$$

The resulting parameter vector $\hat{\boldsymbol{\Theta}}$ and the post-inverse model given in can then be used to calculate the estimate $\hat{x}(n) = \tilde{z}(\tilde{y}(n), \hat{\boldsymbol{\Theta}})$. In order to avoid additional calibration effort, the measured input and output samples of the RF PA are normalized to their maximum absolute values

$$\tilde{x_{\mathrm{n}}}(n) = \frac{\tilde{x}(n)}{\tilde{x}_{\max}}, \quad \tilde{y_{\mathrm{n}}}(n) = \frac{\tilde{y}(n)}{\tilde{y}_{\max}}, \tag{3.18}$$

where the maximum absolute vales $\tilde{x}_{\max} = max[|\tilde{x}(n)|]$ and $\tilde{y}_{\max} = max[|\tilde{y}(n)|]$ are calculated for all $n = 1, ..., L$.

3.2.4. Expected Gain

Since the superposition principle is not obeyed by nonlinear systems, it is crucial to normalize the data used for parameter extraction to the same power level as the input signal. This allows the parameters to be used directly for the DPD process, otherwise a nonlinear scaling needs to be applied. The normalization defines the expected gain of the DPD system as a whole. Normalization of the input and output samples to their respective maximum values as given in Equation 3.18 will lead to the expected gain of

$$G_{\mathrm{ex}} = \frac{max[|\tilde{y}(n)|]}{max[|\tilde{x}(n)|]} = \frac{y_{\max}}{x_{\max}}. \tag{3.19}$$

This expected gain is smaller than the maximum gain of the RF PA but it is equivalent to a DPD system with higher expected gain in terms maximum and average output power. This fact becomes obvious when the AM/AM characteristics of these DPD systems are compared as can be seen in Figure 3.9. For the linearization of the nonlinear AM/AM characteristics of the RF PA two points of its curve (red) are connected with a straight line, whose slope equals the expected gain of the DPD system. Normalization to the maximum absolute values of the input and output samples as in Equation 3.18 leads to the linearized AM/AM characteristics shown by the blue curve. Note that due to gain compression and the resulting peak to average power ratio (PAPR) reduction of the

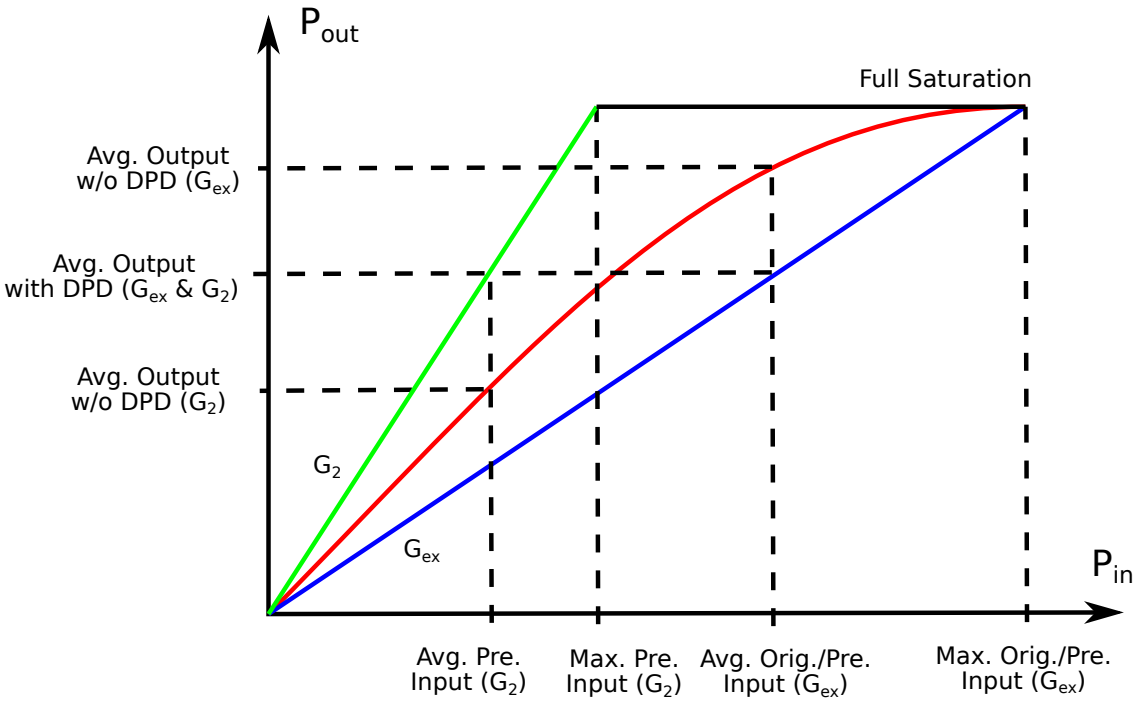


Figure 3.9.: AM/AM characteristics of nonlinearized RF PA (red), DPD system with expected gain G_{ex} (blue) and DPD system with higher gain (green).

output signal, the average output power of the linearized RF PA will be smaller than the average output power of the RF PA without DPD. However, the maximum output power will be the same for both cases.

A DPD system with higher gain G_1 (green curve) can be constructed from the DPD system with gain G_{ex} by scaling the amplitude of input signal with a corresponding factor greater than 1. Since the linearization resulting in G_{ex} is based on the assumption of stationarity, it is necessary to reduce the input power applied to the scaled DPD system with G_1 in order to keep the actual input power applied to the RF PA is the same. This constraint leads to a shift of the maximum and average input power to lower levels, which will ultimately result in the same average and maximum output power levels as for the DPD system with gain G_{ex}. This means that the DPD systems with gain G_{ex} and G_1 are equivalent in terms of linearization performance [19].

Note that the reduction of the average output power of the DPD systems means, that the pre-distorted output signals have to be compared to distorted output signals with the same average output power in order to make a fair performance evaluation of the DPD performance.

3.2.5. Bandwidth

The intermodulation distortions caused by the RF PAs nonlinear transfer characteristic cause spectral regrowth in the output signal. Depending on the order of nonlinearity the increase in bandwidth can be several times the bandwidth of the input signal. As discussed in subsection 2.2.2, a nonlinear system of order P will cause pth-order in-band intermodulation products, which appear at the

same frequencies as the input signal or are located too close to the signal band to be filtered out. In order to accurately capture the nonlinear effects exhibited by the RF PA and extract the model parameters for the DPD, an increased sampling rate is required. Since the pre-distorted signal is also generated by a pth-order nonlinear system, this increased sampling rate is also necessary for proper reconstruction of the pre-distorted signal. This spectral regrowth in the DPD system is shown in Figure 3.10. Conventional DPD systems typically use a sampling rate, which is five times the input

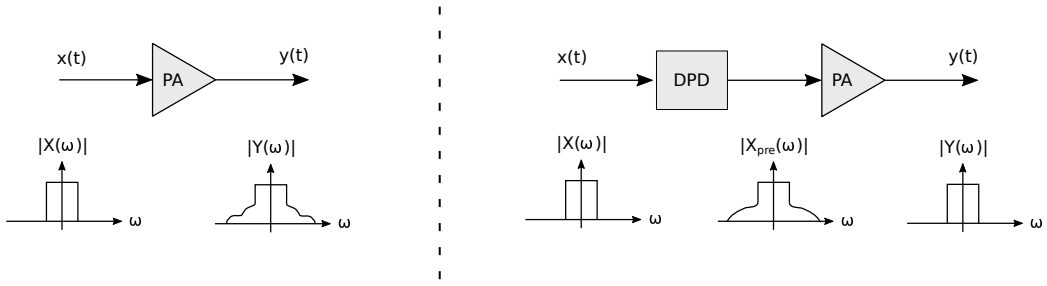


Figure 3.10.: Spectral regrowth in DPD systems.

signal bandwidth [11]. As a result, with increasing input signal bandwidth hardware requirements for the DPD system get more demanding and it becomes more difficult to implement closed-loop DPD systems on dedicated signal processing devices such as DSPs or FPGAs. Open-loop DPD system are generally less effected, since the parameter extraction can be performed in software on a PC. However, also for open-loop DPD systems the waveform generator and oscilloscope need to support the increased sampling rates.

Despite the increase in hardware sampling rates, a higher input signal bandwidth also leads to problems in the analog domain. When analog IQ modulation respectively a direct conversion architecture is used the IQ imbalance over frequency can be come the major source for increased EVM values in the received constellation diagrams. Another aspect, which effects all architectures, are frequency dependent amplitude and phase variations in the transfer characteristics of the analog signal path including cables, up/down converters, filters, amplifiers and oscilloscope input channel [14]. In addition the thermal noise power given by [2]

$$P_{\mathrm{n,thermal}} = k_{\mathrm{B}} T B, \tag{3.20}$$

which is present in all analog devices, increases proportional with the bandwidth B. Here k_{B} represents the Boltzmann constant and T the temperature. The thermal noise together with other unknown noise sources causes distortions in the output signal of the RF PA, which can have an negative impact on the DPD performance.

4. Simulations

The parameter extraction and pre-distortion steps of the open-loop DPD algorithm presented in section 3.2 have been implemented in software using Matlab. Before testing the DPD algorithm on an actual RF PA in hardware simulations have been performed in order to verify the correct implementation of the algorithm. For this purpose a comlpex baseband model for the RF PA based on the dynamic deviation reduction Volterra series with moderate nonlinearity and memory effects has been used. The parameters $\tilde{g}_{2p+1,1}(i)$ and $\tilde{g}_{2p+1,2}(i)$ for this model as given in Equation 3.13 are listed in Table 4.1 and Table 4.2. These parameters are not based on measurements of an actual RF

$2p+1$	$i=0$	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$
1	10	0	0	0	0	0	$-2.3-j1$	-0.7
3	-0.1	0	0	0	0	0	0.03	-0.02
5	-0.001	0	0	$j0.01$	0	0	0	$j0.01$

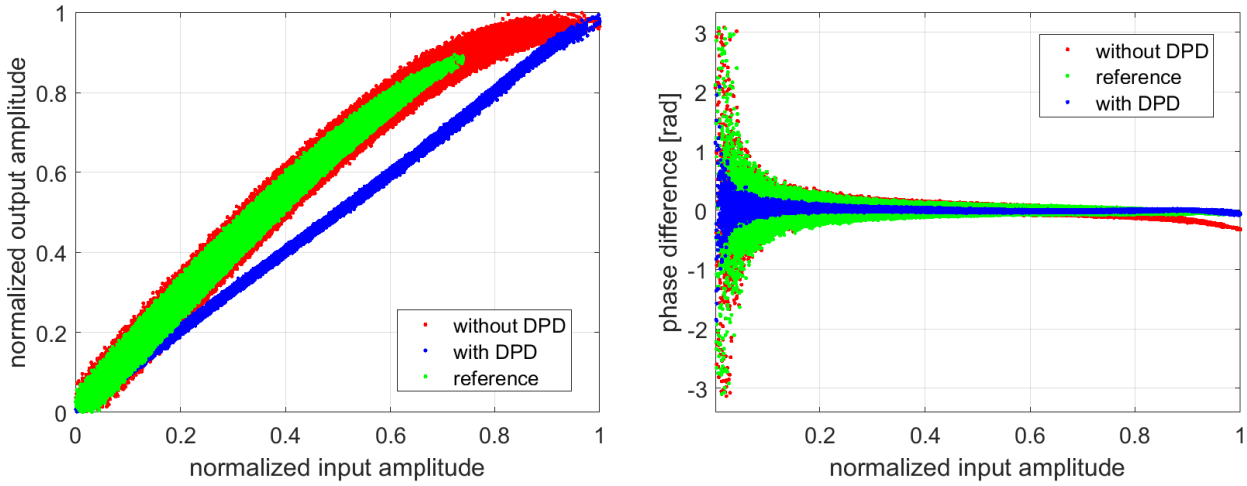
Table 4.1.: Simulation model parameters $\tilde{g}_{2p+1,1}(i)$.

$2p+1$	$i=0$	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$
3	0.03	0	0	0	0	0	0	0.005
5	0	0	0	0	0	-0.001	-0.005	0

Table 4.2.: Simulation model parameters $\tilde{g}_{2p+1,2}(i)$.

PA but they were chosen by adjusting the resulting AM/AM and PM/AM transfer characteristics to look similar to that of the RF PA used in [19]. Further, the fact that the influence of higher order dynamics as well as nonlinearities on the output signal declines has been considered. As can be seen from the dimensions of Table 4.1 and Table 4.2 the order of nonlinearity of the model is $P = 5$ and the memory depth is $M = 8$. In addition white Gaussian noise is added at the model output for modeling the intrinsic noise sources of the RF PA. The model for the post-inverse respectively DPD uses the same order of nonlinearity and memory depth as the RF PA model.
The input signal for the RF PA model is constructed from a uniformly distributed random source, whose output sequence is mapped to complex constellation symbols by a 16-QAM modulator and filtered with a root raised-cosine filter with a rolloff factor of $r = 0.35$. The input signal is then passed through the RF PA model. The corresponding output signal is used for parameter extraction using LS estimation. Finally, the input signal is passed through the DPD with the previously extracted model parameters and again through the RF PA model. The resulting input and output signals

are then used for creating the AM/AM and AM/PM plots as shown in Figure 4.1. As discussed in subsection 3.2.4, in addition to the distorted signal (red) and the pre-distorted signal (blue) a distorted reference signal (green) with the same channel power as the pre-distorted signal is generated in order to ensure a fair performance evaluation of the DPD system. The simulation has been performed with a SNR of 40 dB. The gain compression of the RF PA model is clearly visible in

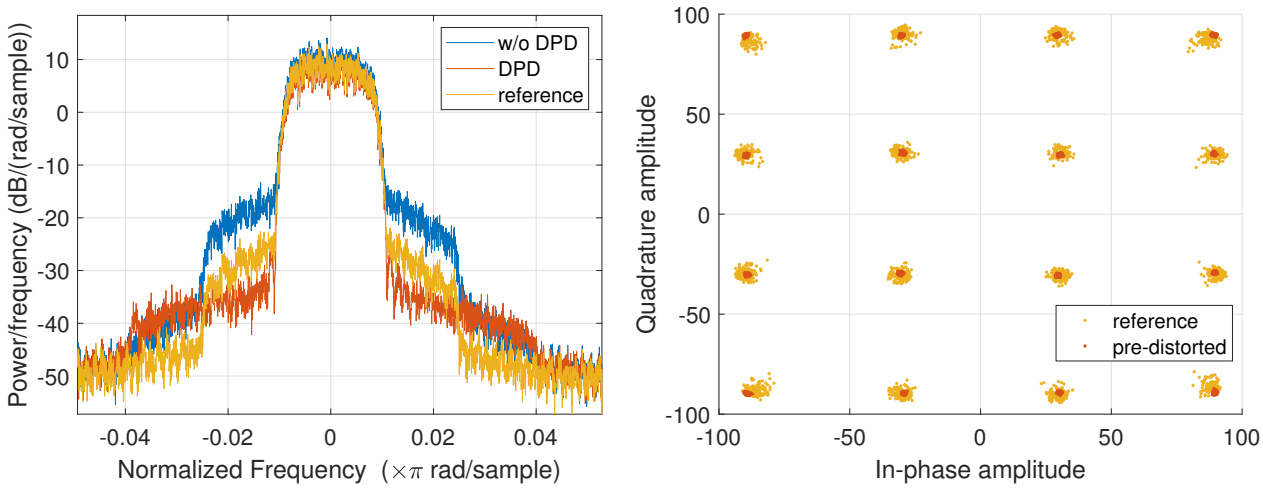


(a) AM/AM characteristics.

(b) AM/PM characteristics.

Figure 4.1.: AM/AM and AM/PM characteristics of nonlinearized RF PA (red), pre-distorted signal (blue) and reference signal (green) for $SNR = 40$ dB.

the AM/AM characteristics in Figure 4.1a when looking at the distorted signal and its slope, which is decreasing with the normalized input amplitude. Also the memory effects can be noticed when comparing the variance of the distorted and pre-distorted signals. The variance of the distorted signal is greater than that of the pre-distorted signal. This is caused by the memory effects, which create slightly different AM/AM characteristics depending on the history of the signal and the superposition of these characteristics in the plot. In the AM/PM characteristics in Figure 4.1b a slight deviation of the phase from the optimal value (0) for large normalized input amplitudes can be observed on the distorted signal (red). For small normalized amplitudes the phase differences for the distorted (red) and reference signal (green) are recognizably bigger than for the pre-distorted signal (blue). This effect can also be explained by the memory effects, since the random distortions caused by them lead to bigger phase differences when the amplitude of the phasor is small.
In addition to the AM/AM and AM/PM characteristics also the power spectral density of the output signals and the constellation diagrams are evaluated, since these forms of representation reveal objective parameters for performance evaluation of the DPD system. The power spectral density of the distorted (blue), the pre-distorted (red) and the reference signals is shown in Figure 4.2a. The ACPR for the pre-distorted and reference signals is calculated based on a channel bandwidth of 30 MHz and a channel offset of 35 MHz relative to the middle channel at 0 Hz. Since the Matlab spectrum analyzer tool requires the sampling rate as an input parameter, a sampling frequency of 3.333 GHz was assumed. Compared to the reference signal the ACPR of pre-distoreted signal



(a) Power spectral density.

(b) Constellation diagram.

Figure 4.2.: Power spectral density (Welch estimate) for distorted signal (blue), pre-distorted signal (red) and reference signal (yellow). Constellation diagram for reference signal (blue) and pre-distorted output signal (red).

was reduced of up to 4.85 dB. The numerical values for the ACPR are listed in Table 4.3. In

	ACPR [dBc]	
Signal	**lower** (−35 MHz)	**upper** (+35 MHz)
refernce	−37.02	−38.94
pre-distorted	−41.87	−43.07

Table 4.3.: ACPR for simulated reference and pre-distorted signals

the constellation diagram shown in Figure 4.2b slightly more noise can be seen on the individual constellation points of the reference signal compared to the pre-distorted signal. This observation is is confirmed by the EVM measurements listed in Table 4.4. These simulation results are compatible

	EVM [%]	
Signal	**RMS**	**peak**
refernce	2.8	9.9
pre-distorted	0.7	1.9

Table 4.4.: EVM for simulated reference and pre-distorted signals.

with the DPD performance presented in [19] but in general the simulation results are not as good as the results presented in the referenced paper. This might be related to the SNR value of 40 dB used for the simulation, which might be less compared to the SNR of the signals in the paper. The SNR is not stated in the referenced paper.

5. Experiments

Several experiments have been performed on different hardware setups and with various signal parameters in order to examine the real-world performance of the open loop DPD algorithm. In this chapter a selection of these experiments is presented. For an extensive list of results see section B.

5.1. EIVE AWG and Mini-Circuits LNA

5.1.1. Setup

The institute of robust power semiconductor systems at the university of Stuttgart is currently developing a payload component for a cubesat mission, which is demonstrating an exploratory in-orbit verification of an E/W-Band satellite communication link. This project is called EIVE and as a part of the payload component an arbitrary waveform generator (AWG) based on a Xilinx Zynq UltraScale+ SoC board has been developed. This AWG consists of a 12-bit dual core DAC, which is operating at a sampling rate of 3.333 GHz and corresponding reconstruction filters. This AWG is used as a signal source for the this experiment. The baseband signal itself is constructed from a uniformly distributed random source followed by a 16-QAM modulator and a root raised-cosine pulse shaping filter with rolloff factor of $r = 0.35$. For this experiment three different signal bandwidths 33 MHz, 66 MHz and 133 MHz are generated. The signal is then up converted in software to a carrier frequency of $F_c = 1$ GHz and outputted on one of the DAC cores. The DUT in this experiment is the Mini-Cicuits ZX60-3018G-S+ low noise amplifier with an operating frequency range from 20 MHz to 3 GHz, a gain of 21.9 dB, a noise figure of 2.7 dB and a $P_{1\mathrm{dB}} = 12.76$ dBm at 1 GHz. For data acquisition a Keysight MSOS804A oscilloscope with 20 GSPS, 8.4 GHz input bandwidth and a 10-bit ADC is used. A 20 *dB* attenuator is connected at the output of the DUT to keep its power inside the specified range of the oscilloscope. The signal alignment, parameter extraction and pre-distortion steps of the DPD algorithm are performed in software on a computer. The setup is shown in Figure 5.1. An obvious disadvantage of this setup is the absence of any

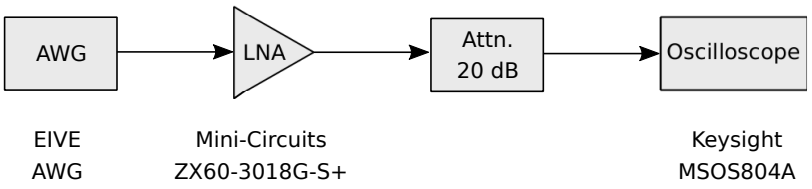
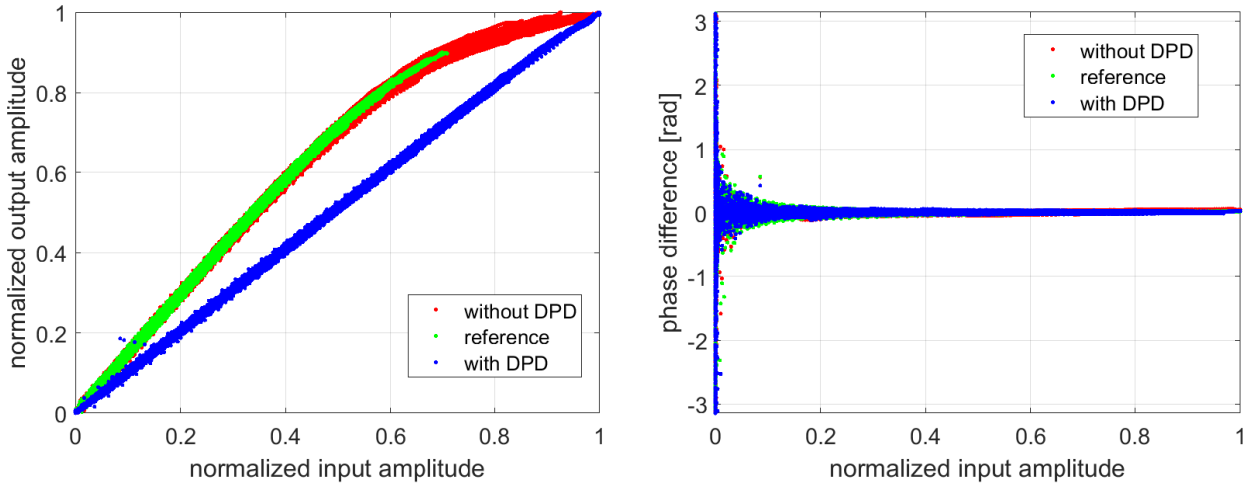


Figure 5.1.: Setup consisting of EIVE AWG, Mini-Circuits LNA and Keysight oscilloscope.

sort of oscillator synchronization between the AWG and the oscilloscope. This means that an extra effort in the signal alignment has to be made in order to compensate for the carrier frequency offset, which is caused by the two free-running local oscillators.

5.1.2. Results

In order to not overload this thesis with figures, only the plots for the signal with 33 MHz bandwidth are shown here. The rest of the measurement results are given in tabular form. See subsection B.1 for the remaining pots. The AM/AM and AM/PM characteristics are shown in Figure 5.2. The

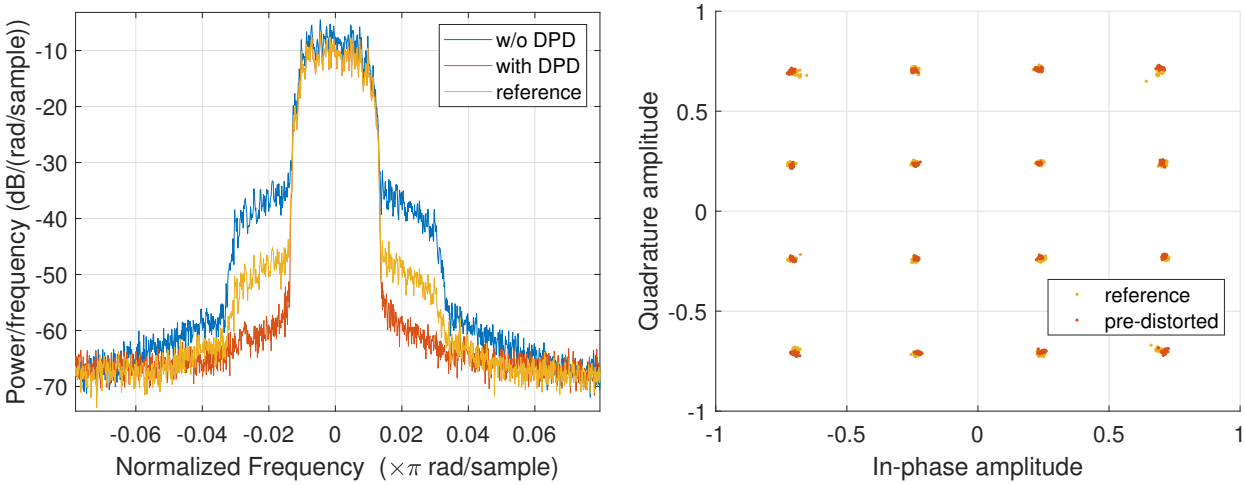


(a) AM/AM characteristics.

(b) AM/PM characteristics.

Figure 5.2.: AM/AM and AM/PM characteristics of nonlinearized RF PA (red), pre-distorted signal (blue) and reference signal (green).

hyperparameters of the post-inverse respectively DPD model have been determined by a manual grid search. This lead to a nonlinearity order of $P = 15$ and a memory depth of $M = 2$ for which optimum EVM results were achieved. In contrast to the rather strong memory effects, which were present in the simulation, the AM/AM and AM/PM characteristics of the distorted signal (red) show only moderate broadening. In combination with the small memory depth of the DPD model this indicates that only moderate memory effects are present. However, the linearization performance of the DPD system – even compared to the reference signal – is excellent as can be seen in the power spectral density plot in Figure 5.3. This is also reflected by the ACPR reduction of a 12.94 dB for the lower and 11.58 dB for the upper adjacent channel. Also the EVM is reduced by roughly 1 % down to 0.52 %, which is even lower than the EVM of the reference signal at the input of the DUT (0.88 %). This is possible, since the DPD does not only compensate for the DUT but the complete signal chain including the DAC analog front end. A list of results for the power, EVM and ACPR measurements for all signal bandwidths is given in Table 5.1. Note that the channel spacing for the ACPR measurement is based on the UMTS standard, which specifies a channel bandwidth of 3.84 MHz and a channel offset of 5 MHz for the first lower and upper adjacent channels [5]. This



(a) Power spectral density.

(b) Constellation diagram.

Figure 5.3.: Power spectral density (Welch estimate) for distorted signal (blue), pre-distorted signal (red) and reference signal (yellow). Constellation diagram for reference signal (blue) and pre-distorted output signal (red).

channel spacing has been scaled according to the signal bandwidth. The exact numerical values are given in the annex. As can be seen in Table 5.1 the performance of the DPD system in terms of EVM is rather constant at around 1 % improvement compared to the reference signal for all tested signal bandwidths. But on the other hand the gain for the ACPR is decreasing with the increase in signal bandwidth with roughly 3 dB for each doubling in bandwidth. This is due to the fact, that the ACPR is measured in dBc and is therefore directly related to the maximum of the signal power spectrum. As the band power is kept roughly constant, the maximum of the signal power spectrum has to decrease with the same factor as the signal bandwidth increases.

5.2. Keysight AWG and Mini-Circuits LNA

5.2.1. Setup

This experiment was performed using a high-performance AWG built by Keysight (model M8195A), which offers a sampling rate of upto 64 GSPS, an analog bandwidth of 25 GHz and four channels. As in Figure 5.2.2 the baseband signal is constructed from a uniformly distributed random source followed by a 16- respectively 64-QAM modulator and a root raised-cosine pulse shaping filter with rolloff factor of $r = 0.35$. For this experiment again three different signal bandwidths 160 MHz, 320 MHz and 640 MHz are generated. The signal is then up converted in software to a carrier frequency of $F_c = 2$ GHz and outputted on one of the four AWG channels. The DUT in this experiment is the same Mini-Cicuits ZX60-3018G-S+ as in Figure 5.2.2. Also for data acquisition the Keysight MSOS804A oscilloscope is used. A 20 dB attenuator is connected at the output of the DUT to keep its power inside the specified range of the oscilloscope. The signal alignment,

Signal	band power [dBm] (input)	band power [dBm] (output)	EVM (input)	EVM (output)	ACPR [dBc] lower	ACPR [dBc] upper
		16-QAM, bandwidth 33 MHz				
distorted	−9.61	10.33	0.86 %	4.32 %	−28.67	−27.18
reference	−12.60	7.38	0.88 %	1.51 %	−37.01	−36.52
pre-distorted	−12.62	7.31	-	0.52 %	−49.95	−48.10
		16-QAM, bandwidth 66 MHz				
distorted	−9.63	9.63	1.55 %	4.33 %	−30.70	−30.28
reference	−12.73	7.29	1.58 %	1.89 %	−40.07	−39.85
pre-distorted	−12.63	7.29	-	1.00 %	−46.51	−46.97
		16-QAM, bandwidth 133 MHz				
distorted	−9.42	9.79	1.97 %	4.77 %	−28.88	−29.37
reference	−12.48	7.48	2.01 %	2.48 %	−38.52	−39.02
pre-distorted	−12.37	7.54	-	1.26 %	−43.36	−43.64

Table 5.1.: EIVE AWG and Mini-Circuits LNA: Measurement results for signal with various bandwidths.

parameter extraction and pre-distortion steps of the DPD algorithm are performed in software on a computer. The setup is shown in Figure 5.4. In contrast to the setup in Figure 5.2.2 this setup con-

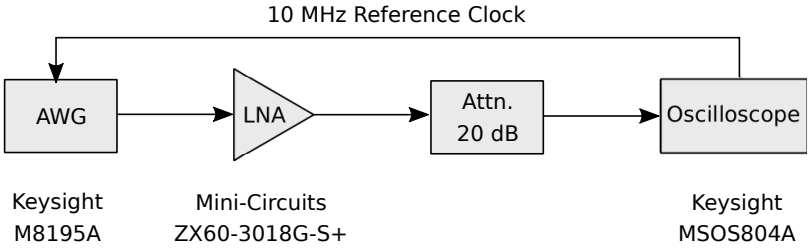
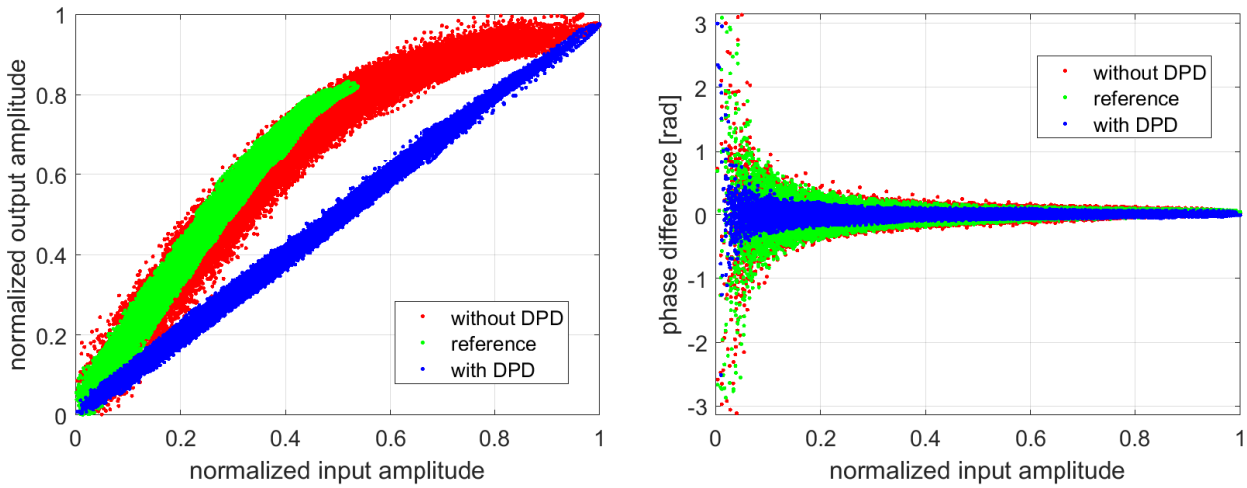


Figure 5.4.: Setup consisting of Keysight AWG, Mini-Circuits LNA and Keysight oscilloscope.

tains a 10 MHz reference clock, which is shared between the AWG and the oscilloscope. It ensures coherence between the signal generation and data acquisition steps. This makes the compensation for the carrier frequency offset obsolete. The nearly 20 times higher sampling rate compared to the EIVE AWG also improves the signal quality. This improvement manifests in higher SNR and lower EVM values at the output of the AWG.

5.2.2. Results

Again, to not overload this thesis with figures, only the plots for the 16-QAM signal with 640 MHz bandwidth are shown here. The rest of the measurement results are given in tabular form. See subsection B.2 for the remaining plots. The AM/AM and AM/PM characteristics are shown in Figure 5.5. In comparison to the AM/AM and AM/PM characteristics in Figure 5.2 the width of the traces for the distorted and reference signal are significantly wider than that of the pre-distorted signal. The increased bandwidth leads to a more severe memory effect, which becomes also evident

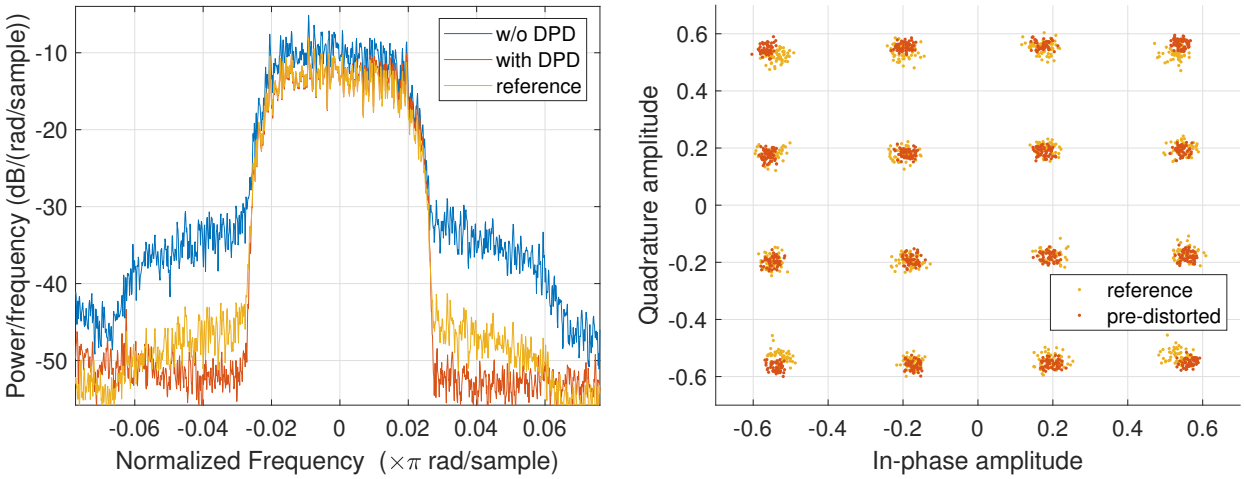


(a) AM/AM characteristics.

(b) AM/PM characteristics.

Figure 5.5.: AM/AM and AM/PM characteristics of nonlinearized RF PA (red), pre-distorted signal (blue) and reference signal (green).

when the DPD model parameters are considered. In this experiment a nonlinearity order of $P = 11$ and a memory depth of $M = 15$ have been applied to get optimum results. The power spectral density and constellation diagram are shown in Figure 5.6. The ACPR of the pre-distorted signal could be reduced by upto 5.13 dB compared to the reference signal and also the EVM has been improved by 2.25 %. The improved EVM leads to more closely grouped constellation points for the pre-distorted signal, which is also visible directly in the constellation diagram. For the ACPR measurements in this experiment the same channel spacing definition was used as for the experiment described in . The remaining measurement results are listed in Table 5.2. It can be seen, that the signal power of the 64-QAM signals is less than that of the 16-QAM signal. This is related to the tighter signal integrity requirements of the 64-QAM constellation, in which the individual points are placed closer together. As a result the distortions introduced by the LNA have to be limited in order to obtain reasonable constellations respectively EVM values. The reduction in signal power and consequently distortions also means, that the nonlinearity of the LNA is less pronounced. As a result, after linearization the signal power is closer to that of the distorted signal. The AM/AM characteristics provide a good visual explanation: As can be seen in Figure B.7 the deviation of the distorted from the linearized curve is less severe for the 64-QAM signal compared to the 16-QAM signal. The ACPR and EVM values of the remaining measurements listed in Table 5.2 are comparable to that of the 16-QAM signal with bandwidth of 640 MHz, which was presented in more detail. The EVM values improved in all measurements by more than 1.26% and the ACPR by more than 3.39 dB compared to the corresponding reference signals.



(a) Power spectral density. **(b)** Constellation diagram.

Figure 5.6.: Power spectral density (Welch estimate) for distorted signal (blue), pre-distorted signal (red) and reference signal (yellow). Constellation diagram for reference signal (blue) and pre-distorted output signal (red).

5.3. Keysight AWG and EIVE E-band Hardware

5.3.1. Setup

In order to increase the linearization bandwidth even further, a third setup consisting of the EIVE analog front end hardware including IQ modulator, medium power amplifier (MPA) and high power amplifier (HPA) and other support equipment operating in the waveguide E band frequency range (60 − 90 GHz) has been used. The block diagram of this setup is shown in Figure 5.7. The

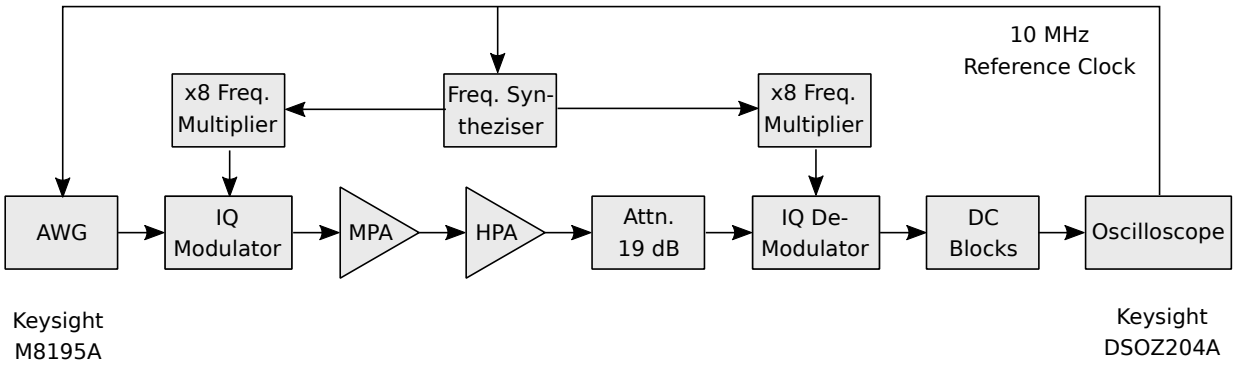


Figure 5.7.: Setup consisting of Keysight AWG, EIVE hardware and Keysight oscilloscope.

setup utilizes the direct conversion architecture. The signal source is again the Keysight M8195A AWG including a built-in equalizer, which compensates for the frequency response of the coax cables on the AWGs outputs. It is connected to the EIVE IQ modulator, which performs the symbol mapping and up conversion to the E band. Since the main amplification is performed at high frequencies and with large bandwidths, a cascade of two amplifiers - EIVE MPA and HPA - is used.

Signal	band power [dBm] (input)	band power [dBm] (output)	EVM (input)	EVM (output)	ACPR [dBc] lower	ACPR [dBc] upper
		16-QAM, bandwidth 160 MHz				
distorted	−5.40	11.11	0.88 %	7.16 %	−24.95	−24.73
reference	−9.60	8.55	0.75 %	2.99 %	−32.87	−32.72
pre-distorted	−9.58	8.55	-	1.30 %	−42.17	−40.31
		16-QAM, bandwidth 320 MHz				
distorted	−5.42	11.06	1.23 %	7.21 %	−25.09	−25.20
reference	−10.06	8.19	1.01 %	2.99 %	−34.28	−34.96
pre-distorted	−9.94	8.21	-	1.43 %	−41.96	−43.59
		16-QAM, bandwidth 640 MHz				
distorted	−5.41	11.13	1.69 %	7.80 %	−25.36	−25.17
reference	−10.26	8.16	1.41 %	3.98 %	−34.84	−35.00
pre-distorted	−9.83	8.32	-	1.73 %	−38.62	−40.13
		64-QAM, bandwidth 160 MHz				
distorted	−9.56	8.48	0.73 %	3.19 %	−33.50	−32.69
reference	−11.54	6.94	0.88 %	1.96 %	−38.76	−38.30
pre-distorted	−11.50	6.94	-	0.71 %	−45.40	−45.53
		64-QAM, bandwidth 320 MHz				
distorted	−9.61	8.49	0.95 %	3.41 %	−31.38	−31.75
reference	−11.98	6.65	1.15 %	2.25 %	−36.78	−37.84
pre-distorted	−11.79	6.64	-	0.99 %	−42.61	−43.42
		64-QAM, bandwidth 640 MHz				
distorted	−9.62	8.52	1.29 %	4.12 %	−31.84	−32.26
reference	−11.22	7.27	1.29 %	3.42 %	−35.39	−35.98
pre-distorted	−11.11	7.26	-	1.60 %	−38.78	−40.00

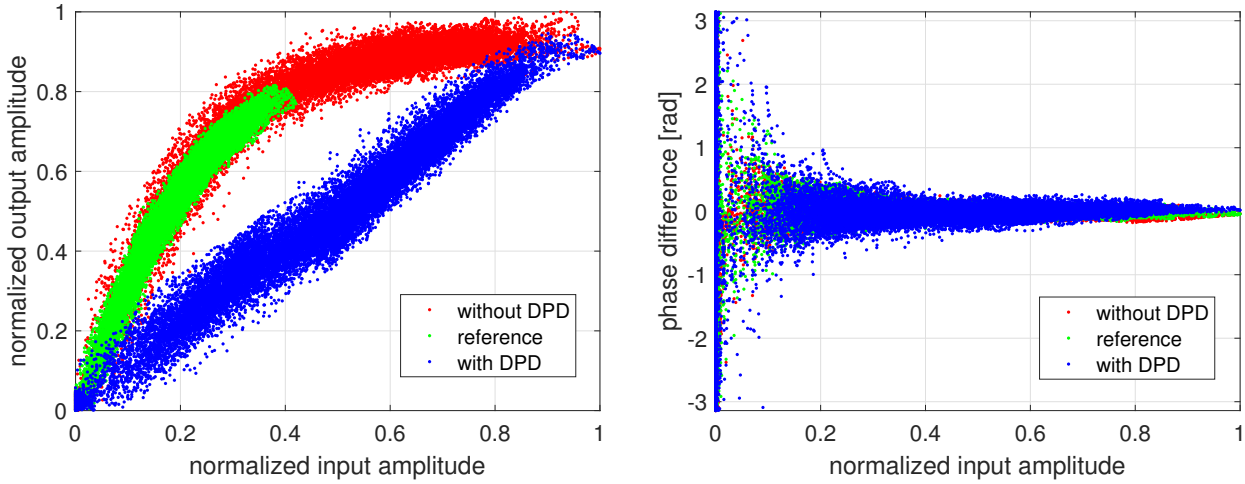
Table 5.2.: Keysight AWG and Mini-Circuits LNA: Measurement results for signal with various bandwidths and constellations.

Note, that due to signal integrity issues not for all experiments both amplifiers were included in the signal chain. After amplification the passband signal is converted back down to baseband using an IQ demodulator, whose input is protected from being driven into saturation by a 30 dB attenuator. The LO signal for up and down conversion is generated by a frequency synthesizer and two distinct x8 frequency multipliers. The frequency of the LO signal is 9.125 GHz respectively 73 GHz. As described in subsubsection 2.1.1.2 this setup also suffers from problems regarding LO leakage and DC offset. In order to compensate for these issues, DC blocking filters are placed at the input of the Keysight DSOZ204A oscilloscope. A 10 MHz reference clock signal is used for synchronization of the AWG, frequency synthezier and oscilloscope. Due to the use of two distinct x8 frequncy multipliers, this does not create coherence in the complete setup but at least limits the carrier frequency offset between up and down conversion.
The complex baseband signal is again constructed from a uniformly distributed random source followed by a 16-QAM modulator and a root raised-cosine pulse shaping filter with rolloff factor of $r = 0.35$. Unfortunately the frequency response of the DC blocking filters is not perfect and

therefore not only the LO leakage but also parts of the baseband signal are removed. This leads to problems in the DPD algorithm and as a consequence the complex baseband signal is shifted in the digital domain to different carrier frequencies ranging from 1 GHz to 5 GHz. However, the complex output of the AWG is still used in order to raise the signal output power.

5.3.2. Results

The targeted linearization bandwidth in this thesis is 5 GHz but the results of several experiments with a bandwidth of upto 4.6 GHz are not satisfying. For that reason the experiment with 941 MHz bandwidth is described here in more detail. The AM/AM and AM/PM characteristics are shown in Figure 5.8. The results for higher bandwidths are given in the annex and in a summarized form in Table 5.3. Although the linearization of the DPD algorithm seems to be working fine as can

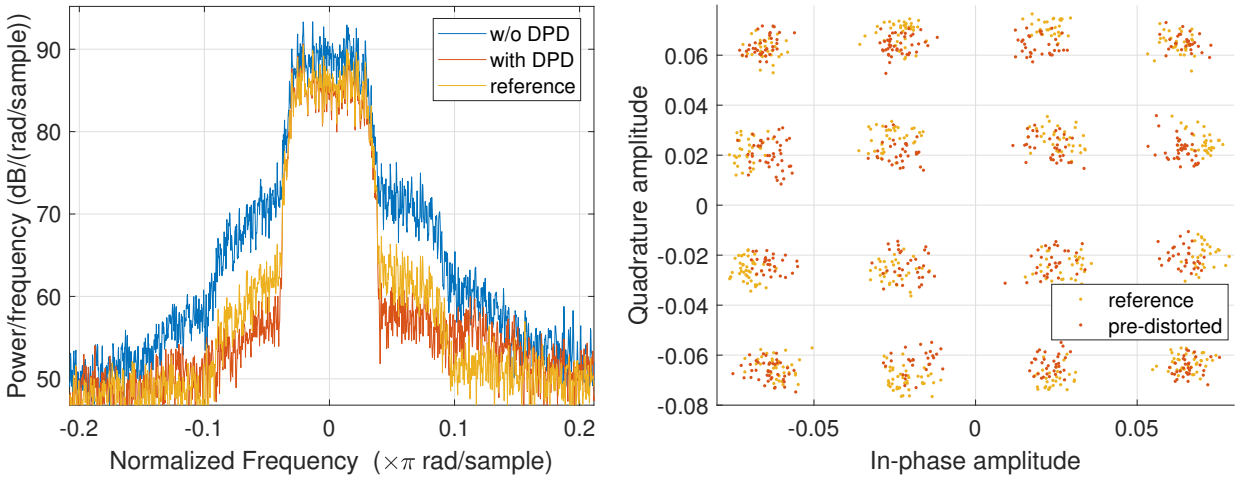


(a) AM/AM characteristics.

(b) AM/PM characteristics.

Figure 5.8.: AM/AM and AM/PM characteristics of nonlinearized RF PA (red), pre-distorted signal (blue) and reference signal (green).

be seen in the AM/AM characteristic, the compensation of the memory effects is obviously not as good as for the previous experiments with the Mini-Ciruits LNA. This can be assessed by the width of the pre-distorted (blue) curves in the AM/AM and AM/PM characteristics, which are nearly as wide as the curves for the distorted (red) signal and the rather small EVM improvements of just 0.8 % in comparison to the reference signal with an EVM of 7.8 %. The power spectral density and constellation diagram are shown in Figure 5.9. The ACPR could be improved by 3.78 dB for the lower and 3.73 dB for the upper adjacent channel. An important fact to note in the power spectral density is the asymmetry in the spectrum. For higher frequencies (beginning at 0.1π rad) the spectral power density of the pre-distorted signal is higher than that of the reference signal. For this experiment a DPD model with nonlinearity order $P = 5$ and a memory depth of $M = 2$ have been used for obtaining optimum results. Considering the high signal bandwidth in this experiment and the higher memory depths used for the previous experiments with lower bandwidths the short



(a) Power spectral density. **(b)** Constellation diagram.

Figure 5.9.: Power spectral density (Welch estimate) for distorted signal (blue), pre-distorted signal (red) and reference signal (yellow). Constellation diagram for reference signal (blue) and pre-distorted output signal (red).

memory depth of 2 appears to be not optimal but higher memory depths resulted in worse EVM and ACPR improvements. The measurement results for the other signal bandwidths and hardware configurations are listed in Table 5.3.

Signal	**band power [dBm]** (output)	**EVM** (output)	**ACPR [dBc]** lower	upper
	16-QAM, bandwidth 941 MHz, **TX-MPA-HPA**			
distorted	−5.23	12.9 %	−21.24	−19.04
reference	−8.33	7.8 %	−27.37	−25.24
pre-distorted	−8.73	7.0 %	−31.15	−28.97
	16-QAM, bandwidth 4600 MHz, **TX-MPA**			
distorted	−0.94	12.8 %	−29.71	−24.02
reference	−3.64	12.3 %	−32.00	−27.59
pre-distorted	−3.44	9.3 %	−30.24	−25.84
	16-QAM, bandwidth 4600 MHz, **TX**			
distorted	−4.28	6.3 %	−25.76	−28.57
reference	−3.54	5.9 %	−26.72	−28.46
pre-distorted	−5.37	5.9 %	−26.48	−28.73

Table 5.3.: Keysight AWG and EIVE E-band Hardware: Measurement results for signal with various bandwidths.

Unfortunately, for this experiment the input band power and input EVM values have not been recorded, because the setup did not allow it. The measurement with the 4600 MHz signal bandwidth and TX (IQ modulator) and MPA does show an improvement in EVM of 3 % for the pre-distorted signal but the ACPR for this measurement is around 2 dB worse compared to the reference signal. This might be caused by the slightly lower band power of the reference signal, which is 0.2 dB less than that of the pre-distorted signal. Nevertheless, the AM/AM characteristic in Figure B.8 reveals, that due to the low input band power practically no gain compression is present in that experiment. For the 4600 MHz signal bandwidth and only the TX no gain compression and no improvements by the DPD are recognizable.

6. Evaluation

The measurement results presented in chapter 5 show that the open-loop DPD algorithm can be successfully applied to signals with much higher carrier frequency (waveguide E-band) and larger bandwidths than for example the 3.84 MHz bandwidth of a standard WCDMA signal as presented in [19]. On the other hand these experiments also revealed that for larger bandwidths and carrier frequencies the problems regarding signal integrity mentioned in subsection 2.1.1 especially the LO leakage as well as noise, which is not included in the signal model, become significant and can deteriorate the DPD performance.

6.1. LO Leakage and IQ Imbalance

Since LO leakage and IQ imbalance are inherent problems of the IQ modulator, it makes sense to study these effects with a setup that includes only the IQ modulator and no other RF components as the MPA or HPA, which could cause additional distortions in the signal. Such a setup has been used for generating the measurement results given in Figure B.9 and Table B.9. It only included the TX, attenuator and RX modules in the signal chain. The attenuator ensures that the RX module is not over driven and does not cause any significant distortions during the down conversion. As a result, all the distortions in the RF signal can be solely attributed to the TX module. As mentioned in the section 5.3 the measurement results have been obtained for a 16-QAM signal with 4600 MHz bandwidth, a IF frequency of 5 GHz and a carrier frequency of 70 GHz, so that the RF signal is located at a center frequency of 75 GHz. This frequency shift ensures that the LO does not leak directly into the signal band. This becomes obvious when the PSD of the complex baseband signal, which was captured by the oscilloscope with a sampling frequency of 40 GHz, in Figure 6.1 is considered. Although DC blocking filters have been present at the input of the oscilloscope, a significant DC signal component is visible in the PSD. In addition there are also HM distortions of the LO signal present in the spectrum at a normalized frequency of ± 0.4375, which is equivalent to a frequency offset of ± 8.75 GHz. Note, that in this experiment the LO signal originates from the frequency synthesizer signal and its frequency is multiplied by a factor of 8 to get a a final LO frequency of 70 GHz. In the interval between -0.4 and -0.1 of the normalized frequency more distortions are present in the PSD. These distortions appear in a frequency band, which is symmetric to frequency band of the actual signal and therefore they might be related to IQ imbalance issues. Due to the fact, that the distortions do not have the same spectral shape as the original signal (just mirrored at the y-axis), the magnitude of the IQ imbalance is not constant over the complete frequency band.

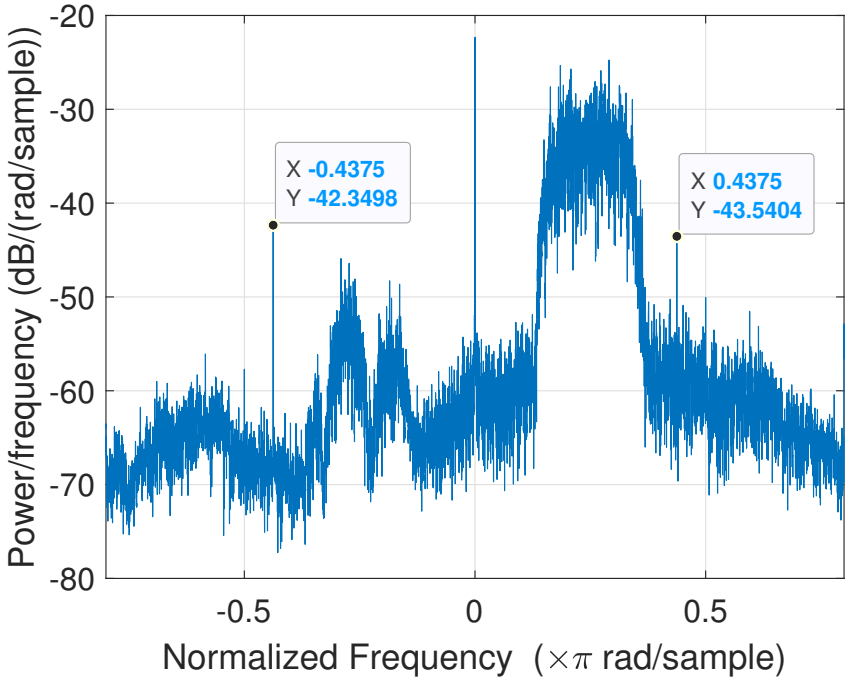


Figure 6.1.: PSD of 16-QAM signal with 4600 MHz bandwidth and an IF of 5 GHz passed through EIVE TX module at carrier frequency of 70 GHz.

The LO leakage and LO HM distortions (at 0.4375) might be filtered out with a sufficiently narrow band filter but as they fall into the same frequency band as the third order IM distortions, this filtering will also have an impact on the linearization performance of the DPD algorithm. The IQ imbalance distortions between -0.4 and -0.1 might be filtered out by shifting the center of the distortions to 0 and high-pass filter this signal. But this does not undo the effects of the IQ imbalance caused on the original signal. So in order to improve the linearization performance of the DPD algorithm, it is crucial to minimize LO leakage and IQ imbalance issues by careful IQ modulator design or utilizing other compensation techniques such as LO leakage cancellation, which is based on the principle of destructive interference for the leaked LO signal. Even if a small IF frequency is used for the baseband signal.

6.2. Noise Impact on DPD Performance

The impact of white additive Gaussian noise on the parameter estimation for an extended envelope memory polynomial DPD model has already been studied in [1]. The results from this paper can directly be applied to the parameter extraction implemented in this thesis, since both approaches use LS estimation and from the estimation perspective the used DPD model does not matter as long as the model is linear in its parameters.
Due to the fact, that the noise distribution in the RF PA output signal is unknown, a simple LS estimator is used in the parameter extraction step of the DPD algorithm. As described in

subsection 3.2.3 the LS estimator does not take into account any noise. It just minimizes the cost function

$$J(\mathbf{\Theta}) = \sum_{n=M+1}^{L} |e(n)|^2 = \sum_{n=M+1}^{L} |\tilde{x}(n) - \tilde{z}(n)|^2 = \|\mathbf{x} - \mathbf{Y}\mathbf{\Theta}\|_2^2, \tag{6.1}$$

where $\mathbf{x}$ is the desired output vector and $\mathbf{Y}\mathbf{\Theta}$ is the actual vector signal of the post-inverse. This minimization and its closed-form solution is given as

$$\hat{\mathbf{\Theta}} = arg \min_{\mathbf{\Theta}} J(\mathbf{\Theta}) = (\mathbf{Y}^H\mathbf{Y})^{-1}\mathbf{Y}^H\mathbf{x} \tag{6.2}$$

Of course the assumption, that the output signal of a real-world RF PA is free of noise is not correct. In order to understand the impact of the unmodeled noise on the DPD performance, additive white Gaussian noise with zero mean and σ^2 variance is added to the RF PAs output signal $\mathbf{Y}$. As a result the regression matrix does contain an additional noise part

$$\mathbf{Y} = \mathbf{Y}_0 + \mathbf{Y}_{\text{AWGN}}, \tag{6.3}$$

which is given by $\mathbf{Y}_{\text{AWGN}}$ and the original signal is $\mathbf{Y}_0$. The solution for the minimization of the cost function is now given as

$$\hat{\mathbf{\Theta}} = ((\mathbf{Y}_0 + \mathbf{Y}_{\text{AWGN}})^H(\mathbf{Y}_0 + \mathbf{Y}_{\text{AWGN}}))^{-1}(\mathbf{Y}_0 + \mathbf{Y}_{\text{AWGN}})^H\mathbf{x}. \tag{6.4}$$

This expression does contain the cross covariance matrices $\mathbf{Y}_0^H\mathbf{Y}_{\text{AWGN}}$ and $\mathbf{Y}_{\text{AWGN}}^H\mathbf{Y}_0$ but since the noise and the input and output signals are uncorrelated, the Frobenius norm of the covariance matrices do tend to zero for a sufficiently large number of samples and therefore these matrices can be ignored. This finding can be used to simplify the result to

$$\hat{\mathbf{\Theta}} = (\mathbf{Y}_0^H\mathbf{Y}_0 + \mathbf{Y}_{\text{AWGN}}^H\mathbf{Y}_{\text{AWGN}})^{-1}(\mathbf{Y}_0 + \mathbf{Y}_{\text{AWGN}})^H\mathbf{x}. \tag{6.5}$$

Finally, the well-known matrix inversion lemma can be used to bring it into the following form

$$\hat{\mathbf{\Theta}} = (\mathbf{Y}_0^H\mathbf{Y}_0)^{-1}\mathbf{Y}_0^H\mathbf{x} + \mathbf{B} \tag{6.6}$$

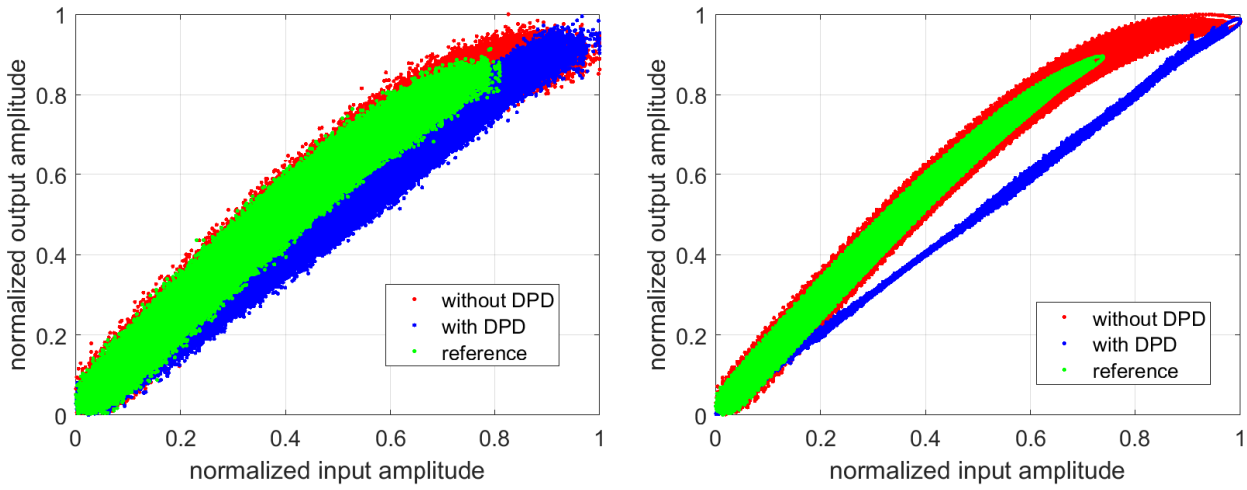
with

$$\mathbf{B} = \mathbf{D}\mathbf{x} - \left[\mathbf{C}(\mathbf{I} + \mathbf{C})^{-1}(\mathbf{Y}_0^H\mathbf{Y}_0)^{-1}\right](\mathbf{Y}_0^H + \mathbf{Y}_{\text{AWGN}}^H)\mathbf{x} \tag{6.7}$$

and

$$\begin{aligned} \mathbf{D} &= (\mathbf{Y}_0^H\mathbf{Y}_0)^{-1}\mathbf{Y}_{\text{AWGN}}^H \\ \mathbf{C} &= (\mathbf{Y}_0^H\mathbf{Y}_0)^{-1}(\mathbf{Y}_{\text{AWGN}}^H\mathbf{Y}_{\text{AWGN}}). \end{aligned} \tag{6.8}$$

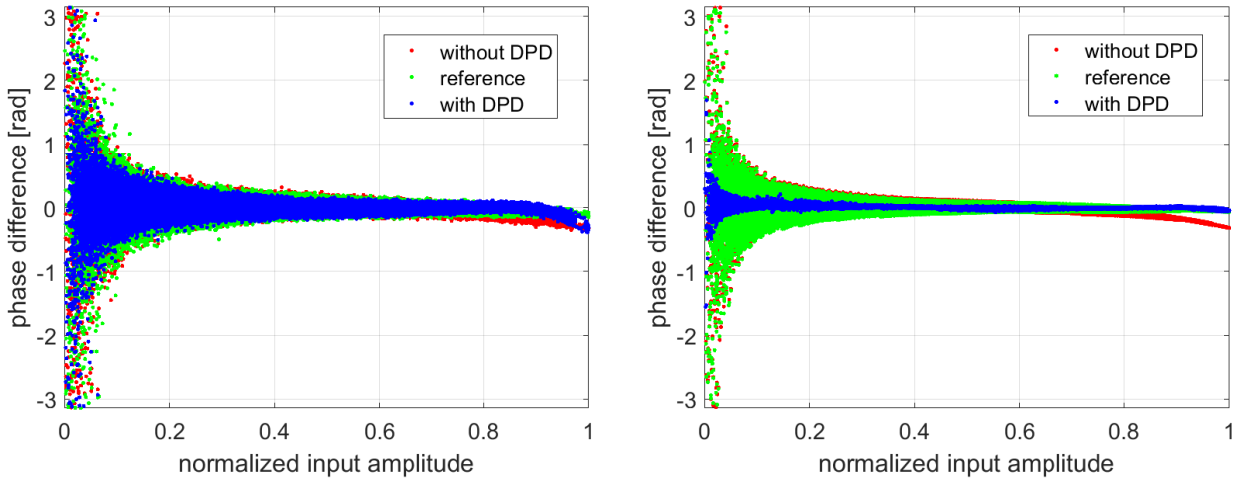
Equation 6.6 shows that the solution of the LS parameter estimation for the noisy RF PA output signal does contain the original solution for the noiseless case as given in Equation 6.4 plus an additional bias term **B**, which depends on the noise level. This bias term is an unwanted distortion in the DPD model parameters and deteriorates the performance of the DPD.
To illustrate the negative influence of additive white Gaussian noise on the performance of the DPD, the simulation described in chapter 4 is repeated with two different SNR level of 25 dB and 60 dB. The simulation results for both SNR values are plotted side by side to allow an easy comparison. The noise impact is especially noticeable in the AM/AM and AM/PM characteristics. The AM/AM characteristics for the signal with 25 dB SNR shown in Figure 6.2a do have a significantly higher variance as the AM/AM characteristics for the signal with 60 dB SNR shown in Figure 6.2b. This is true for all three signals - the distorted (red), the pre-distorted (blue) and the reference signal - but in particular the difference between the two pre-distorted signals is highly distinctive. This means, that the compensation of the memory effects is strongly influenced by the SNR. If the unmodeled noise represents a significant part of distortion on the RF PAs output signal, the DPD is not able to distinguish between distortions caused by memory effects or the unmodeled noise. This leads to worse DPD performance. However, Figure 6.2 reveals, that the linearization performance of the DPD seems to be less effected by unmodeled noise. The same observation can be made in the



(a) AM/AM characteristics (SNR 25 dB). **(b)** AM/AM characteristics (SNR 60 dB).

Figure 6.2.: Simulated AM/AM characteristics for two noisy RF PA output signals.

AM/PM characteristics shown in Figure 6.3. The AM/PM characteristic of the pre-distorted signal shows a slight deviation from the desired 0 rad for larger normalized input amplitudes (> 0.9). This might be a negative of the unmodeled noise on the linearization performance of the DPD.



(a) AM/PM characteristics (SNR 25 dB). **(b)** AM/PM characteristics (SNR 60 dB).

Figure 6.3.: Simulated AM/PM characteristics for two noisy RF PA output signals.

Since constellation diagram and PSD do not show such distinct differences between the two different SNR scenarios, these plots are not shown. Instead the ACPR and EVM results are listed in Table 6.1 and Table 6.2.

	ACPR [dBc]	
Signal	**lower** (−35 MHz)	**upper** (+35 MHz)
	16-QAM, SNR 25 dB	
refernce	−34.62	−36.31
pre-distorted	−40.83	−42.35
	16-QAM, SNR 60 dB	
refernce	−37.03	−38.96
pre-distorted	−43.04	−44.11

Table 6.1.: ACPR for simulated reference and pre-distorted signals with different SNR values (16 QAM, bandwidth 30 MHz).

The ACPR improvements for the 25 dB signal are about 3 dB worse compared to the 60 dB signal.

	EVM [%]	
Signal	**RMS**	**peak**
	16-QAM, SNR 25 dB	
refernce	3.2	11.1
pre-distorted	1.5	12.0
	16-QAM, SNR 60 dB	
refernce	2.9	9.0
pre-distorted	0.7	2.0

Table 6.2.: EVM for simulated reference and pre-distorted signals with different SNR values (16 QAM, bandwidth 30 MHz).

The differences in EVM improvements between the two SNR scenarios are also significant. In terms of RMS the EVM improvement is 0.5% better for the higher SNR and the peak value of the EVM improvements differs even more. Notice, that for the signal with 25 dB the peak EVM is worse compared to the reference signal.

7. Conclusion and Outlook

This thesis showed that the DPD algorithm described in [19] can be applied successfully to transmitters operating at E band carrier frequencies (e.g. 75 GHz) and with high bandwidths of close to 1 GHz. Especially when the IQ modulation and up/down conversion process is performed in the digital domain, the improvements in terms of EVM and ACPR are significant and enable the transmitter to operate at a higher power level than without DPD while still maintaining linearity in its transfer characteristic. The strong LO leakage exhibited by the EIVE E-band hardware made it necessary to shift the baseband signal to a small IF. As a result the hardware, which was designed to be used in a direct conversion architecture, was actually used in a mode of operation for which it was not designed. This might be a possible explanation for the IQ imbalance issues described in section 6.1. The LO leakage and IQ imbalance issues in combination with unmodeled noise lead to a low SNR on the RF PAs output signal, which deteriorated the DPD performance.
Due to the bandwidth expansion caused by the odd-order IM distortions and the resulting increase in input and output bandwidth of the DPD, the hardware requirements imposed by the DPD are demanding for signals with high bandwidths. The open loop architecture represents a solution, which can be easily implemented without the need for any additional DSP hardware such as FPGAs. Only standard measurement equipment is necessary, which is usually available in every well equipped laboratory. Therefore it is a good choice for evaluating the potential performance improvements gained by the integration of a DPD into an already existing system. Since the open loop approach is not adaptive and requires the DPD parameter extraction to be repeated when ever the external conditions (temperature, supply voltage, ...) or the signal statistics change significantly, it is certainly not suitable for use in an actual base station transmitter or similar devices.
In the scope of this thesis signal integrity was identified as the main problem for the DPD when applied on the EIVE E band hardware. Therefore future work should be focused on eliminating the LO leakage, IQ imbalance and noise problems in order to reduce distortions, which are not modeled by the DPD respectively post-inverse model and prevent successful application of the DPD on signals with high bandwidths. The LO leakage may be reduced by careful redesign of the IQ modulator and trying to increase the isolation between the LO and RF port. In addition a technique called LO leakage cancellation, which is based on the concept of destructive interference and uses a separate observation receiver connected directly to the transmitters output for determining the phase and amplitude of the leaked LO signal, could be used. A DC offset with opposite phase and same amplitude as the leaked LO signal is added to the complex baseband signal and destructively interferes with the leaked LO signal at the RF output of the IQ modulator. For more details on LO leakage cancellation see [4]. Even with careful design of the analog front end an

actual IQ modulation system typically only provides an image rejection of 30 to 40 dB, which is not sufficient for high-quality communication systems [17]. As a consequence, the compensation of IQ imbalances in IQ modulation systems has been studied quite extensively in literature. Today a lot of different compensation approaches exist for frequency-dependent as well as frequency-independent gain and phase imbalances. For signals with high bandwidths frequency-dependent compensation techniques should be applied. In [15] a pre-distortion technique for compensating frequency-dependent IQ imbalances caused by mismatches between the reconstruction filters (LPF), which is based on the LS estimation of FIR filter parameters using test sequences, is presented. Unfortunately, the mismatches caused by the IQ modulator are not considered in this paper. On the other hand, this technique is similar to the DPD algorithm described in [19] and may be included into the DPD algorithm implemented in the scope of this thesis. Adapting the IQ imbalance compensation to also include the imbalances introduced by the IQ modulator should be possible, because the DPD algorithm includes capturing the complex baseband signal by the oscilloscope and therefore all IQ imbalance effects are captured. Assuming that the noise in the RF PAs output signal is uncorrelated with the data carrying signal, coherent averaging over a sufficiently large number of signals may be used to lower the noise power and therefore increase the SNR, which will improve DPD performance as shown in section 6.2.
Since the presented DPD algorithm does not require any additional hardware (no observation path is needed) and can be pre-computed for a given set of signals and external conditions (temperature, supply voltage, etc.), it could be applied to the test signals of the EIVE cubesat mission, which is currently in development at the institute of robust power semiconductor systems at the university of Stuttgart. During thermal vacuum testing of the E band hardware the DPD algorithm could be used to create a set of pre-distorted test signals for specific temperatures (e.g. for every 1 $\circ$C of the operating temperature range a pre-distorted signal is generated). This set of signal could be used when the cubesat is in orbit to send a pre-distorted signal corresponding to the actual temperature of the E-band hardware. This could help to reduce the distortions introduced by the E-band PAs and increase the SNR of the test signal, which may be beneficial for achieving the mission goal of characterizing the atmospheric distortions on the E-band link.

Bibliography

[1] Amin, Shoaib ; Zenteno, Efrain ; Landin, Per ; Rönnow, Daniel ; Isaksson, M. ; Händel, Peter: Noise impact on the identification of digital predistorter parameters in the indirect learning architecture, 2012. – ISBN 978–1–4673–4761–7, S. 36–39

[2] Couch, Leon W.: *Digital and Analog Communication Systems*. 8th ed. Pearson Education, 2014

[3] Cripps, Steve: *Advanced Techniques in RF Power Amplifier Design*. Artech House, 2002

[4] Frizelle, Dave: *Transmit LO Leakage (LOL)-An Issue of Zero-IF That Isn't Making People Laugh Out Loud*. https://www.analog.com/en/analog-dialogue/articles/transmit-lo-leakage-lol-an-issue-of-zero-if-that-isn-t-making-people-laugh-out-loud.html

[5] Generation Partnership Project (3GPP) 3rd: *Universal Mobile Telecommunications System (UMTS); User Equipment (UE) radio transmission and reception (FDD) (3GPP TS 25.101 version 16.1.0 Release 16)*. https://www.etsi.org/deliver/etsi_ts/125100_125199/125101/16.01.00_60/ts_125101v160100p.pdf. Version: November 2020. – [Online; accessed 09.09.2021]

[6] Ghannouchi, Fadhel M. ; Hammi, Oualid: Behavioral modeling and predistortion. In: *IEEE Microwave Magazine* 10 (2009), Nr. 7, S. 52–64. http://dx.doi.org/10.1109/MMM.2009.934516. – DOI 10.1109/MMM.2009.934516

[7] Ghannouchi, Fadhel M. ; Hammi, Oualid ; Helaoui, Mohamed: *Behavioral Modeling and Predistortion of Wideband Wireless Transmitters*. John Wiley & Sons, 2015. – ISBN 978–1119–00443–1

[8] Gregorio, Fernando ; González, Gustavo ; Schmidt, Christian ; Cousseau, Juan: *Signal Processing Techniques for Power Efficient Wireless Communication Systems*. Springer, 2020

[9] Gu, Qizheng: *RF System Design of Transceivers for Wireless Communications*. Berlin Heidelberg : Springer Science & Business Media, 2006. – ISBN 978–0–387–24161–6

[10] Gupta, A. ; Jha, R. K.: A Survey of 5G Network: Architecture and Emerging Technologies. In: *IEEE Access* 3 (2015), S. 1206–1232. http://dx.doi.org/10.1109/ACCESS.2015.2461602. – DOI 10.1109/ACCESS.2015.2461602

[11] Li, Yue ; Wang, Xiaoyu ; Zhu, Anding: Sampling Rate Reduction for Digital Predistortion of Broadband RF Power Amplifiers. In: *IEEE Transactions on Microwave Theory and Techniques* 68 (2020), Nr. 3, S. 1054–1064. http://dx.doi.org/10.1109/TMTT.2019.2944813. – DOI 10.1109/TMTT.2019.2944813

[12] Rahkonen, Timo ; Vuolevi, Joel: *Distortion in RF Power Amplifiers*. Artech House, 2003

[13] Schetzen, M.: *The Volterra and Wiener Theories of Nonlinear Systems*. John Wiley & Sons, 1980

[14] Solutions, Skyworks ; Maleksadeh, Foad A. ; Sarbishaei, Hassan ; Xu, Jin-Biao ; Moreau, Jean-Marc: *Keysight Technologies and Skyworks Solutions Applying a Very Wide-Bandwidth Millimeter-Wave Testbed to Power Amplifier DPD*. http://www.unitest.com/pdf/5992-1954EN.pdf. Version: November 2016. – [Online; accessed 04.09.2021]

[15] Tuthill, J. ; Cantoni, A.: Optimum precompensation filters for IQ modulation systems. In: *IEEE Transactions on Communications* 47 (1999), Nr. 10, S. 1466–1468. http://dx.doi.org/10.1109/26.795813. – DOI 10.1109/26.795813

[16] Valkama, M. ; Renfors, M. ; Koivunen, V.: Advanced methods for I/Q imbalance compensation in communication receivers. In: *IEEE Transactions on Signal Processing* 49 (2001), Nr. 10, S. 2335–2344. http://dx.doi.org/10.1109/78.950789. – DOI 10.1109/78.950789

[17] Witt, D. P. D. ; Jacobus, Josias: Modelling, estimation and compensation of imbalances in quadrature transceivers, 2011

[18] Zhu, Anding ; Dooley, John ; Brazil, Thomas J.: Simplified Volterra Series Based Behavioral Modeling of RF Power Amplifiers Using Deviation-Reduction. In: *2006 IEEE MTT-S International Microwave Symposium Digest*, 2006, S. 1113–1116

[19] Zhu, Anding ; Draxler, Paul ; Yan, Jonmei ; Brazil, T.J. ; Kimball, Donald ; Asbeck, Peter: Open-Loop Digital Predistorter for RF Power Amplifiers Using Dynamic Deviation Reduction-Based Volterra Series. In: *Microwave Theory and Techniques, IEEE Transactions on* 56 (2008), 08, S. 1524 – 1534. http://dx.doi.org/10.1109/TMTT.2008.925211. – DOI 10.1109/TMTT.2008.925211

8. Appendix

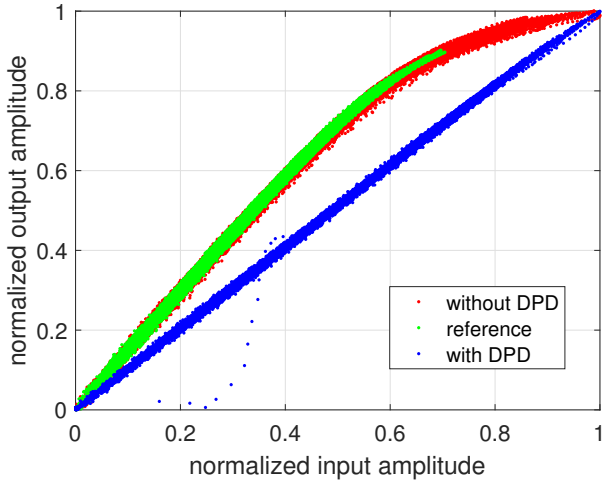
A. List of Abbreviations

Abbreviation	Meaning
3G	Third Generation
4G	Fourth Generation
5G	Fifth Generation
ACPR	Adjacent Channel Power Ratio
ADC	Analog to Digital Converter
AM	Amplitude Modulation
Attn	Attenuator
AWG	Arbitrary Waveform Generator
DAC	Digital to Analog Converter
DC	Direct Current
DPD	Digital Pre-distortion
DSB-SC	Double-Sideband Suppressed-Carrier
DSP	Digital Signal Processor
DUT	Device under Test
EIVE	Exploratory In-orbit Verification of an E-band link
EVM	Error Vector Magnitude
FET	Field-Effect Transistor
FIR	Finite Impulse Response
FM	Frequency Modulation
FPGA	Field Programmable Gate Array
HM	Harmonic
HPA	High Power Amplifier
IF	Intermediate Frequency
IM	Intermodulation
IQ	In-Phase Quadrature
ISI	Intersymbol Interference
LNA	Low Noise Amplifier
LO	Local Oscillator
LPF	Lowpass Filter
LTI	Linear Time-Invariant
LUT	Look Up Table
MeSFET	Metal Semiconductor Field-Effect Transistor
MPA	Medium Power Amplifier
PA	Power Amplifier

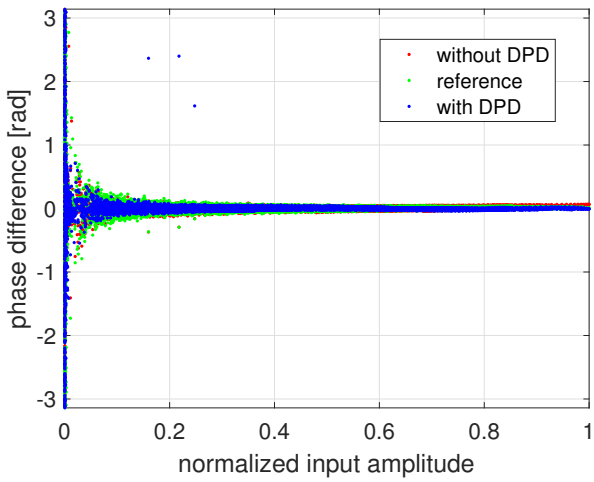
Abbreviation	Meaning
PAPR	Peak to Average Power Ratio
PM	Phase Modulation
PSD	Power Spectral Density
PSK	Phase-Shift Keying
QAM	Quadrature Amplitude Modulation
QM	Quadrature Modulation
QPSK	Quadrature Phase-Shift Keying
RF	Radio Frequency
RX	Receiver
SNR	Signal to Noise Ratio
SoC	System on Chip
TX	Transmitter
UHF	Ultra High Frequency
UMTS	Universal Mobile Telecommunication System
VGA	Variable Gain Amplifier
VHF	Very High Frequency
WCDMA	Wideband Code Division Multiple Access

B. Measurement Results

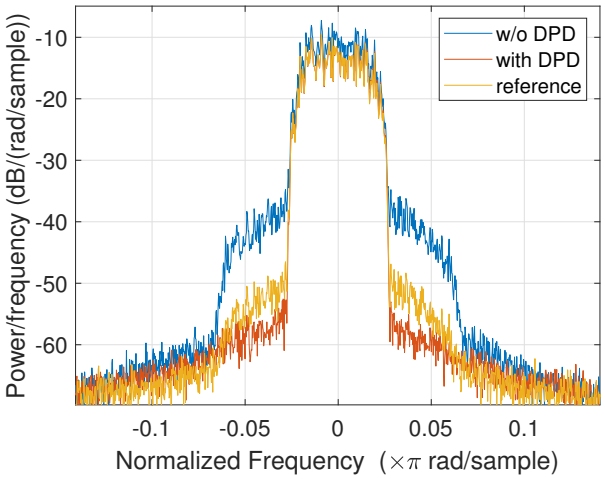
B.1. EIVE AWG and Mini-Circuits LNA



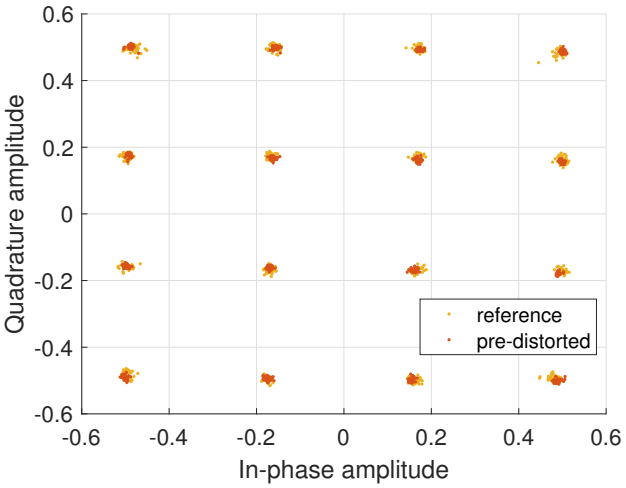
(a) AM/AM characteristics.



(b) AM/PM characteristics.



(c) Power spectral density.

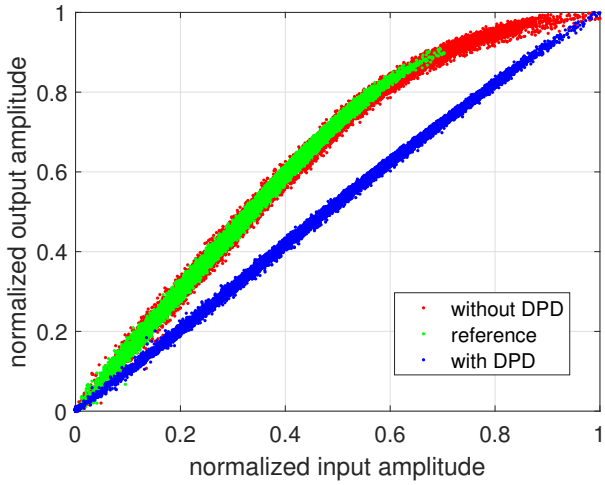


(d) Constellation diagram.

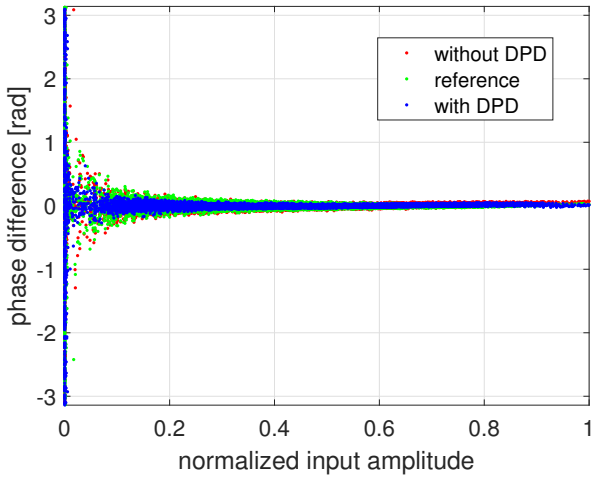
Figure B.1.: Plots for signal with 66 MHz bandwidth.

Signal	band power [dBm] (input)	band power [dBm] (output)	EVM (input)	EVM (output)	ACPR [dBc] lower (−81.31 MHz)	ACPR [dBc] upper (+81.31 MHz)
distorted	−9.63	9.63	1.55 %	4.33 %	−30.70	−30.28
reference	−12.73	7.29	1.58 %	1.89 %	−40.07	−39.85
pre-distorted	−12.63	7.29	-	1.00 %	−46.51	−46.97

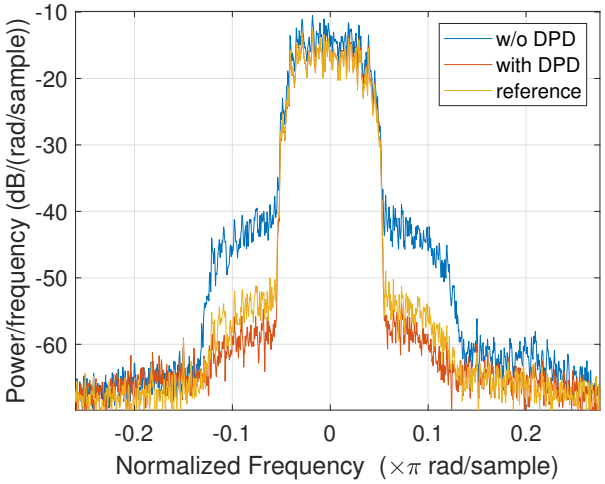
Table B.1.: Measurement results for signal with 66 MHz bandwidth.



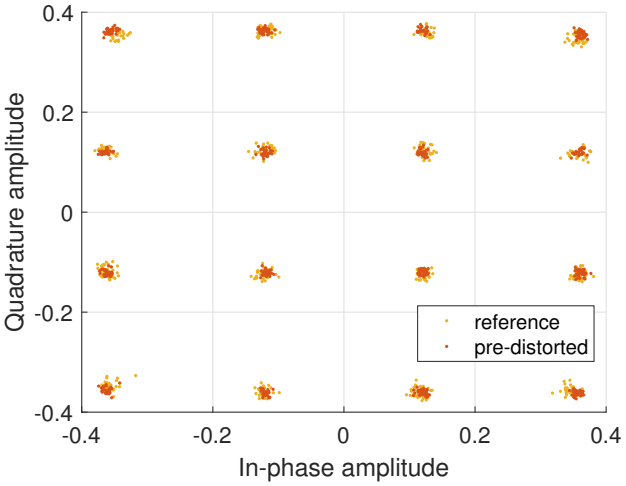
(a) AM/AM characteristics.



(b) AM/PM characteristics.



(c) Power spectral density.



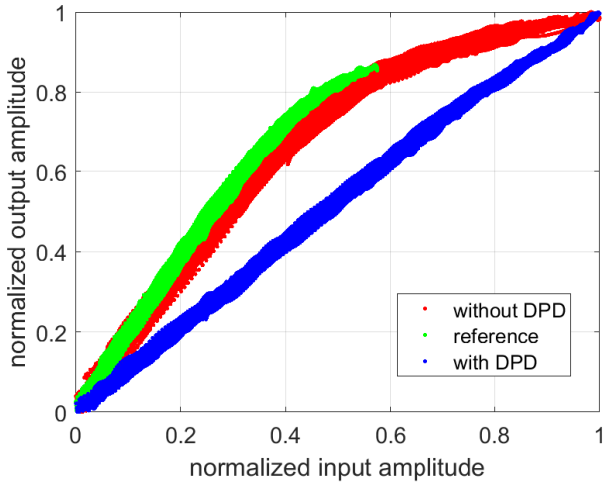
(d) Constellation diagram.

Figure B.2.: Plots for signal with 133 MHz bandwidth.

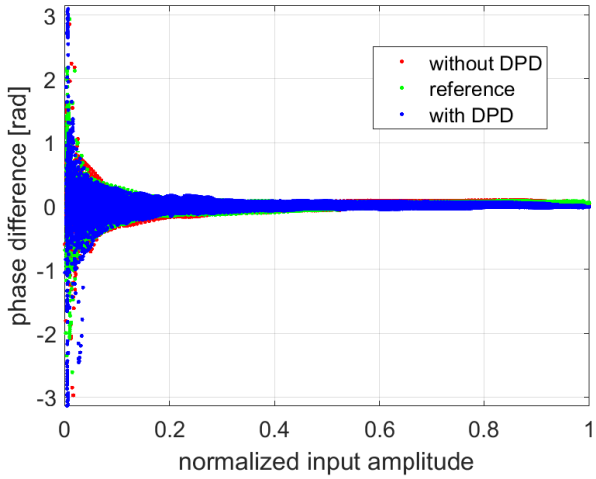
Signal	band power [dBm] (input)	band power [dBm] (output)	EVM (input)	EVM (output)	ACPR [dBc] lower (−163.86 MHz)	ACPR [dBc] upper (+163.86 MHz)
distorted	−9.42	9.79	1.97 %	4.77 %	−28.88	−29.37
reference	−12.48	7.48	2.01 %	2.48 %	−38.52	−39.02
pre-distorted	−12.37	7.54	-	1.26 %	−43.36	−43.64

Table B.2.: Measurement results for signal with 133 MHz bandwidth.

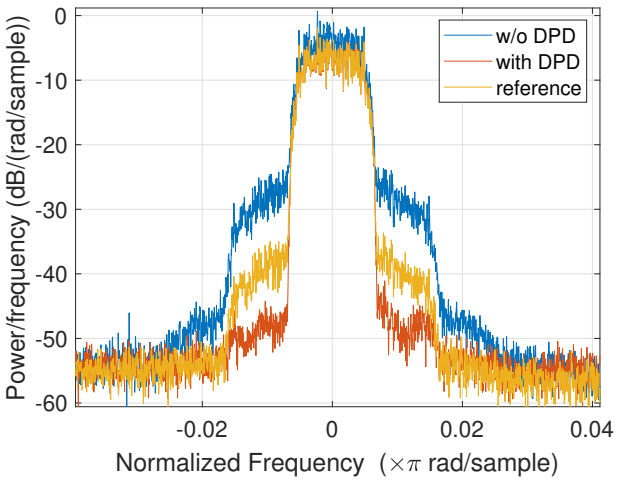
B.2. Keysight AWG and Mini-Circuits LNA



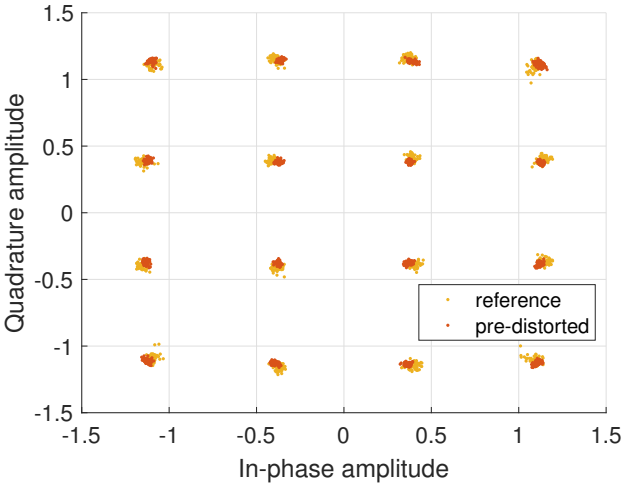
(a) AM/AM characteristics.



(b) AM/PM characteristics.



(c) Power spectral density.

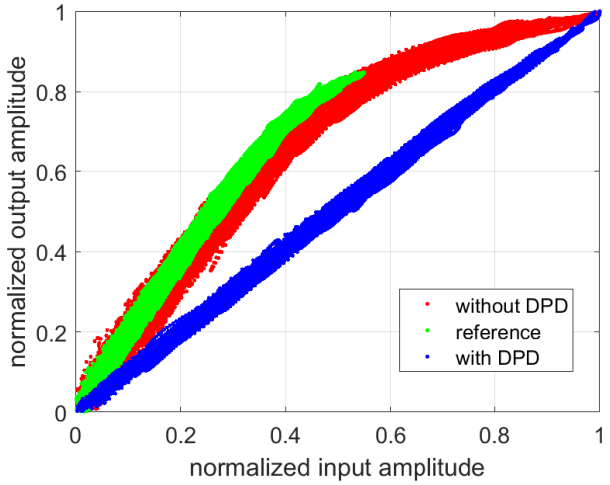


(d) Constellation diagram.

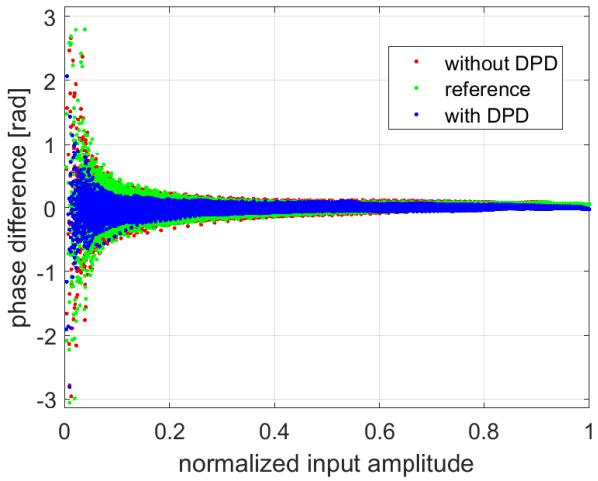
Figure B.3.: Plots for 16-QAM signal with 160 MHz bandwidth.

Signal	band power [dBm] (input)	band power [dBm] (output)	EVM (input)	EVM (output)	ACPR [dBc] lower (−197.12 MHz)	ACPR [dBc] upper (+197.12 MHz)
distorted	−5.40	11.11	0.88 %	7.16 %	−24.95	−24.73
reference	−9.60	8.55	0.75 %	2.99 %	−32.87	−32.72
pre-distorted	−9.58	8.55	-	1.30 %	−42.17	−40.31

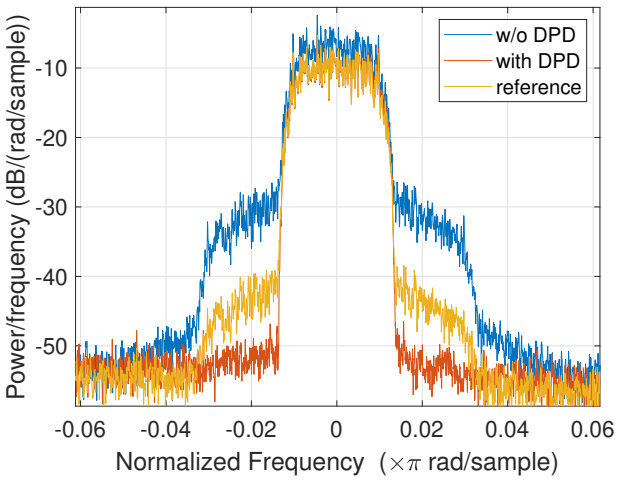
Table B.3.: Measurement results for 16-QAM signal with 160 MHz bandwidth.



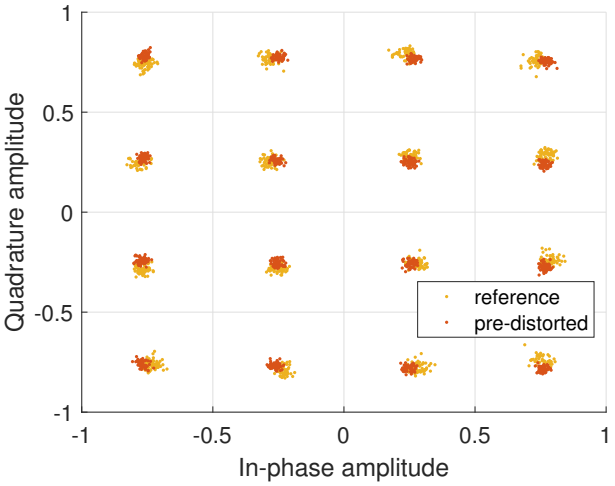
(a) AM/AM characteristics.



(b) AM/PM characteristics.



(c) Power spectral density.

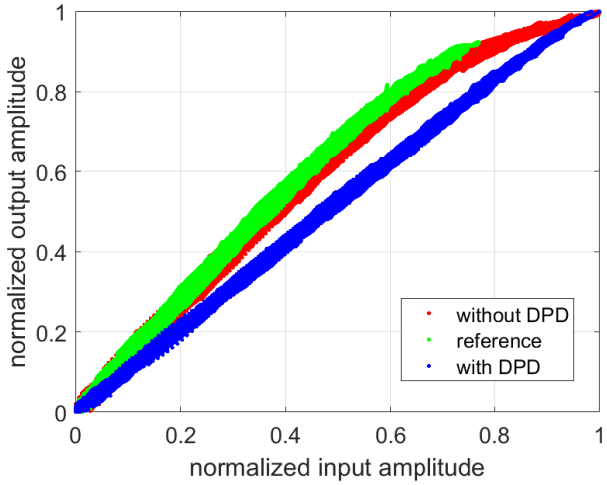


(d) Constellation diagram.

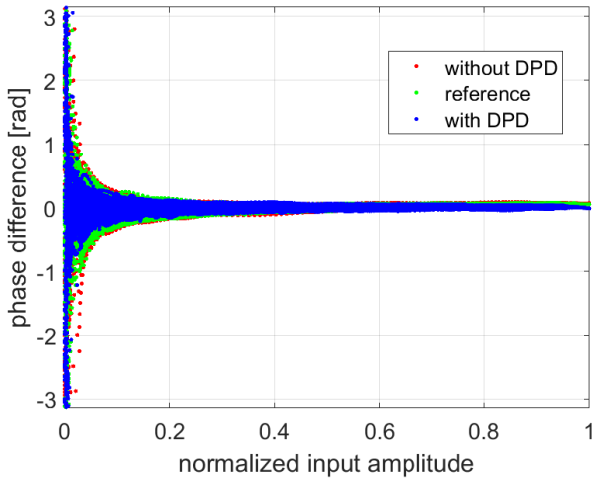
Figure B.4.: Plots for 16-QAM signal with 320 MHz bandwidth.

Signal	band power [dBm] (input)	band power [dBm] (output)	EVM (input)	EVM (output)	ACPR [dBc] lower (−394.24 MHz)	ACPR [dBc] upper (+394.24 MHz)
distorted	−5.42	11.06	1.23 %	7.21 %	−25.09	−25.20
reference	−10.06	8.19	1.01 %	2.99 %	−34.28	−34.96
pre-distorted	−9.94	8.21	-	1.43 %	−41.96	−43.59

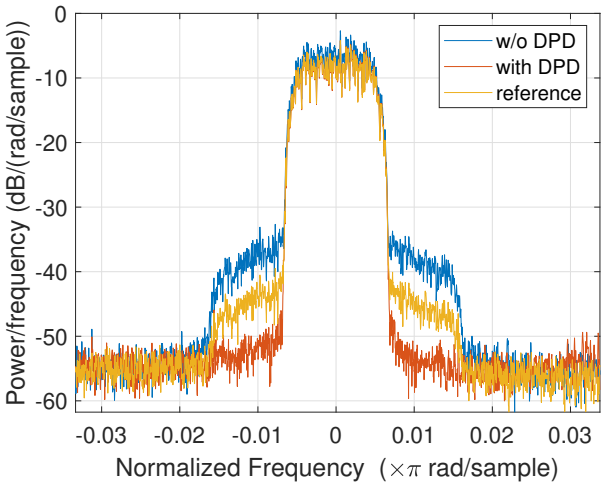
Table B.4.: Measurement results for 16-QAM signal with 320 MHz bandwidth.



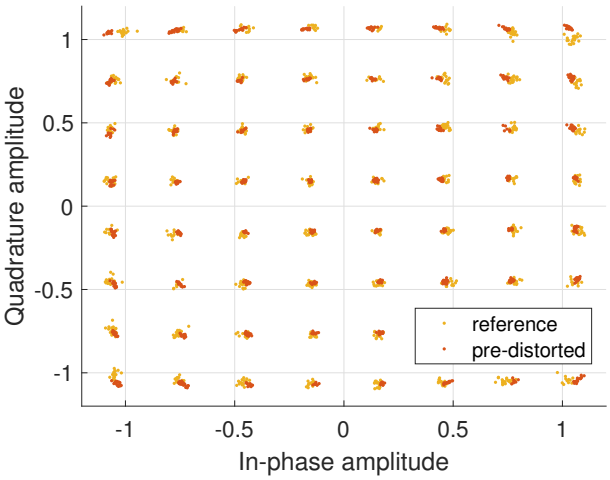
(a) AM/AM characteristics.



(b) AM/PM characteristics.



(c) Power spectral density.

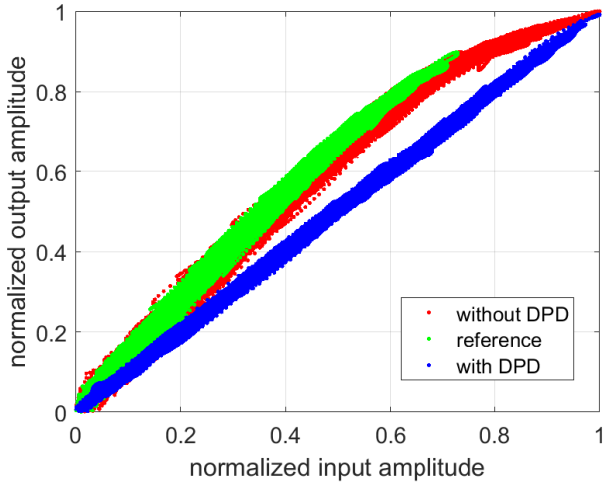


(d) Constellation diagram.

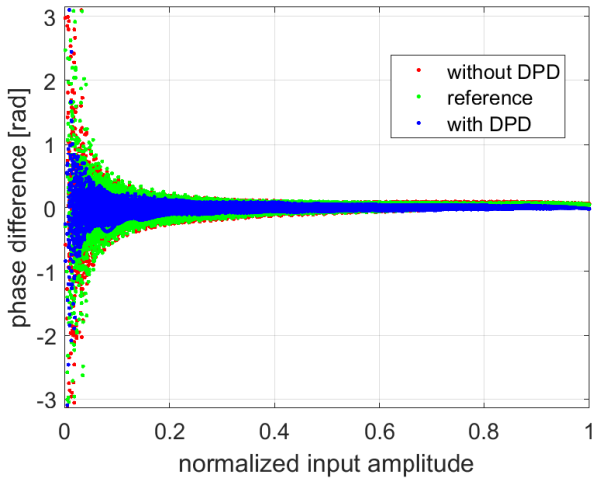
Figure B.5.: Plots for 64-QAM signal with 160 MHz bandwidth.

Signal	band power [dBm] (input)	band power [dBm] (output)	EVM (input)	EVM (output)	ACPR [dBc] lower (−197.12 MHz)	ACPR [dBc] upper (+192.12 MHz)
distorted	−9.56	8.48	0.73 %	3.19 %	−33.50	−32.69
reference	−11.54	6.94	0.88 %	1.96 %	−38.76	−38.30
pre-distorted	−11.50	6.94	-	0.71 %	−45.40	−45.53

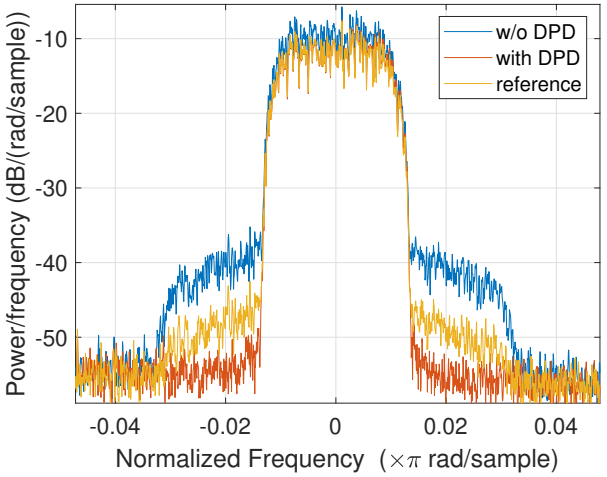
Table B.5.: Measurement results for 64-QAM signal with 160 MHz bandwidth.



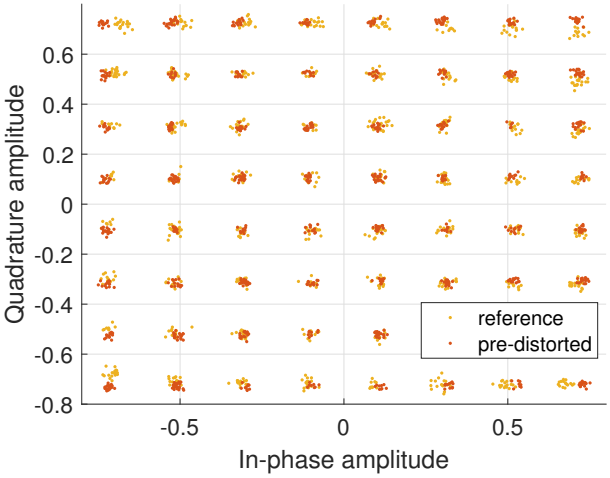
(a) AM/AM characteristics.



(b) AM/PM characteristics.



(c) Power spectral density.

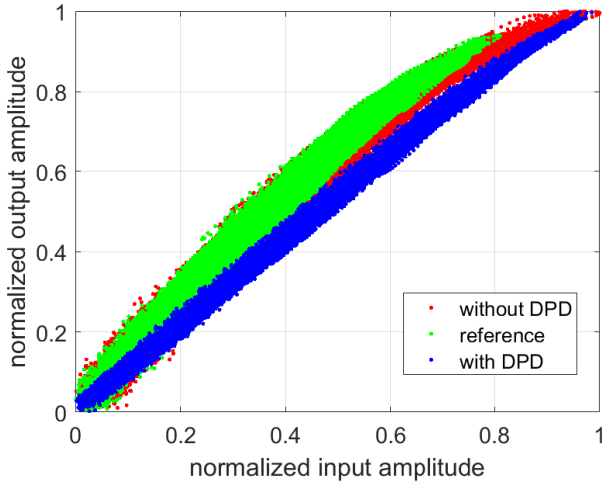


(d) Constellation diagram.

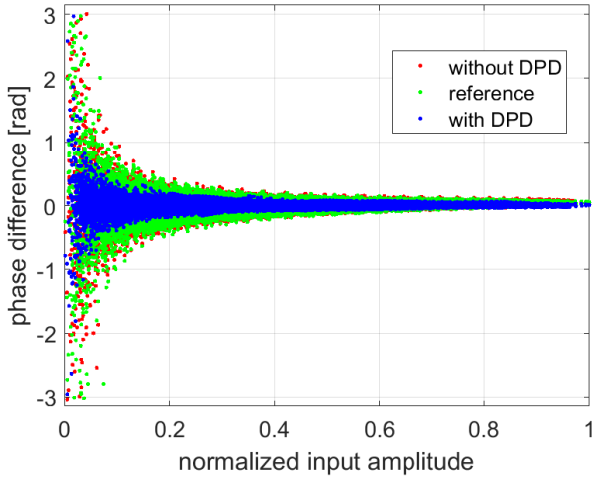
Figure B.6.: Plots for 64-QAM signal with 320 MHz bandwidth.

Signal	band power [dBm] (input)	band power [dBm] (output)	EVM (input)	EVM (output)	ACPR [dBc] lower (−394.24 MHz)	ACPR [dBc] upper (+394.24 MHz)
distorted	−9.61	8.49	0.95 %	3.41 %	−31.38	−31.75
reference	−11.98	6.65	1.15 %	2.25 %	−36.78	−37.84
pre-distorted	−11.79	6.64	-	0.99 %	−42.61	−43.42

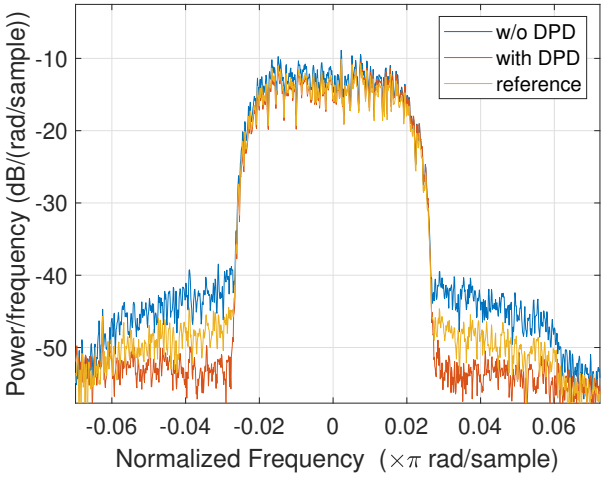
Table B.6.: Measurement results for 64-QAM signal with 320 MHz bandwidth.



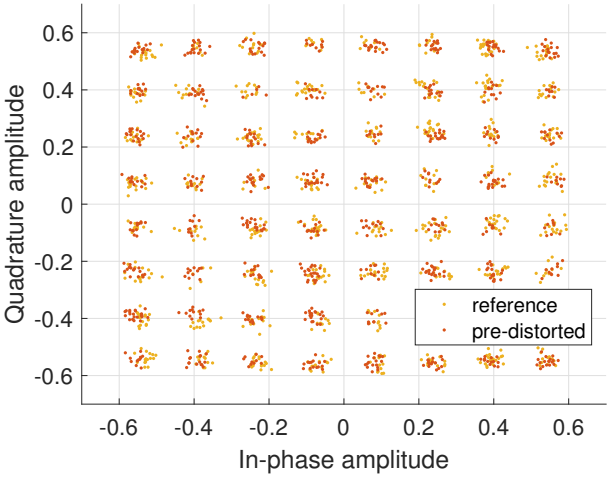
(a) AM/AM characteristics.



(b) AM/PM characteristics.



(c) Power spectral density.



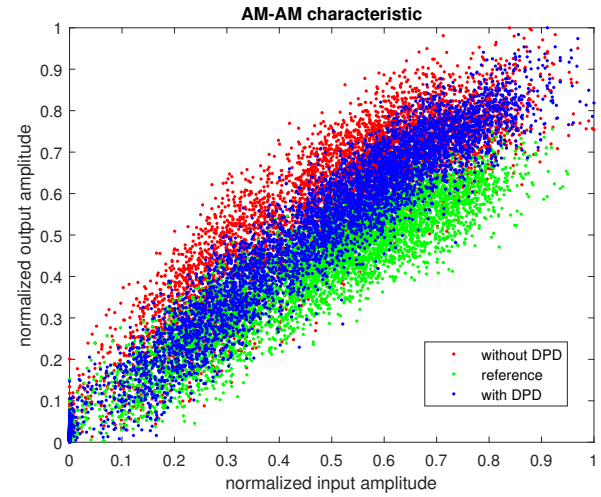
(d) Constellation diagram.

Figure B.7.: Plots for 64-QAM signal with 640 MHz bandwidth.

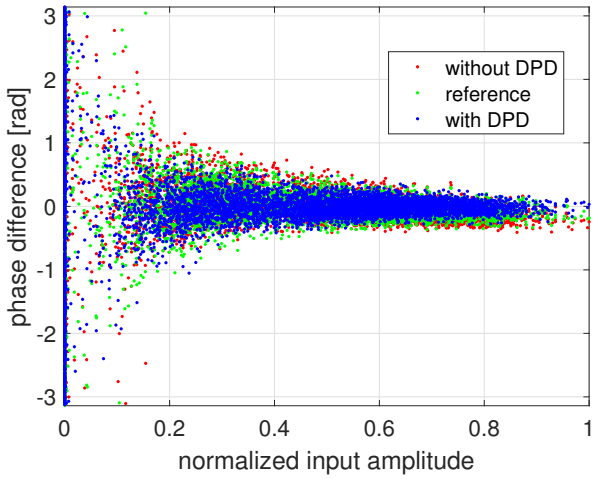
Signal	band power [dBm] (input)	band power [dBm] (output)	EVM (input)	EVM (output)	ACPR [dBc] lower (−788.48 MHz)	ACPR [dBc] upper (+788.48 MHz)
distorted	−9.62	8.52	1.29 %	4.12 %	−31.84	−32.26
reference	−11.22	7.27	1.29 %	3.42 %	−35.39	−35.98
pre-distorted	−11.11	7.26	-	1.60 %	−38.78	−40.00

Table B.7.: Measurement results for 64-QAM signal with 640 MHz bandwidth.

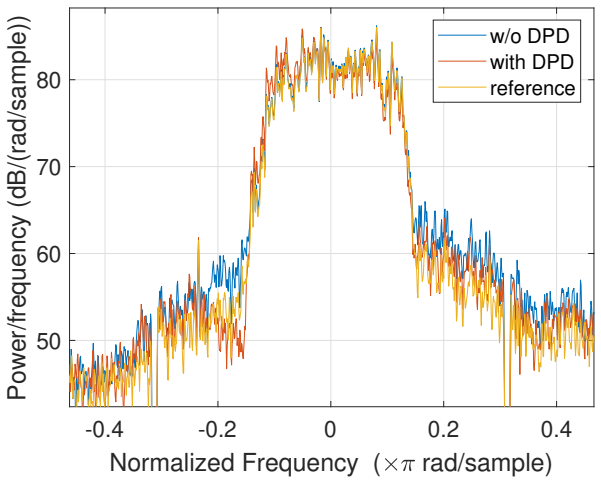
B.3. Keysight AWG and EIVE E-band Hardware



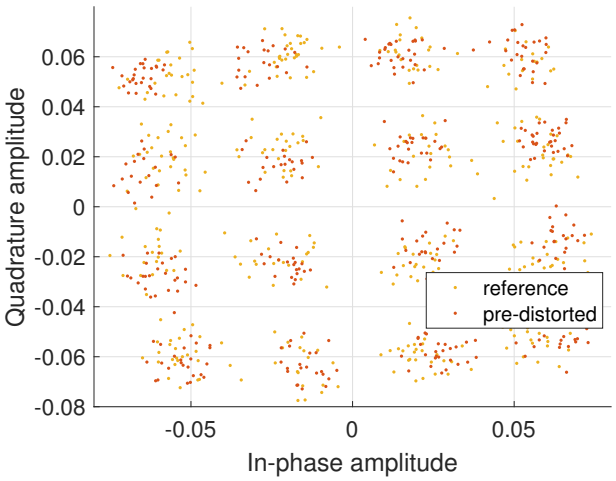
(a) AM/AM characteristics.



(b) AM/PM characteristics.



(c) Power spectral density.

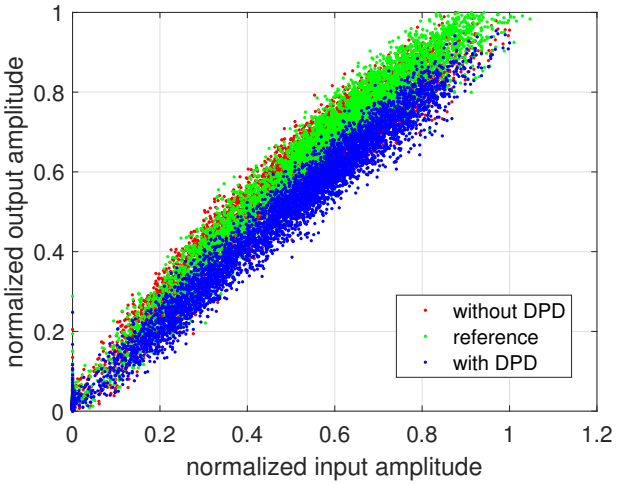


(d) Constellation diagram.

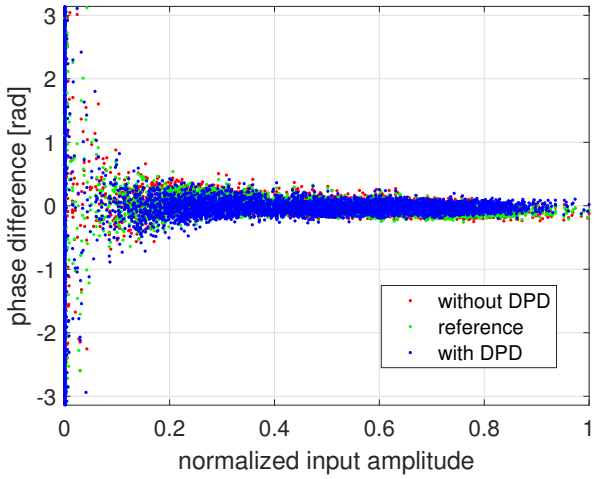
Figure B.8.: Plots for 16-QAM signal with 4600 MHz bandwidth; TX and MPA.

Signal	band power [dBm] (output)	EVM (output)	ACPR [dBc] lower (−5667.2 MHz)	ACPR [dBc] upper (+5667.2 MHz)
distorted	−0.94	12.8 %	−29.71	−24.02
reference	−3.64	12.3 %	−32.00	−27.59
pre-distorted	−3.44	9.3 %	−30.24	−25.84

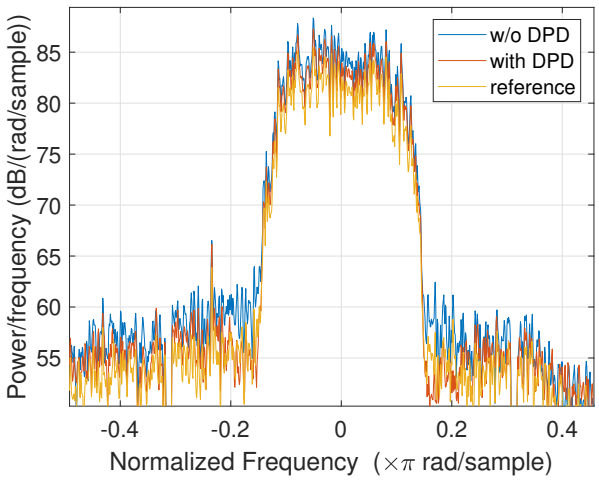
Table B.8.: Measurement results for 16-QAM signal with 4600 MHz bandwidth; TX and MPA.



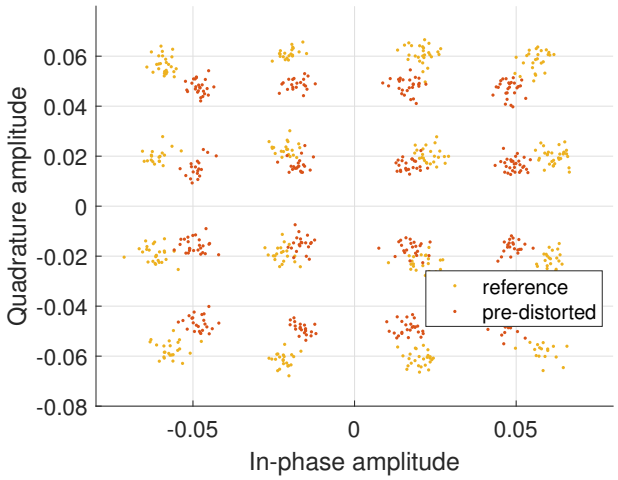
(a) AM/AM characteristics.



(b) AM/PM characteristics.



(c) Power spectral density.



(d) Constellation diagram.

Figure B.9.: Plots for 16-QAM signal with 4600 MHz bandwidth; TX.

Signal	band power [dBm] (output)	EVM (output)	ACPR [dBc] lower (−5667.2 MHz)	ACPR [dBc] upper (+5667.2 MHz)
distorted	−4.28	6.3 %	−25.76	−28.57
reference	−3.54	5.9 %	−26.72	−28.46
pre-distorted	−5.37	5.9 %	−26.48	−28.73

Table B.9.: Measurement results for 16-QAM signal with 4600 MHz bandwidth; TX.