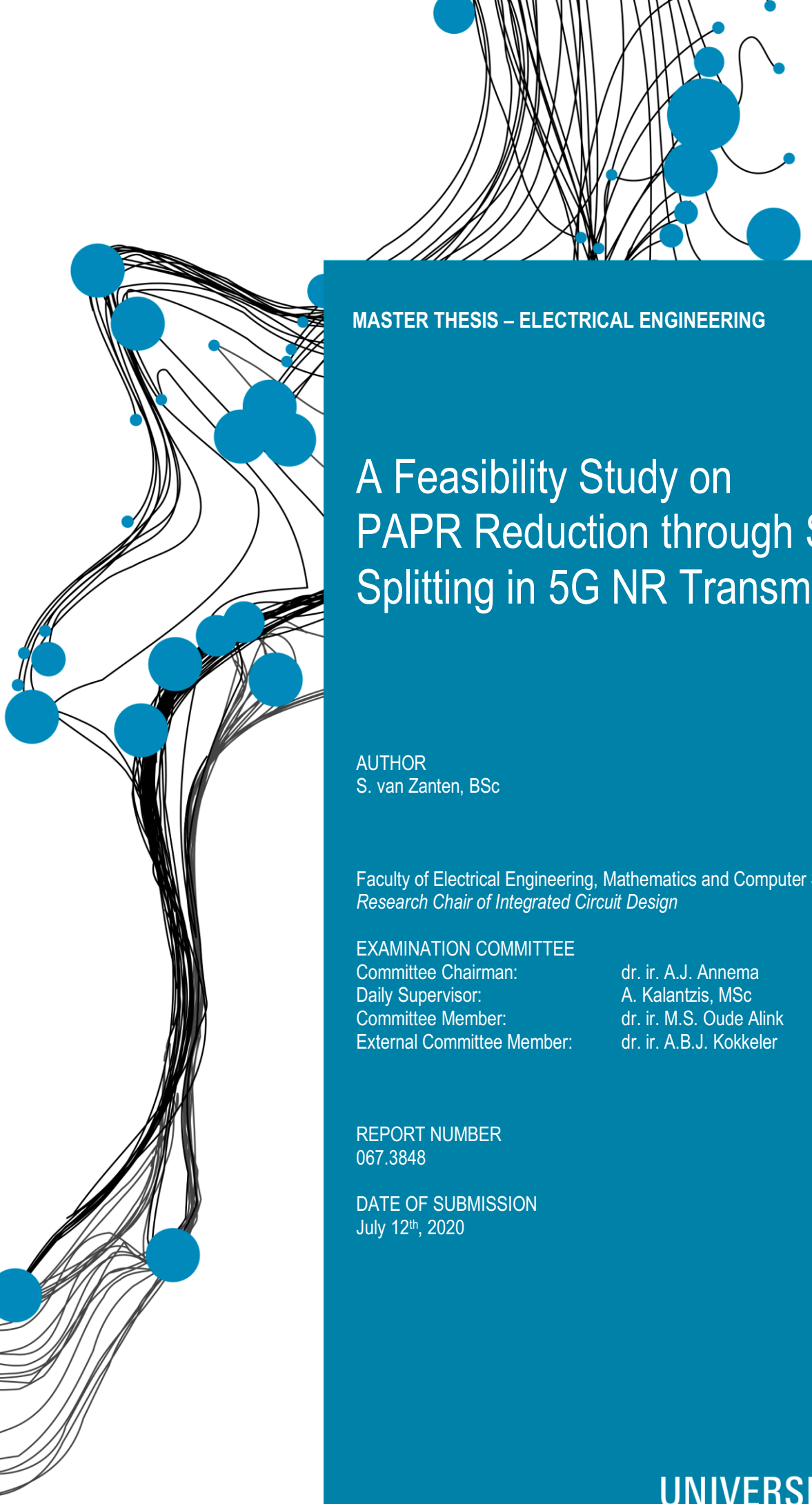


MASTER THESIS – ELECTRICAL ENGINEERING

A Feasibility Study on PAPR Reduction through Signal Splitting in 5G NR Transmitters

AUTHOR
S. van Zanten, BSc

Faculty of Electrical Engineering, Mathematics and Computer Science
Research Chair of Integrated Circuit Design

EXAMINATION COMMITTEE

Committee Chairman:	dr. ir. A.J. Annema
Daily Supervisor:	A. Kalantzis, MSc
Committee Member:	dr. ir. M.S. Oude Alink
External Committee Member:	dr. ir. A.B.J. Kokkeler

REPORT NUMBER
067.3848

DATE OF SUBMISSION
July 12th, 2020

UNIVERSITY OF TWENTE.

ACKNOWLEDGEMENTS

First of all, I would like to thank my supervisors for their time and effort invested in helping me during the graduation process. Anne-Johan, Andreas, Mark and André, your feedback has been very valuable and has helped me a lot in both conducting the research and writing the thesis. The year 2020 has proven to be challenging to many and, even though the university had to close due to the COVID-19 pandemic, the flexibility with which you have managed to continue graduation guidance using online platforms deserves recognition, I would like to thank you for that.

I would also like to thank Bram for his time and the critical reflection on some of the concepts I presented. You have helped to shape the results and brought the research forward.

Finally, I want to thank my partner and other family and friends for making my time so far at the University of Twente 6 awesome years to remember. I have truly enjoyed all the social activities that were a great distraction to all the serious studying. The people at Scintilla have offered many great activities that have allowed me to meet new friends and do new things, for which I am grateful.

Stef van Zanten, Doornspijk, July 12th, 2020

ABSTRACT

The introduction of 5G NR enables mobile communications in many new domains. As the usage of telecommunication soars, the demand for more efficient implementations grows. A major bottleneck in the development of efficient 5G NR transmitters is the negative impact of a high PAPR on its analog hardware, which is a result of combining many subcarriers into an OFDM signal. The aim of this thesis is to investigate to what degree the negative impact of a high PAPR as a result of the use of OFDM modulation on the transmitter's efficiency and linearity could be alleviated by making use of a combination of signal splitting and parallel hardware.

Two system concepts have been explored. The first uses splitting in the frequency-domain through decomposition of the OFDM signal into groups of one or multiple subcarriers. Each is processed by its own analog hardware track of which the parallel outputs concurrently deliver power to the load. As these outputs can deliver power concurrently, also the possibility to omit the inefficient PA stage was investigated. The second concept uses time-domain splitting. Based on the magnitude of the baseband OFDM signal, it is processed by a particular hardware track, resulting in PAPR reduction for the signals in each of these tracks. For each of these concepts, the main research question has been answered through a combination of theoretical analysis and simulations. Statistical analysis on OFDM signal properties was used to evaluate PAPR reduction and the impact of this reduction on the transmitter's hardware was considered through both literature studies and mathematical analysis. These theoretical components were verified with simulations on transmitter models in MATLAB/Simulink and LTSpice.

Analysis on frequency-domain signal splitting showed that to profit from splitting in terms of PAPR, only few subcarriers may be combined for each hardware track. This proved problematic when attempting to profit from the PAPR reduction as excessive amounts of hardware are required for the number of subcarriers used in 5G NR. Also, the parallel DAC outputs drive the same load concurrently with orthogonal signals resulting in partial cancellation of the generated power in the load. Using frequency-domain splitting for PAPR reduction in 5G NR transmitters is therefore not recommended. Research also explored using frequency-domain splitting to omit the inefficient PA stage. Here, the strict linearity specifications complicate implementation. If the DACs would be used to provide the required output power of the transmitter, large signals in the mixers lead to problems with gain compression. Also, full power would be generated before upconversion leading to increased dissipation in the mixers, which is why this approach is also not recommended.

Then, the feasibility of time-domain signal splitting to reduce PAPR was investigated. Here, significant PAPR reduction can be achieved with relatively little additional hardware. However, switching between hardware tracks was shown to introduce transient effects that degraded the signal's linearity and the ACPR. Both the tough linearity requirements and the large signal bandwidths requiring fast switching complicate the implementation of functionality to disable or enable the unused hardware tracks. Based on the resemblance of this design problem to considerations in the design of envelope tracking PAs, the preliminary conclusion in this thesis is that the large bandwidths required prevent a feasible implementation of this transmitter concept. However, the limitations to this research should be recognized and more research into disabling hardware tracks is recommended to come to a well-considered final verdict.

Considering the conclusions as presented in this thesis, the answer to the main research question is: no, the use of signal splitting and parallel hardware tracks does not benefit the negative impact of a high PAPR in 5G NR transmitters without significant degradation of other performance parameters. The identified root causes for the limitations that have surfaced in this thesis are the strict linearity requirements and the large signal bandwidth in 5G applications.

LIST OF ACRONYMS

5G 5th Generation Networking

AA Anti-Alias

ACPR Adjacent Channel Power Ratio

AM Amplitude Modulation

AWGN Additive White Gaussian Noise

BER Bit Error Rate

BO Back-Off

CM Common-Mode

CMOS Complementary Metal Oxide Semiconductor [Technology]

DAC Digital-to-Analog Converter

DNL Differential Non-Linearity

DPD Digital Pre-Distortion

ET Envelope Tracking

EVM Error Vector Magnitude

FDSOI Fully Depleted Silicon-on-Insulator

FFT Fast Fourier Transform

FIR Finite Impulse Response

iFFT Inverse Fast Fourier Transform

IM Intermodulation

INL Integral Non-Linearity

ISI Inter-Symbol Interference

LO Local Oscillator

LSB Least Significant Bit

LTE Long Term Evolution

MER Modulation Error Ratio

mmWave Millimeter Waves

MOSFET Metal Oxide Semiconductor Field-Effect Transistor

NMOS N-type Metal Oxide Semiconductor (Field-Effect Transistor)

NR New Radio

OFDM Orthogonal Frequency Division Multiplexing

PA Power Amplifier

PAE Power Added Efficiency

PAPR Peak-to-Average Power Ratio

pdf Probability Density Function

PM Phase Modulation

PTS Partial Transmission Sequence

QAM Quadrature Amplitude Modulation

RF Radio Frequency

RMS Root Mean Square

RX Receiver

RZ Return-to-Zero

SEM Spectrum Emission Mask

SLM Selective Mapping

SNR Signal-to-Noise Ratio

SSEC Small-Signal Equivalent Circuit

TF Transfer Function

TI Tone Injection

TL Transmission Line

TR Tone Reservation

TX Transmitter

Wi-Fi Wireless Fidelity

ZOH Zero-Order Hold

CONTENTS

1 INTRODUCTION

To sustain the rapidly expanding field of mobile communications, new hardware is being developed that supports the increasing demand for high-speed and reliable over-the-air networks. A part of the recently introduced 5th Generation Networking (5G) equipment contributes to this goal through the development of a novel air interface named New Radio (NR) (5G NR). 5G NR hardware targets higher speeds, improved reliability and significantly reduced latency, which will enable expanding the use of mobile networking to many new domains such as precision agriculture, remote healthcare and automated industry [1].

The current generation of Long Term Evolution (LTE) networking, typically referred to as 4G, uses sub-6 GHz frequency bands, as did its predecessors. Now that wireless network evolution has spanned several decades, allocating sufficient bandwidth to a new standard in a crowded spectrum for the desired 5G speeds is challenging. This is why the novel 5G air interface uses two frequency ranges: sub-6 GHz frequency bands that are located in the same frequency range as those used in 4G LTE networking and frequency bands in the Millimeter Waves (mmWave) range that have not been used in LTE before [2]. The much larger bandwidths available in the mmWave range help make higher data throughput feasible.

As did 4G LTE, 5G uses Orthogonal Frequency Division Multiplexing (OFDM) modulation for over-the-air data transmission, both for the sub-6 GHz and mmWave frequency ranges. OFDM modulation offers several distinct advantages over other forms of data modulation and especially the increased robustness to delay spread as a result of multi-path propagation is paramount to obtaining high enough reliable throughput in urban environments.

Unfortunately, many of the problems related to designing hardware that supports OFDM modulation are aggravated at higher frequencies; research is necessary to redefine bottlenecks and find renewed optima for classical design trade-offs. Within this context, this Master's Thesis is a contribution to 5G NR transmitter design and solving its associated problems.

1.1 Problem statement

One of the classical problems in the design of transmitters supporting OFDM is the high Peak-to-Average Power Ratio (PAPR) of OFDM signals. Because of this high PAPR, the Power Amplifier (PA) has to be sufficiently linear over a large dynamic range to support the signal's complex modulation. Achieving such a level of linearity under widely varying incoming signal power and antenna load variations with sufficient efficiency is challenging, especially in the millimeter range frequency bands. Attempts to improve the efficiency of the PA, predominantly in power back-off, have yet to yield the desired acceptable performance.

Since the introduction of OFDM, research has focused on identifying methods to alleviate the inherently high PAPR. Findings of the past decades on digital PAPR reduction techniques can roughly be grouped into five categories [3]: the use of clipping and filtering, the use of coding schemes, the use of Partial Transmission Sequence (PTS) or Selective Mapping (SLM), the use of nonlinear companding techniques and, finally, the use of methods related to Tone Reservation (TR) and Tone Injection (TI). Many of these techniques make changes merely to the digital hardware of the transmitter and to its software.

Alternatively, techniques that are deployed completely in the analog hardware domain also exist and focus on reducing the negative impact of a high PAPR signal on the hardware itself, generally without altering the PAPR. A well-known example is the Doherty-structure [4] in which the use of an additional

amplifier reduces the impact of the high PAPR on the transmitter's efficiency and linearity.

The effects of the aforementioned techniques are limited, whilst they come at increased degradation of several other design factors such as spectral efficiency and system complexity. To further improve transmitters, it is therefore worthwhile to characterize alternative methods to reduce the impact of a high PAPR that may have less negative side-effects. Opportunities in this regard have been identified in this thesis through analysis of the statistical signal properties of OFDM waveforms.

At the very basis of the high PAPR of OFDM modulated signals is the combination of many statistically independent orthogonally-spaced subcarriers modulated with Quadrature Amplitude Modulation (QAM) into a single signal. In line with the Central Limit Theorem, such signals tend towards a normal distribution and the PAPR is increasing when combining more and more subcarriers. Considering the typical negative impact of a high PAPR on analog hardware performance, it will be investigated if it will be beneficial for transmitter performance to process the desired waveform in separate parts as long as possible.

Research will therefore focus on splitting-up the OFDM waveform along either the frequency-domain (e.g. grouping subsets of subcarriers together) and the time-domain (e.g. smartly processing the samples of the time-domain OFDM waveform). Advantages and disadvantages of possible implementations will be mapped and a comparison to existing PAPR reduction techniques will be made to evaluate their effectiveness.

1.2 Research questions

In this thesis, the feasibility of novel system design concepts will be explored that are aimed at exploiting the aforementioned combination of signal splitting and parallel hardware tracks. The main research question in this thesis will be:

> *To what degree can the negative impact of high PAPR OFDM signals on transmitter efficiency and linearity be alleviated by applying a combination of signal splitting and parallel hardware tracks?*

In support of answering the main research question, a set of additional questions has been formulated.

First, the design space has to be defined. The transmitter design has to be compatible with existing networking equipment and must therefore adhere to global networking standards [2]. These standards characterize the design space available to the designer. An additional subquestion is therefore:

- What set of representative system requirements can be derived for 5G NR transmitters in the mmWave range based on the specifications in [2]?

Then, for both the frequency- and time-domain splitting, the following subquestions should be posed and answered:

- What type of concepts related to frequency-/time-domain splitting in transmitters can be identified in literature?
- How can the optimal method for splitting the signal in terms of PAPR reduction be mathematically modelled?
- How can hardware be designed in such a way that it benefits most from this PAPR reduction with minimal negative side-effects?

Just as important as the design is the verification of the theoretical analysis through simulations or measurements. In this thesis, the majority of the designs will be verified in MATLAB/Simulink®. Questions related to these simulations have been phrased as follows.

- How can basic functionality of the designed system concepts effectively be verified in simulations?
- What non-idealities should be incorporated into simulations to provide a comprehensive picture of system performance?

Finally, for both the frequency- and time-domain, a comparison will be made to existing PAPR reduction techniques in order to evaluate effectiveness.

- How do the effects of PAPR reduction through signal splitting stack up against existing PAPR reduction techniques?

1.3 Thesis organization

Chapter 2 will present fundamental concepts related to the design of transmitters processing OFDM waveforms. A formal definition for the PAPR that will be used in this thesis will be introduced. Chapter 2 will also contain a detailed literature review of existing techniques to reduce the high PAPR or alleviate its impact on transmitter hardware. An overview of the 5G NR transmitter design space will conclude the chapter.

Readers familiar with these topics may skip ahead to Chapter 3, in which important design considerations related to frequency-domain signal splitting will be treated. The chapter will contain a theoretical analysis on the PAPR reduction that can be obtained using this technique and provides a discussion on related design challenges. A comparison with the in Chapter 2 identified existing techniques will conclude Chapter 3.

Then, in the second half of the thesis, time-domain splitting techniques will be discussed. Chapter 4 addresses PAPR reduction through time-domain sample processing in a similar way as Chapter 3 did. A theoretical analysis on the PAPR reduction will be provided, followed by a discussion on the required hardware. The thesis is concluded with a discussion on the validity of these results in Chapter 5 and a conclusion in Chapter 6.

2 BACKGROUND INFORMATION ON 5G TRANSMITTER DESIGN

In support of the main chapters of this thesis, an overview of important design concepts with respect to 5G NR transmitter design will be provided. The reader is assumed to have a basic understanding of the working of transceiver systems. For refreshing design considerations related to transmitters in general, the reader is referred to extensive treatment in literature (e.g. in [5]).

To make the problem context concrete, a section has been devoted to signal statistics of transmitters based on OFDM and design considerations related to OFDM modulation will be discussed. The formal definition as used in this thesis to describe the PAPR will be introduced. Subsequently, an exploration into existing PAPR reduction techniques and amplifier structures that are able to (partly) alleviate a high PAPR input signal will be provided. Finally, based on the specifications from the 3GPP 5G NR standard, requirements for the transmitters treated in this thesis will be included to mark the boundaries of the design space to which this thesis adheres.

2.1 OFDM and its relation to PAPR

Information on OFDM waveform signal statistics provides insight into the origin of the relatively large PAPR in OFDM transmitters. To that end, this section comprises a discussion on basic OFDM related design considerations and on its mathematical background.

2.1.1 Signal properties of OFDM signals

In meeting the increased demands on wireless link throughput, the bandwidth of the baseband signal increases to facilitate higher data rates. As a result, the symbol duration of a typical QAM symbol becomes shorter and shorter. These short symbol durations complicate symbol decoding in the receiver. Especially in environments susceptible to multi-path propagation, a single symbol may travel over multiple paths with different delays, leading to overlapping symbol instances. This phenomenon is known as Inter-Symbol Interference (ISI) and leads to significant degradation of the signal's quality at the receiver.

To limit the impact of ISI, the symbol duration can be increased. The absolute difference in delay between the various propagation paths does not change and will constitute a smaller part of a symbol if its duration is longer. To still meet the requirements on the high data rate, many of these slower signals can be combined in what is known as an OFDM signal. An additional advantage of this approach is that the modulation-factor can locally be adjusted to the channel's quality for each of the subcarriers. Figure 2.1 shows an example of the combination of 4 subcarriers in the frequency domain [6].

The imposed orthogonality requirement on the so-called frequency spacing Δf avoids inter-subcarrier leakage and can mathematically be described as follows:

$$\Delta f = \frac{1}{T} \tag{2.1}$$

in which T is the useful symbol duration in seconds.

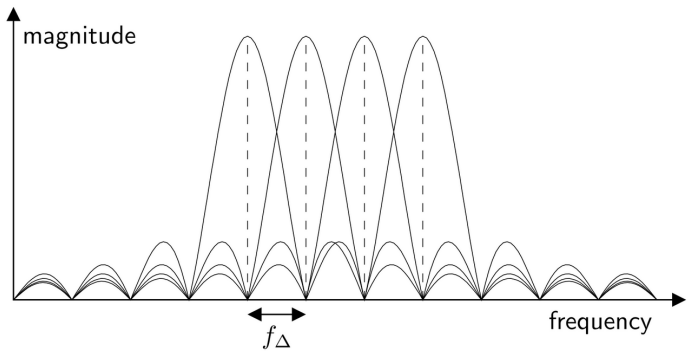


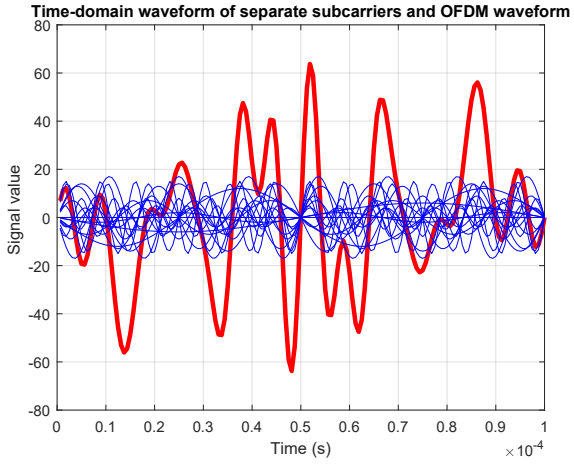
Figure 2.1: Simplified visualization of four orthogonally spaced subcarriers in the frequency domain [6].

As a result, the combination of N subcarriers into an OFDM signal should mathematically be defined as [6]:

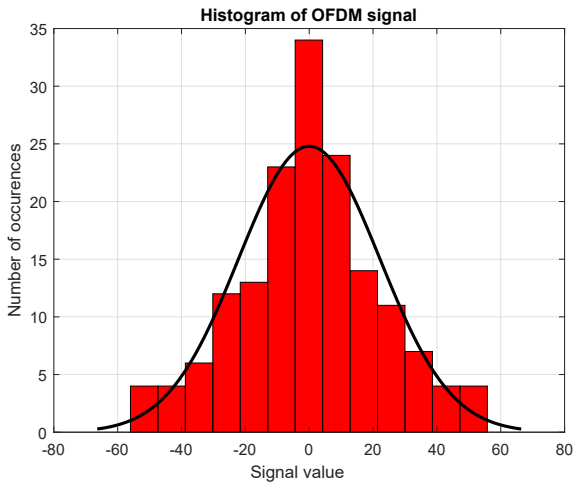
$$x(t) = \sum_{i=0}^{N-1} m_i(t) \cdot e^{j2\pi \frac{i \cdot t}{T}} \tag{2.2}$$

where $m_i(t)$ is the transmitted symbol or message modulating the carrier's phase and amplitude, mapped onto the i^{th} subcarrier running at frequency $\frac{i}{T}$.

The result in Equation 2.2 has major mathematical significance as it provides insight into many of the required signal properties in PAPR estimation, such as the distribution of the signal values over the dynamic range (the distance between the largest and smallest value). To further increase insight, Figure 2.2a visualizes 16 individual subcarriers running at orthogonally spaced frequencies (in blue). In line with Equation 2.2, these subcarriers collectively form either the in-phase or quadrature-phase component of an OFDM signal (depicted in red). Note that due to the combination of the independently modulated subcarriers, the PAPR of the signal has increased with respect to that of the original QAM signals.



(a) Time-domain waveform of separate subcarriers and of the I- or Q-component of an OFDM waveform in which these subcarriers have been combined orthogonally.



(b) Histogram of the I- or Q-component of an OFDM waveform as shown in Figure 2.2a. Superimposed is a fit to a Gaussian distribution.

Figure 2.2: Visualization of a typical I- or Q-component of an OFDM waveform in the time-domain.

A histogram-plot of the red waveform in Figure 2.2a shows that the extreme amplitude levels are much less likely to occur, see Figure 2.2b. On the contrary, the levels centered around the mean of zero are much more likely.

Through combining more and more independently modulated orthogonal subcarriers, the combined Probability Density Function (pdf) approaches a Gaussian, or normal, distribution, as stated by the Central Limit Theorem [7]:

> When independent random variables are added, their properly normalized sum tends toward a normal (or Gaussian) distribution even if the original variables themselves are not normally distributed.

Referring back to Figure 2.2b, to aid this understanding, a fit to a Gaussian distribution has been superimposed on the histogram. Already after combining 16 subcarriers, strong resemblance with a Gaussian distribution can be identified. Increasing the number of subcarriers further would improve the quality of the match with the Gaussian curve.

Mathematically, the pdf of a Gaussian distribution can be described as [7]:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{2.3}$$

> where $p(x)$ indicates the probability that the signal takes amplitude level x, μ is the mean value of the signal and σ is the standard deviation.

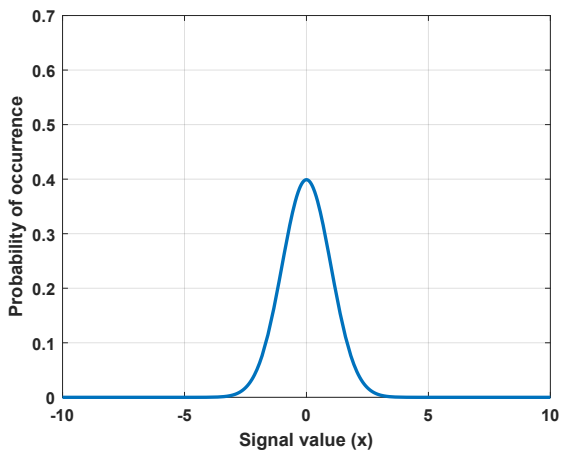
A pdf, such as the one in Equation 2.3, provides information on the probability for a certain signal value to occur, allowing direct estimation of the average signal value over time and hence of its PAPR.

The result in Equation 2.3 applies to the I- and Q-components of the OFDM waveform. Orthogonal combination of these two Gaussian distributions yields a new distribution for the resulting amplitude; the so-called Rayleigh distribution [7], [8]. The pdf of a Rayleigh distribution can be described as follows:

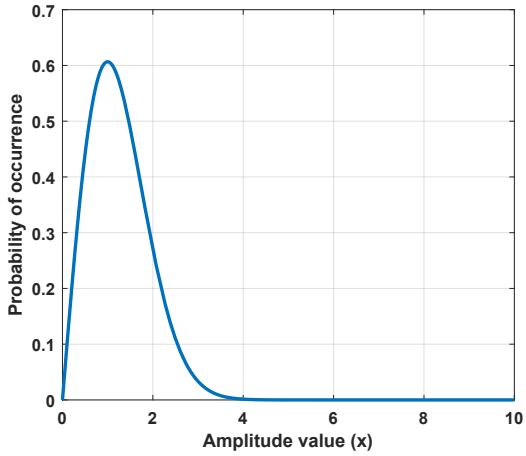
$$p(x) = \frac{x}{\sigma^2} e^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2} \tag{2.4}$$

> in which σ is the standard deviation of the original Guassian distributions that formed this Rayleigh-distributed signal.

The pdfs in Equations 2.3 and 2.4 have been plotted in Figure 2.3. Note that the Rayleigh distribution has only been defined for positive signal values and relates to the absolute amplitude value of the OFDM waveform.



(a) An example of a Gaussian distribution with $\mu = 0$ and $\sigma = 1$.



(b) An example of a Rayleigh distribution with $\sigma = 1$.

Figure 2.3: Visualization of the pdfs of the I- and Q-components of an OFDM waveform and of the complex magnitude of the combined OFDM signals.

2.1.2 Defining PAPR

Combination of an increasing amount of subcarriers into a single signal leads to an increase of the PAPR and a better approximation of the pdfs as plotted in Figure 2.3. In mathematical terms, the PAPR can be described as follows:

$$PAPR = \frac{|x_{peak}|^2}{x_{RMS}^2} \tag{2.5}$$

in which x_{RMS} is the Root Mean Square (RMS) value of the time-domain waveform.

A related quantity is called the Crest-factor and relates the amplitude quantities as opposed to the power ratios. When expressing the result in Equation 2.5 in decibels, it is equal to the Crest-factor due to the way decibels are calculated for power- and amplitude-quantities:

$$PAPR_{dB} = 10 \cdot log_{10}\left(\frac{|x_{peak}|^2}{x_{RMS}^2}\right) = C_{dB} = 20 \cdot log_{10}\left(\frac{|x_{peak}|}{x_{RMS}}\right) \tag{2.6}$$

in which C_{dB} indicates the Crest-factor in decibels.

A pdf of an OFDM signal also provides information on the PAPR of a waveform. If the pdf fully describes the distribution of the signal values, then the expected value of the distribution is equal to the Root Mean Square (RMS) value of the original waveform. The peak value is harder to relate to the pdf as the pdf describes a statistical process. For a sufficiently long observation window, the peak value will be equal to the outer edge of the pdfs. In the rest of this thesis, the observation will be assumed to be sufficiently long, meaning that for the OFDM signals, simply the peak value can be used. Note that in any real system, phenomena such as gain compression and intentional clipping limit the PAPR.

The increase in PAPR as a result of the combination of multiple independently modulated subcarriers has a negative impact on the efficiency and linearity of the hardware processing the signal. In 5G NR networking equipment in particular, the large number of subcarriers that is used results in both a large signal bandwidth of hundreds of megahertz as well as a high PAPR. The subsequent section devotes a few words to this topic.

2.2 Impact of OFDM modulation and high PAPR on transmitter hardware

As a result of the large PAPR due to the use of OFDM, the transmitter hardware has to process a signal with a large dynamic range of which most amplitude values are only rarely used. This impacts not only the PA, but also the mixer and Digital-to-Analog Converter (DAC) preceding it, see Figure 2.4 for a generic transmitter schematic.

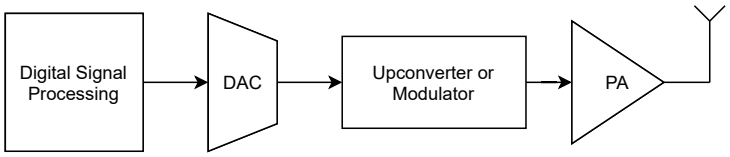


Figure 2.4: Generic transmitter schematic.

In the DAC, a large dynamic range results in relatively large Least Significant Bit (LSB) steps for constant resolution. If the majority of samples to be converted is concentrated around the smaller values, this results in a relatively large quantization error, degrading the Signal-to-Noise Ratio (SNR). In the mixer, the large dynamic range predominantly leads to issues with linearity retainment as compression of the outer amplitude values has to be avoided which is challenging due to the larger signal swings.

The majority of problems, however, are caused by the PA. To avoid linearity problems due to compression, the PA is typically operated in power Back-Off (BO). In addition, as a result of the large PAPR, the average value the PA has to amplify is relatively small. This results in even larger average power BO

over time. As PAs operate most efficiently when the big output Metal Oxide Semiconductor Field-Effect Transistor (MOSFET) is operated in the saturation region, power BO deteriorates efficiency. This efficiency degradation is severe and limits the development of more efficient OFDM transmitters. Figure 2.5 provides a schematic illustration of the impact of these considerations on the PA's efficiency curve.

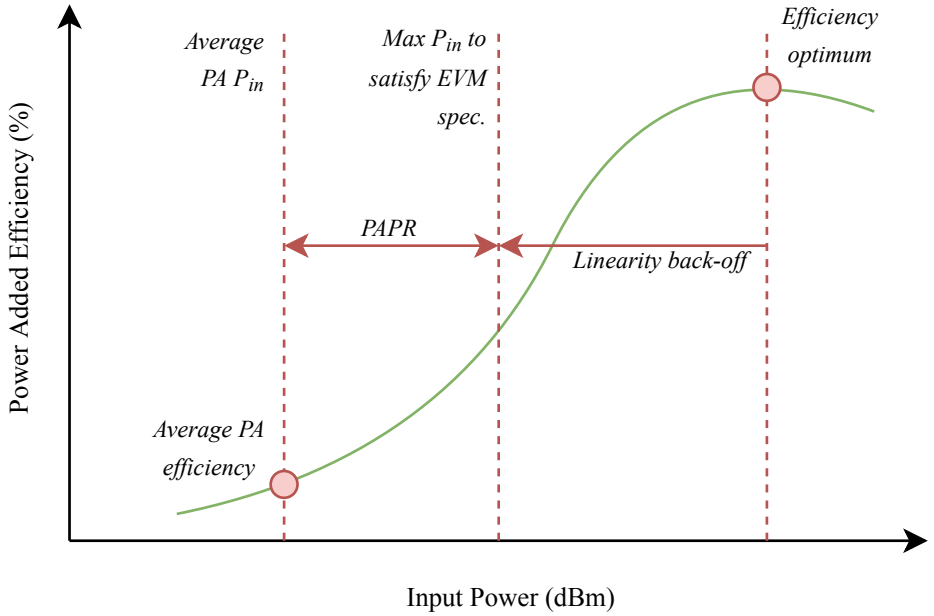


Figure 2.5: Schematic representation of the efficiency curve of a typical linear CMOS Power Amplifier. Typical TX operating regions have been annotated.

As the PA is the most power-hungry component of a typical transmitter [5], the Transmitter (TX) efficiency is heavily impacted by the efficiency of the PA. In the mmWave frequency range, the design of the PA is even more complicated due to extra high power losses and physical limitations to transistor technologies [9]. PAs in this frequency range are typically class AB to strike the right balance between linearity and efficiency [9].

To provide an indication of the efficiency figures that can be achieved with stand-alone Complementary Metal Oxide Semiconductor [Technology] (CMOS) amplifiers suitable for 5G applications in the mmWave range, two recent examples are worth including here based on the overview that Camarchia et al. have presented in [9]. In [10], a two-stacked PA has been designed that achieves a Power Added Efficiency (PAE) of 25 % at an average output power of 14.6 dBm when processing an LTE signal. In [11], a compensated distributed output network has been designed to improve the PA's efficiency. The authors report 14.6 % PAE at average output power for 5G signals.

Both examples and other designs as published in [9] show that the average PAE for CMOS PAs used in mmWave telecommunication systems is typically below 30 %, which is low. Considering the intention to use 5G in many more applications, the telecommunications share of global energy consumption would increase rapidly. The research in this thesis is a contribution to overcoming the negative impact of a large PAPR on the PA's efficiency, and therefore also on that of the transmitter.

The next two sections will discuss existing common techniques to limit the negative impact of a high PAPR on the transmitter hardware. The first focuses on digital techniques that alter the signal to reduce its high PAPR and the second section provides an overview of typical analog hardware implementations aimed at reducing the impact of a high PAPR signal on the hardware's performance (e.g. in terms of efficiency and linearity). Both sections combined provide a reference frame for performance characterization of the transmitter concepts that will be introduced in Chapters 3 and 4.

2.3 Overview of digital PAPR reduction techniques for OFDM transmitters

Problems with the high PAPR of OFDM signals are not new due to the use of OFDM in existing Wireless Fidelity (Wi-Fi) and LTE technologies and techniques to suppress high PAPR have been subject to thorough research. Most PAPR reduction techniques make changes to the digital hardware and software of the transmitter and change the signal properties of the waveform in such a way that the high PAPR is suppressed. These techniques can roughly be divided into five categories [3]. Table 2.1 lists these categories and provides citations to selected recent examples for each category that will be discussed in this section.

In [3], Tao and Jiang also provide an estimate of the typical PAPR reduction that can be achieved with these techniques. A column has been included in Table 2.1 that provides an estimate of the achievable reduction for each category based on the results published by Tao and Jiang. In advance of the requirements on 5G signals that will be treated in Section 2.5, the PAPR of the original OFDM signal used to characterize the reduction will be set to 10 dB. Note that no figure has been provided for coding schemes as the obtainable reduction in PAPR depends strongly on the used coding scheme.

Table 2.1: Overview of PAPR reduction techniques.

PAPR reduction techniques	**Examples since 2010**	**Typical PAPR reduction in dB wrt an OFDM signal with 10 dB PAPR [3]**
Clipping and filtering	[12]	-3.6 dB
Coding schemes	[13]	*Varies strongly with used scheme.*
PTS and SLM	[14], [15]	-4.0 dB (PTS)
Nonlinear Companding Techniques	[16]	-6.0 dB
TI and TR	[17]	-3.9 dB (TR)

In evaluating PAPR reduction techniques, the obtained PAPR improvement as presented in Table 2.1 has to be characterized in terms of the degradation of other system properties. In [3], Tao and Jiang name five system properties that should be considered in any PAPR reduction technique performance evaluation: increase in implementation complexity, bandwidth expansion, Bit Error Rate (BER) degradation, additional power consumption and an increase in spectral spillage.

The clipping techniques for instance, clips parts outside of a pre-defined allowed region which comes at the cost of increased spectral pollution as a result of the rapid transitions after clipping. Filtering may be used to improve out-of-band radiation. Clipping of the outer amplitude levels also leads to significant quality degradation of the clipped samples. An example of this technique is the research in [12], which introduces a compressed sensing system to partly overcome the linearity degradation, but requires changes to the receiver side of the wireless link as well.

Similar trade-offs can be identified in the other techniques. Coding schemes work by reducing the occurrence probability of the highest amplitude levels which generally increases the complexity of the transmitter and reduces the bandwidth efficiency as more data has to be transmitted. Taking [13] as an example; the PAPR is reduced by slightly disturbing the symbols (e.g. changing the original code) and by sending dummy symbols in unused carriers.

Partial Transmission Sequence (PTS) and Selective Mapping (SLM) are especially demanding for computational complexity. These techniques work by trying multiple phase corrections to limit the highest amplitude levels and sending information on the phase rotation to the receiver. There are typically two sides to recent research in the PTS and SLM domains: increasing complexity even further to yield higher PAPR reduction such as in [14] or additional research into reducing algorithm complexity whilst retaining the typical PAPR reduction (see [15]).

Nonlinear companding could be considered the opposite of clipping; instead of reducing the large amplitudes one increases the smaller-sized samples. The example in [16] addresses efficient alogrithms to realize such a technique. Similar to clipping, nonlinear companding techniques introduce additional nonlinearity to the signal.

Finally, tone injection (TI) and tone reservation (TR) techniques add time-domain PAPR reduction samples to the data-carrying subcarriers using additional 'empty' tones. This comes at significant bandwidth cost; the additional tones required limit bandwidth efficiency. [17] is a recent example of this technique being applied in optical-OFDM systems.

These trade-offs can be loosely generalized and visualized, refer to Table 2.2. As the implementation details vary, and therefore also the trade-offs between the various performance characteristics, Table 2.2 should be taken as an indication.

Table 2.2: Comparison of digital PAPR reduction techniques.

PAPR reduction technique	**PAPR improvement**	**Implement. complexity**	**Bandwidth expansion**	**BER degradation**	**Power consumption**	**Spectral spillage**
Clipping and filtering	$+/-$	$+$	$++$	$-$	$++$	$--$
Coding schemes	Varies	$+/-$	$-$	$+/-$	$+$	$++$
PTS and SLM	$+/-$	$--$	$++$	$+$	$-$	$++$
Nonlinear companding schemes	$++$	$+/-$	$++$	$+/-$	$+$	$+$
TR and TI	$+/-$	$-$	$--$	$+$	$+/-$	$+$

Although some of the discussed PAPR reduction mechanisms provide significant PAPR reduction (e.g. 4 dB or more) with reasonable sacrifices on other fronts, many of these techniques make changes to the data of the original signal and/or require changes to the receiver. This especially complicates the implementation of coding schemes, PTS, SLM, TR and TI within the boundaries of the 5G transmitter design space.

The two remaining techniques, clipping and nonlinear companding, could be relatively easily applied in 5G transmitters. However, both degrade the signal's linearity. In advance of the strict linearity demands following from the 3GPP-standard that will be presented in Section 2.5, there is little headroom to coarse PAPR reduction techniques.

An alternative approach to reducing the negative impact of a high PAPR is to refrain from trying to reduce the PAPR itself but try limiting the negative impact of it on the analog hardware instead. The next section will discuss some important examples of this design approach.

2.4 Analog hardware implementations to limit the negative impact of a high PAPR signal

As discussed in Section 2.2, transmitter performance is degraded strongly by a high PAPR signal at the input of a PA. Three common PA implementations suited for use in telecommunication systems that limit the impact of a high PAPR signal on their efficiency will be summarized here and related to the 5G NR design space. In literature, these three PA architectures, the polar amplifier, the outphasing amplifier and the Doherty amplifier, are also referred to as efficiency enhancement PAs [5], [18]. As with the digital PAPR reduction techniques, the design considerations in this Section will be related later to the transmitter concepts that will be introduced in Chapters 3 and 4 to evaluate their effectiveness.

Polar modulation

Figure 2.6 provides a schematic overview of what is today typically known as a polar-modulated amplifier, introduced in 1952 by Kahn in [19].

Polar modulation is based on decomposing the original baseband signal into two components: an envelope signal and a phase signal. Both are amplified separately and recombined in the final amplifier. Amplitude information is restored by modulating the supply of the phase amplifier with the envelope signal.

Applying polar modulation gives rise to several new issues in the design of the PA. First, due to mismatch between the two paths, signal corruption will occur. Because both paths contain different circuitry, it is

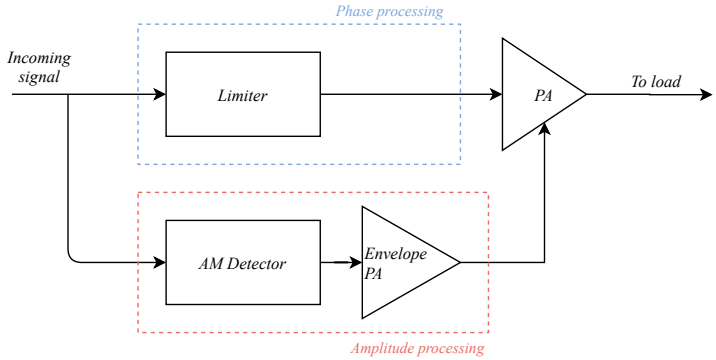


Figure 2.6: A schematic representation of a polar-modulated amplifier, as introduced in [19].

difficult to tune out possible phase delays. Second, although the original baseband signal is no longer amplified directly, the envelope signal still has variations in amplitude. These will degrade the signal's linearity when amplified with a nonlinear PA. Also, because the supply has to provide modulated signals, it needs to have a sufficiently high bandwidth. This is especially challenging in the design of 5G communications equipment for the mmWave-range where the baseband signal has hundreds of MHz of bandwidth.

The third issue is caused by the large output capacitance of the final power amplifier, which, due to supply modulation, gives rise to severe AM/PM-conversion, corrupting the phase information. Finally, the requirement for an additional buffer in the envelope path limits the voltage headroom for the phase PA. A problem that is especially stringent in modern CMOS nodes with a low supply voltage.

As a result of these limitations, implementing a polar amplifier for 5G applications is challenging. Camarchia et al. report no polar amplifier topology in their overview of technologies and design techniques for mmWave PAs [9].

Outphasing amplifier

Pioneered by Chireix in 1935 [20], the outphasing amplifier works by converting the original baseband signal into two phase-modulated components that together form the desired waveform. Figure 2.7 shows a schematic representation of an outphasing amplifier.

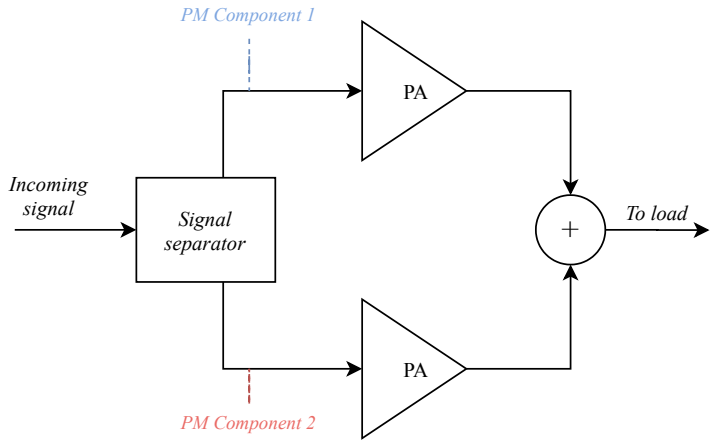


Figure 2.7: Schematic representation of an outphasing amplifier as introduced in [20].

The outphasing amplifier was designed to avoid amplitude modulation in its entirety and as a result, this architecture can operate completely nonlinear power amplifiers. However, this advantage comes at a cost.

First, generation of the two phase-modulated components is nontrivial and requires advanced circuitry. Second, the matching requirements for the two PA paths increase as the bandwidth of the separate phase-modulated components is larger than that of the original signal and gain/phase mismatches between the paths will corrupt the signal and lead to spectral regrowth. As with the polar amplifier, the large signal bandwidth required in 5G transmitters make this design problem especially challenging.

Finally, the outputs of both PAs still have to be summed. The circuitry required to restore envelope and phase information of the original signal from the two phase components introduces additional losses. Also, integrating full isolation between the two PAs into this circuitry can be shown to lead to inherent losses in a three-port network [21]. As a result, the PAs have undesired interaction, complicating their design.

These issues combined typically yield such high losses at mmWave design frequencies that outphasing PAs show very limited efficiency improvement at power BO [9]. In fact, due to their complicated designs, outphasing PAs exhibit degraded linearity with respect to simple class B PAs, while improving efficiency only slightly.

Doherty amplifier

A well-known example of an amplifier designed to handle high PAPR signals is the Doherty amplifier which was introduced in 1936 by Doherty [4]. Figure 2.8 shows a schematic representation of a Doherty amplifier.

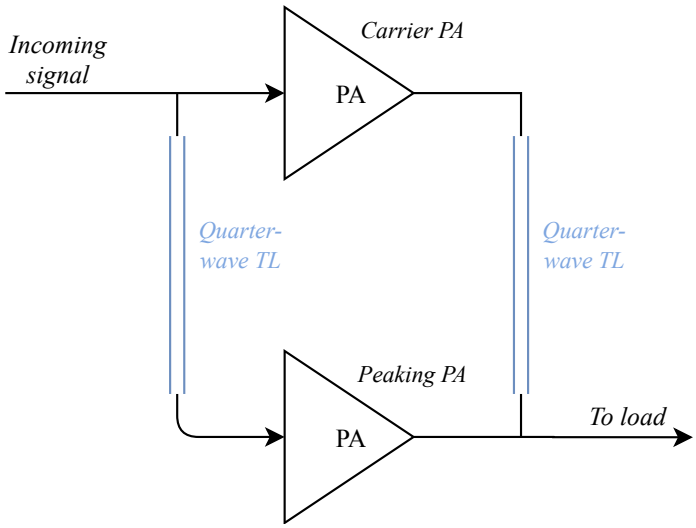


Figure 2.8: Schematic representation of a Doherty amplifier as introduced in [4].

In a Doherty structure, two amplifiers biased in two different operating regions are used: one operated typically as class A, AB or B (carrier amplifier) and the other operated as class C (peaking amplifier). The carrier amplifier amplifies continuously. However, when the input signal drives this amplifier into compression, the peaking amplifier kicks in and provides the required output power. The benefit is that gain compression can be avoided.

The Doherty amplifier has two major issues. Most important: the Transmission Line (TL) sections required for providing the correct loading of the PAs introduce significant losses and limit the bandwidth of the amplifier due to their quarter-wavelength requirement. This quarter-wavelength transformer inverts the impedance so that the resistive impedance seen by the main PA drops when the amplification of the peak PA increases, maintaining maximum voltage swing and efficiency for the main PA [18]. Second, for very large swings, the transistor in the peaking amplifier turns on and off frequently, introducing transient effects that degrade the Adjacent Channel Power Ratio (ACPR) in particular [5].

Note that the implementation of Doherty PAs is not as strongly limited by the wideband 5G NR signals as were outphasing and polar amplifiers and can better deal with high PAPR. Actually, many publica-

tions for designs of Doherty PAs at the mmWave frequency range exist that make use of the its structure as an efficiency enhancement technique at power BO. Examples from the overview as presented by Camarchia et al. in [9] include a mixed-signal Doherty PA that achieves 27.8 % PAE on average for a 64-QAM single-carrier 12 Gb/s signal [22] and a Doherty PA that achieves a PAE of 28 % at 6 dB power BO [23]. These efficiency figures are roughly 10 percentage-points higher than the efficiency figures reported in [9] for standard CMOS PAs.

The Doherty PA was the last of three efficiency-enhancement PAs that have been discussed in this Chapter. As for the design examples discussed in this Section, the to-be-designed hardware has to satisfy certain transmitter requirements to guarantee sufficient quality and intercompatibility with other networking equipment. The next section will define the requirements for the transmitter concepts that will be introduced in Chapters 3 and 4.

2.5 Defining the transmitter requirements

For 5G NR, the spectral regulations for 5G wireless links as derived by the organisation 3GPP dictate the design space boundaries. These spectral regulations for the operating region of the transmitter will determine many of its properties and other specifications logically follow from these requirements.

The spectral regulations for 5G NR have been included for many different types of equipment in [2]. The main aim of the research in this thesis is to determine the feasibility of the proposed transmitter architectures. To support this feasibility study, the 3GPP specifications have been interpreted and generalized to present a set of representative requirements that will be used to evaluate the designs in this thesis, refer to Table 2.3. In this evaluation, the impact of the specifications on system complexity and therefore also on simulation time has been taken into account.

For instance, hundreds of subcarriers may be used in 5G systems [2], which would impose great strain on computer systems having to perform simulations on transmitter designs processing so much data. Therefore, the number of subcarriers has been limited to strike the right balance between model simplicity and model accuracy. As will be shown in Chapter 3, combining 64 subcarriers already approximates the Gaussian distribution closely, as explained in the beginning of this Chapter[1].

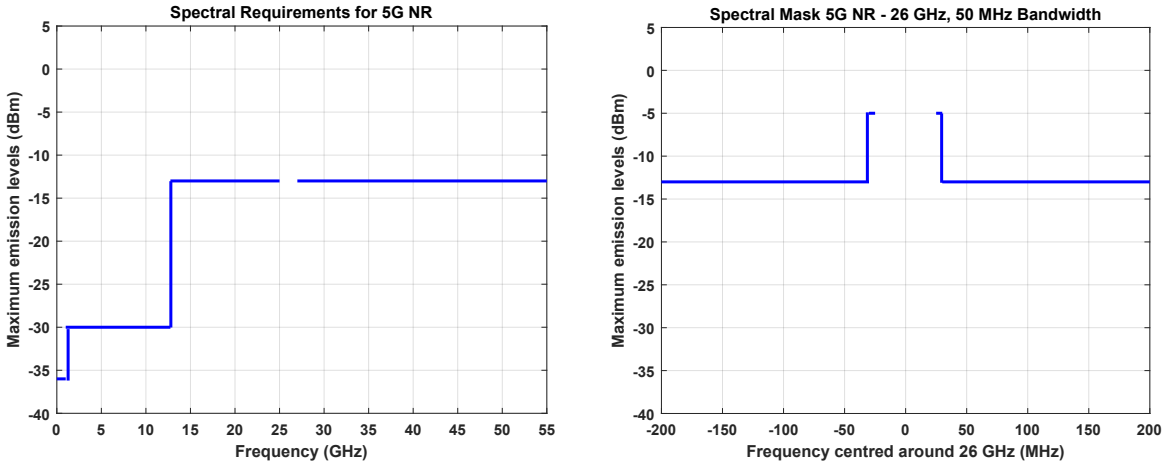
Output power and PAPR requirements have been based on both the specifications in [2] and additional information from industrial applications available to the author. The modulation scheme (and therefore also the Error Vector Magnitude (EVM)) is more strict than the typically specified 64-QAM in [2] to be able to also test future-proofing of the ideas. The remaining parameters in Table 2.3 have directly been derived from the specifications in [2].

Table 2.3: Overview of requirements that will be used in this thesis. These requirements are a generalized interpretation of the 5G NR – n258 operating band - Standalone requirements as included in [2].

Property	**Value**
Carrier frequency	26.0 GHz
Subcarrier spacing	60 kHz
Number of subcarriers	64
Modulation scheme	1024-QAM
Signal's PAPR	10.0 dB
Average output power	10.0 dBm
Maximum output power	20.0 dBm
Error vector magnitude (for 1024-QAM)	1.0 %
Adjacent Channel Power Ratio	-30.0 dBc

In addition to the specifications included in Table 2.3, the spectrum as transmitted by the designed transmitters has to adhere to the spectral mask in place for the complete frequency spectrum. The Spectrum Emission Mask (SEM) has also been included in the 3GPP standards [2]. These spectral masks can best be visualized and have been included in Figures 2.9a and 2.9b.

[1] In fact, already with 16 subcarriers the approximation is close enough to calculate the clipping distortion accurately enough for an SNR up to 50 dB, refer to Figure 3.12.



(a) Spectral Requirements for 5G NR. *(b) Spectral Mask 5G NR - 26 GHz, Zoomed.*

Figure 2.9: Spectral requirements for 5G NR Standalone - 26 GHz.

Combined with the functional requirements related to OFDM modulation, the specifications in Table 2.3 provide the boundaries of the transmitter design space for the ideas in Chapter 3 and 4.

3 PAPR REDUCTION BASED ON FREQUENCY-DOMAIN SPLITTING

Chapter 2 has shown that the combination of many independently modulated subcarriers is at the basis of the relatively high PAPR of OFDM waveforms. This is the first of two chapters to treat hardware solutions that process the waveform in parts. In this specific chapter, splitting of the signal along the frequency-domain will be treated.

A system overview of the transmitter concept based on frequency-domain splitting will be introduced first. Then, analysis will be performed to determine the PAPR reduction with respect to a 'traditional' OFDM transmitter that can be obtained. Design opportunities and implementation issues will be discussed. Finally, the chapter will be concluded with an overview of the results.

3.1 System overview of frequency-domain signal splitting in 5G transmitters

To understand how the OFDM signal can be split along the frequency-domain, first refer to the 'traditional' transmitter design for processing OFDM signals. Figure 3.1 provides a schematic of the digital hardware of a typical transmitter using OFDM modulation.

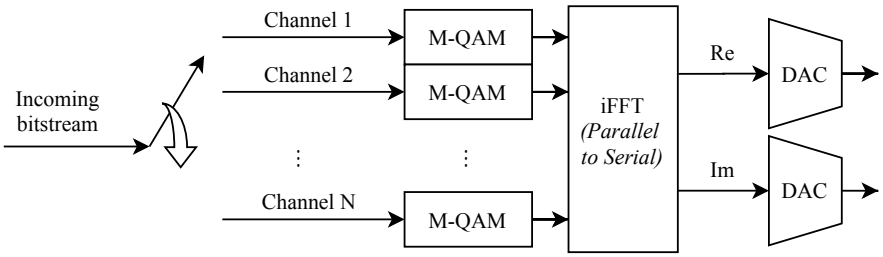


Figure 3.1: A simplified overview of the digital hardware of a typical OFDM transmitter. To process the OFDM signal in separate part, the iFFT block has to be discarded.

As Figure 3.1 illustrates, the individual channels (e.g. the orthogonally spaced subcarriers) in an OFDM signal are combined in the Inverse Fast Fourier Transform (iFFT). Therefore, to realize the system as intended in this Chapter where the OFDM signal is processed in split parts by separate analog hardware tracks, the implementation of the iFFT should be discarded in the new transmitter concept.

Consequently, next to the intended loss of signal combination functionality, also the ability to mix the channel symbols with the orthogonally spaced subcarriers is lost. As retaining orthogonality is paramount to maintaining the robustness to delay spread, the frequency spacing still has to be integrated. There are three options to do so: subcarrier mixing in the analog domain, subcarrier mixing in the digital domain or a combination of both. These three options will now be considered.

3.1.1 Integrating subcarrier mixing

When looking at subcarrier mixing in the analog domain, the subcarrier frequency is mixed with the QAM symbol after digital-to-analog conversion within the DAC. A benefit of this approach is that the DACs only have to support the bandwidth of the data rate which is much lower than the fastest running subcarriers [24]. However, this approach would heavily increase the required analog hardware, as an analog mixer

would be required for each subcarrier. Such a large number of analog mixers would add considerably to system complexity and introduce lay-out related problems such as leakage between the tracks.

Attempts to reduce the required number of analog mixers encounter limitations due to the required subcarrier spacing. To see why, consider the 'M-QAM' blocks in Figure 3.1, each provides an I- and Q-value at the same frequency; the data rate. If one would now add multiple of these outputs before mixing each with its own subcarrier frequency, the individual components can no longer be distinguished because they run at the same frequency. A hybrid combination of digital subcarrier mixing and analog mixing would be a solution to this problem; the digital mixers provide the relative spacing between the channels and the analog mixers translate the spectrum to the correct location in the intended spectrum, refer to Figure 3.2. Note that summing channels without digitally integrating spacing between the carriers would result in overlapping data.

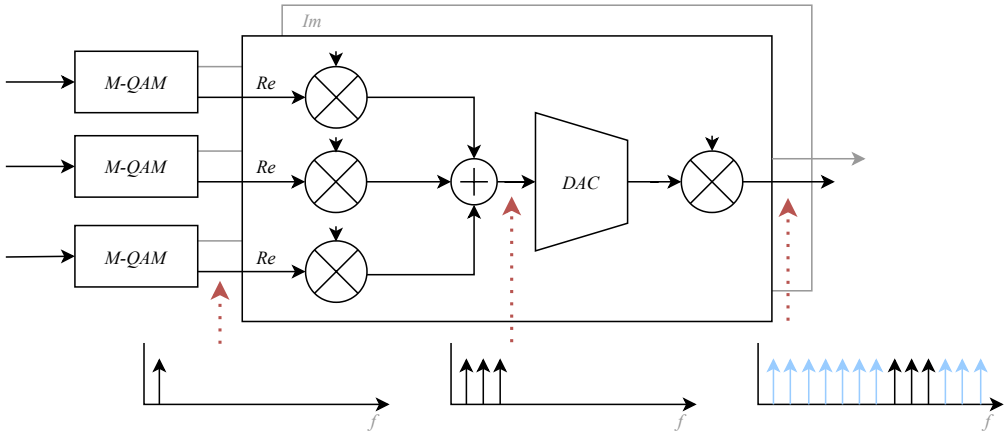


Figure 3.2: Hybrid mixing could be used to ease the requirements on the DACs while limiting the number of analog mixers. The digital mixers provide the relative spacing between the subcarriers. The analog mixers translate a group of digitally spaced subcarriers (in this case three) to the correct location in the intended output spectrum. The location of other subcarriers in that spectrum has been indicated.

As an alternative to the hybrid approach, it is also possible to perform subcarrier mixing fully in the digital domain. This will reduce system complexity as no analog mixers are required to provide the correct subcarrier frequencies and is closest to the original OFDM implementation. The disadvantage is that (at least some of) the DACs will have to support the full bandwidth of the OFDM signal, increasing their bandwidth requirements[1]. Figure 3.3 illustrates how this process of subcarrier mixing in the digital domain impacts the frequency spectra.

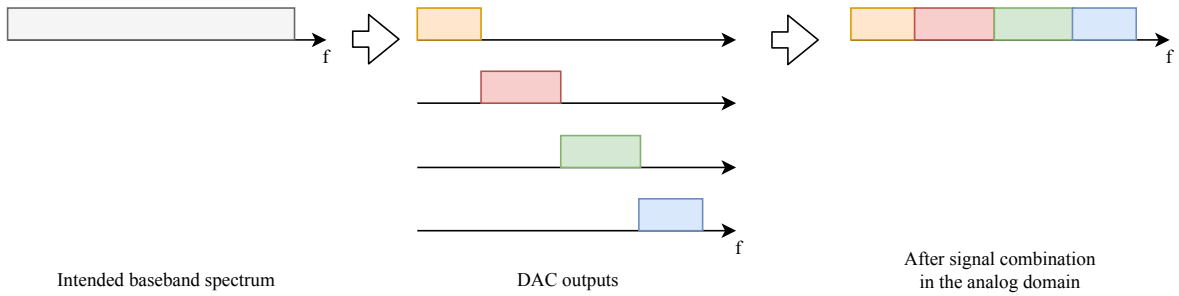


Figure 3.3: Splitting of the baseband constellation into smaller portions for processing in separate hardware tracks.

Because of its simplicity and the fact that the DACs are typically fast enough to handle the full bandwidth of the OFDM baseband signal (in the system as depicted in Figure 3.1 the DACs also have to process

[1]As opposed to hybrid mixing, where the OFDM signal is effectively cut in pieces with narrower bandwidth, that are translated to the correct frequency in the analog domain.

the full baseband signal), in this chapter, the subcarrier frequencies will be mixed in the digital domain. A discussion on this design decision will be provided in Chapter 5.

3.1.2 Digital hardware overview

In hardware terms, the digital part of a typical transmitter as included in Figure 3.1 will be replaced by a new digital hardware section as shown in Figure 3.4. The optimum number of subcarriers to be processed per DAC will be the subject of analysis in this chapter (in Figure 3.2 for example, 3 subcarriers are converted using each DAC) as it will impact the PAPR of the signal at the input of each DAC.

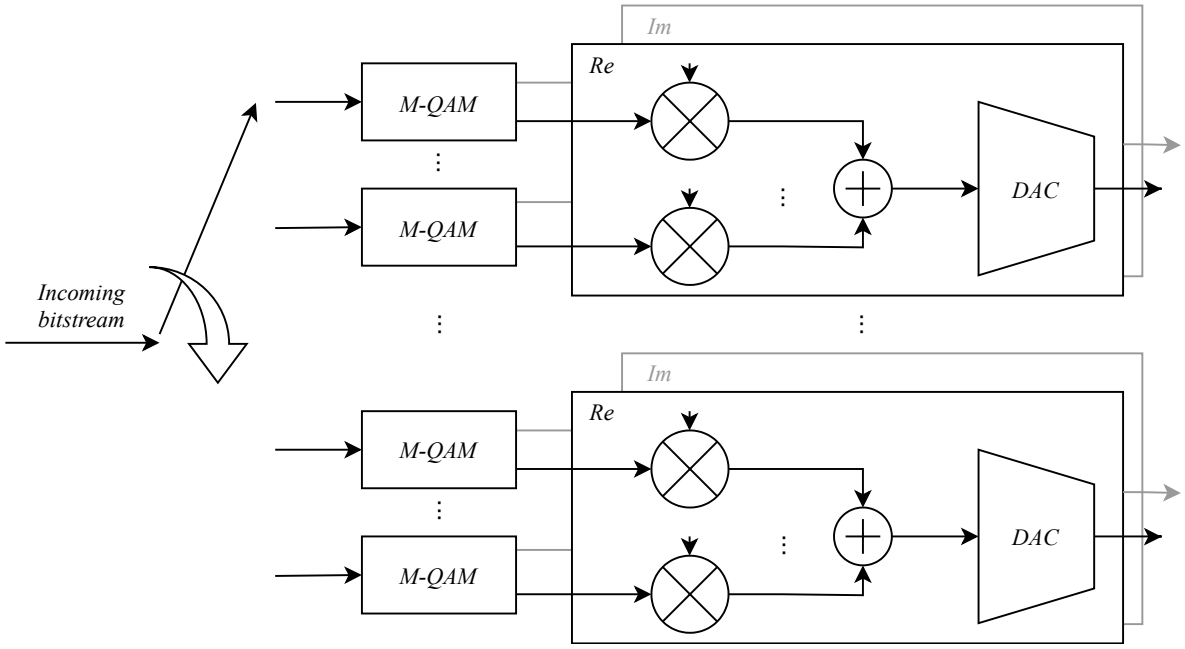


Figure 3.4: A simplified overview of the digital hardware of the proposed transmitter based on frequency-domain signal splitting and parallel hardware tracks. By limiting the number of subcarriers to be converted per DAC, its requirements can be eased. The optimum number of subcarriers will be subject to analysis in this chapter. This design is based on subcarrier mixing in the digital domain.

The left side of Figure 3.4 shows the channels similar to those in an OFDM based transmitter (see Figure 3.1), see the 'M-QAM' blocks. Each of these channels is mixed with its subcarrier. A number of these subcarriers can be combined before conversion to the analog domain using a DAC. As mentioned, the optimal number of subcarriers to combine will be explored in this chapter. Each of the rectangles in Figure 3.4 containing the subcarrier mixing, subcarrier summation and DAC will be referred to as a hardware track.

3.1.3 Analog hardware considerations

The transmitter's analog hardware may also benefit significantly from splitting along the frequency domain as the parallel hardware tracks will deliver power concurrently to the load at a particular time instance. To optimally benefit from the reduced PAPR, the required signal power should be generated when the signal is still split, reducing the dynamic range of the to-be-generated signal and preventing part of the problems related to a high PAPR as discussed in Section 2.2. There are three options to generate the required analog power when the signal is still split:

- Implement the typical chain of DAC $\rightarrow$ Mixer $\rightarrow$ PA for each parallel hardware track.
- Increase the size of the DAC so that it can directly drive the mixer/load and omit the PA.
- Implement a 'regular-sized' DAC and amplify this signal with a PA before mixing.

The overview in Chapter 2 showed that PAs in the mmWave frequency range for 5G applications typically achieve efficiency figures less than 30 % at average output power. If the implementation that relies on the DACs to deliver the output power would be feasible, the inefficient PA stages would be rendered unnecessary, hopefully benefiting transmitter efficiency.

Therefore, in the remainder of this chapter, the design considerations related to the digital hardware in Figure 3.4 and the complications of using DACs to provide the required output power for the analog hardware will be analyzed.

Still, there are many options to implement the required analog hardware. Section 3.4 will provide a detailed discussion on the design consequences related to its implementation.

3.1.4 Simulink model of important signal flows

To increase insight into the signal flows in the digital hardware schematic of Figure 3.4 and a straightforward implementation of the analog hardware by directly summing the DAC outputs, a Simulink model has been designed, refer to Figure 3.5. Note that this example converts a single subcarrier using a single DAC.

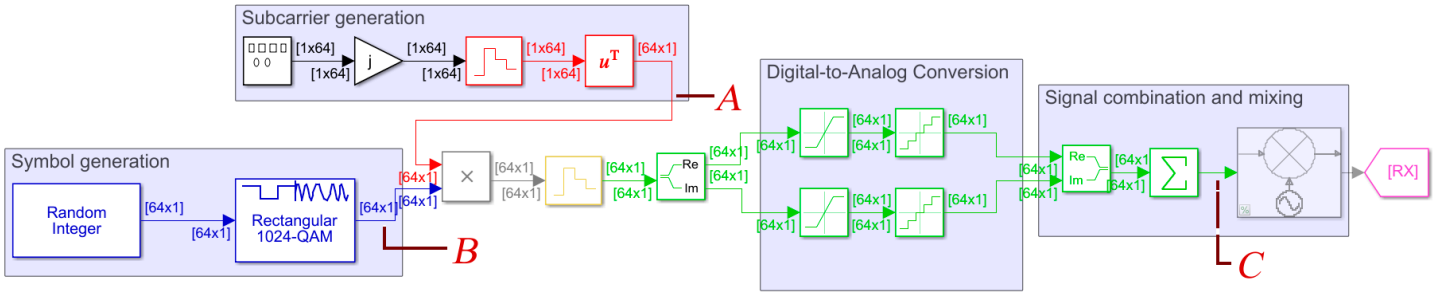


Figure 3.5: Simulink model of the (baseband) transmitter schematic as included in Figure 3.4. Several relevant signals have been indicated with a character, referring to the subfigures in Figure 3.6. In this model, 1 subcarrier is converted with 1 DAC and the analog hardware consists of a simple summation of DAC outputs.

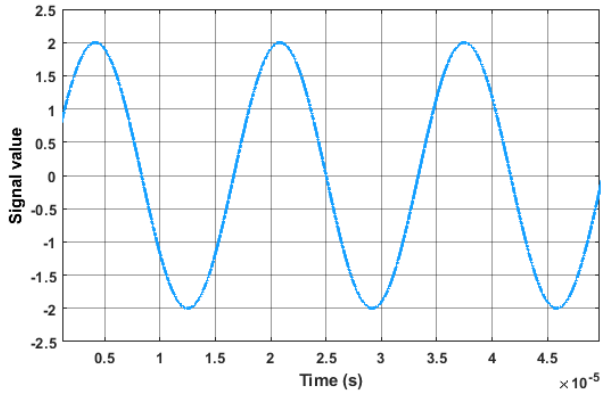
Model details for this and all other Simulink models used in this chapter can be found in Appendix A, but a brief explanation of the used blocks will be provided here.

Figure 3.5 illustrates how the parallel hardware tracks can be modelled in Simulink; each frame represents a single symbol and the number of elements is equal to the number of subcarriers used, in this case 64 (refer to the requirements in Section 2.5).

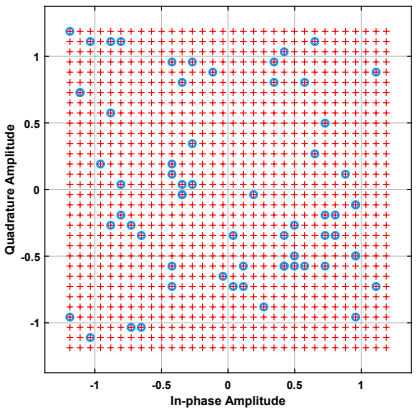
Three important signals have been visualized using the model, refer to Figure 3.6. The data generated in the 'Random Integer' block are converted into QAM symbols using 1024-QAM mapping, refer to Figure 3.6b. Subcarriers spaced at orthogonal frequencies are then mixed with these QAM symbols. An example of such a subcarrier has been depicted in Figure 3.6a.

After conversion to the analog domain, the signals are combined in the summation element, resulting in the 'regular' output of the iFFT in a typical transmitter. Figure 3.6c shows the real and imaginary parts of this waveform. After upconversion, this signal can be transmitted.

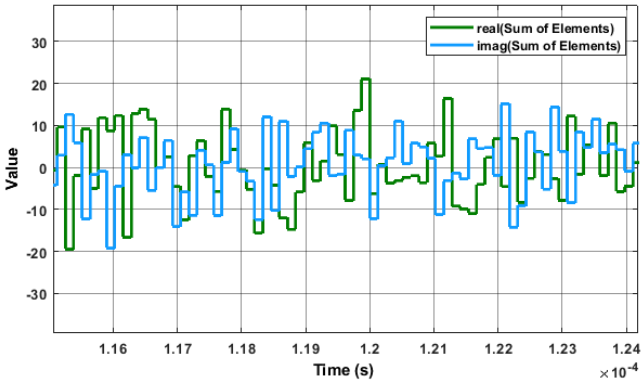
Now that the transmitter basics have been covered, implementation details will be treated. The first step is to determine the optimum number of subcarriers to convert using one DAC, the next section will address this topic.



(a) Visualization of a subcarrier to be mixed digitally with a generated QAM symbol.



(b) Visualization of 64 1024-QAM symbols in a constellation diagram. These QAM symbols will be mixed digitally with the orthogonally spaced subcarriers.



(c) After summation of the DAC outputs in the analog domain, the same signal has been restored that typically leaves the iFFT block in an OFDM transmitter.

Figure 3.6: Visualization of three important signals as designated with the characters A, B and C in the Figure of the Simulink model used to model transmitter signal flows, refer to Figure 3.5.

3.2 Analysis of PAPR reduction through frequency-domain splitting

This section will analyze to what extent the PAPR can be reduced when splitting up the OFDM signal in multiple groups of one or multiple subcarriers. Subsequently, the impact of the number of subcarriers at the input of a DAC on its requirements will be determined. As discussed in Chapter 2, the PAPR of an OFDM waveform increases as more and more subcarriers are combined and a Gaussian distribution is approximated.

As was also discussed in Chapter 2, statistical processes are involved in determining the PAPR. To make the comparison between the cases with varying numbers of subcarriers fair, an estimate of the typical PAPR of an OFDM signal will be determined by applying some clipping to yield an optimum in the balance between quantization and clipping distortion, limiting the dynamic range and thus the PAPR. As will be shown later, this optimum can be related to the BER of the processed signals, which makes comparing different cases easier[2].

The first part of the theoretical analysis in this chapter will therefore focus on this optimum in the digital-to-analog conversion, which will now be analyzed for a varying number of subcarriers converted by a

[2]As the number of subcarriers that are combined increases, the PAPR increases. However, the relative likeliness that the highest peaks occur decreases. This makes clipping generally more effective when the PAPR is high. In the end, what matters is the linearity that can be achieved (e.g. the BER), hence, an optimum between clipping and quantization distortion will be used that can be used to relate the PAPR to the BER.

DAC.

3.2.1 The optimum between clipping and quantization distortion

In a transmitter, the discrete time-domain waveforms are converted into the analog domain by a DAC. Fundamentally, the converted waveform has limited accuracy due to this quantization process. Linearity requirements in 5G transmitters dictate a minimal SNR to enable the receiver to decode transmitted symbols in presence of noise and distortion with an acceptable BER. Through scaling the number of quantization levels (e.g. the number of conversion bits), the quantization distortion can be reduced. Similarly, reducing the dynamic range while keeping the number of quantization levels constant will also reduce the quantization distortion.

Hence, reducing the dynamic range by clipping the outer amplitude levels will benefit quantization distortion at the cost of the clipped signal power. Figure 3.7 illustrates this concept.

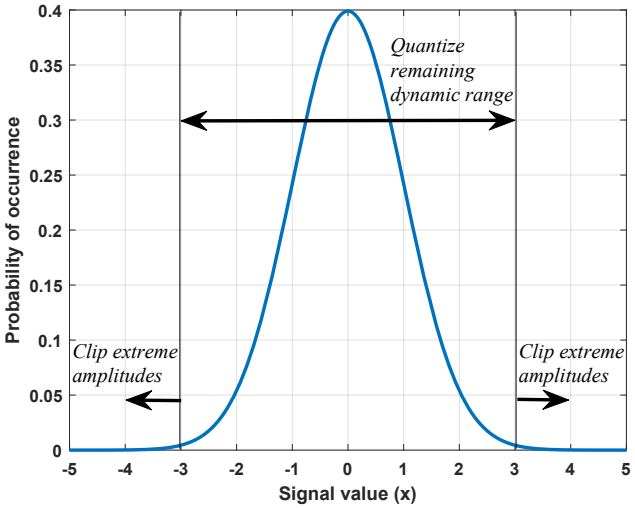


Figure 3.7: Tolerating some clipping distortion to limit quantization distortion may yield lower overall distortion. Visualized in the Figure is the distribution to which the I- and Q-parts of an OFDM waveform adhere, a normalized Gaussian distribution.

When considering in-band signal quality only, the optimum SNR can be found by balancing the contributions of clipping and quantization; this equilibrium will be used in the rest of this Chapter to determine the required resolution for the digital-to-analog conversion process. The reader should note, however, that the out-of-band signal leakage will be aggravated under very heavy clipping of the signal; if spectral mask requirements are particularly strict, a different trade-off between clipping and quantization may be required [8]. In this thesis, based on the requirements as presented in Chapter 2 and considering that the presented design is meant as a coarse proof-of-concept, analysis will be based on in-band signal quality only.

The following sections will first analyze the required DAC resolution for conversion of 'traditional' OFDM signals to form a baseline for the resolution improvement as a result of splitting up the DAC to reduce the PAPR. A verification of this theoretical analysis will be performed in Simulink. In subsequent sections, various methods to split up the OFDM signal into multiple subsignals along the lines of Figure 3.4 and their impact on the required conversion resolution will be identified and characterized. For each of these methods, verification will be performed in Simulink. The chapter will be concluded with an overview of these results.

3.2.2 Quantization and clipping distortion in OFDM signals

The goal of this Section is to analyze the impact of balancing clipping and quantization distortion on the PAPR for a typical OFDM signal to be able to determine the PAPR *improvement* when considering splitting the signal. To be able to compute what the effect would be of clipping the outer amplitude levels

on the distortion, the pdf of the signal can be used. As discussed in Chapter 2, when many subcarriers are combined, the I- and Q-components of the time-domain OFDM samples will closely approximate a Gaussian distribution[3].

The probability density function of an OFDM signal

In Chapter 2, the mathematical definition for a Gaussian distribution was provided in Equation 2.3 [7]. As discussed, both the I- and Q-components of an OFDM signal adhere to this distribution.

For now, it will be assumed that enough subcarriers have been combined to approach the Gaussian distribution with sufficient accuracy, this assumption will be verified later. Integration over the full pdf yields the total signal power P_{signal} which is thus equal to the variance (σ^2). The RMS value of the time-domain waveform is then equal to the standard deviation (σ).

Normalization of the signal power of the pdf in Equation 2.3 to unity will aid subsequent calculations and will allow for comparison with results as published in literature (for instance in [8]):

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{\left(-\frac{x^2}{2}\right)} \tag{3.1}$$

in which x designates the signal values of either the I- or Q-component of an OFDM signal.

The closed-form expression in Equation 3.1 will now be used as basis in determining the distortion as a result of clipping. Note the significance of the normalization of the signal power to unity; any distortion components that will be derived in subsequent sections using the result in Equation 3.1 can be related to a signal power of 1 to find the SNR.

Clipping distortion in OFDM signals

Referring back to Figure 3.7, if clipping is applied to the outer levels starting at some level $\pm\, x_c$, then the signal values exceeding these levels form the total power of the clipped signal portions, Figure 3.8 illustrates this concept for the positive signal values.

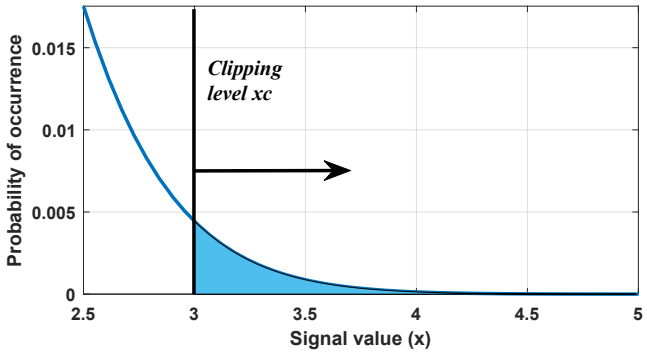


Figure 3.8: Zoomed part of Figure 3.7. Integration of the squared signal values out from the clipping level gives the power of the clipped signal portion. As an example, clipping is applied at $x = 3$.

As Figure 3.8 illustrates, the signal values beyond the clipping level can be integrated to determine the distortion as a result of clipping. By squaring the signal values, the distortion power can be determined. Mathematically, this can be described as in Equation 3.2.

$$P_{clipped} = 2 \cdot \int_{x_c}^{\infty} (x - x_c)^2 \cdot p(x) \cdot dx \tag{3.2}$$

[3]In this Chapter, when OFDM signals are mentioned in relation to a DAC without explicitly mentioning the I- or Q-component, still either the I- or Q-component is meant as each is converted by a separate DAC. In other words, analysis is based on normally distributed signals.

Note that it was assumed that the clipping distortion for the positive levels is equal to the clipping power of the negative levels, hence the factor 2.

In [8], a closed-form expression for Equation 3.2 after substitution of the Gaussian distribution has been provided:

$$P_{clipped} = (x_c^2 + 1) \cdot erfc\left(\frac{x_c}{\sqrt{2}}\right) - \frac{\sqrt{2} \cdot x_c \cdot e^{\frac{x_c^2}{2}}}{\sqrt{\pi}} \tag{3.3}$$

in which $erfc$ is the complementary error function.

Equation 3.3 will be evaluated to determine the power of the portion of the signal that has been clipped. Since the result in Equation 3.3 is based on a normalized signal power for x of 1, the signal-to-distortion ratio can be easily calculated by inverting the result in Equation 3.3. First, the distortion as a result of quantization will be examined.

Quantization distortion in OFDM signals

The remaining dynamic range after clipping is equal to $2 \cdot x_c$ and will be divided into quantization levels along the number of conversion bits the DAC provides. This process creates distortion as an instantaneous signal value will be rounded to the nearest quantization level available. In literature, this so-called quantization distortion is considered to be uniformly distributed over the Nyquist bandwidth for sine-waves if the number of bits is equal to 7 or higher [25]. For a lower number of bits, quantization distortion applied to sine-waves shows up as harmonic distortion.

For the (combination of) subcarriers as used in 5G systems, the number of bits required is typically higher than 7 bit to achieve sufficient SNR. In this case, the concept of uniformly distributed distortion is valid, which could be described by the following mathematical expression [25]:

$$P_{quantization} = \left(\frac{2 \cdot x_c}{2^N \cdot \sqrt{12}}\right)^2 \tag{3.4}$$

in which N stands for the number of bits used in the conversion.

In case the optimum resolution is lower than 7 bit, this analysis will be revisited to determine its impact on the validity of the results. Note that, similar to the result on clipping distortion, the result in Equation 3.4 is based on a normalized signal power of 1 and that the signal-to-distortion ratio can be determined by inverting the result in Equation 3.4.

Optimum resolution for conversion of OFDM signals

A balance will now be sought between clipping and quantization distortion. First, assumptions that were made in the derivation of both the expression for clipping distortion (Equation 3.2) and quantization distortion (Equation 3.4) will be summarized:

- A Gaussian distribution was assumed to sufficiently model the combination of independently modulated subcarriers into the I- and Q-components of an OFDM signal.
- The portion of the signal to be clipped was assumed to be symmetrically divided over the extreme positive and extreme negative amplitude levels.
- Distortion as a result of quantization was assumed to be accurately modelled with uniformly distributed distortion over the Nyquist bandwidth.

In addition to these assumptions, an extra assumption is necessary to find the optimum resolution. As both distortion mechanisms have different origins, careful consideration of their effect on the signal's linearity is required. To determine if both contributions can be summed, one has to know if the clipping distortion is uniformly distributed such that it can be added to the uniformly distributed quantization distortion [8].

Clipping the signal leads to higher order harmonic distortion due to the abrupt signal changes. These higher-order components are beyond the Nyquist frequency and will, without excessive Anti-Alias (AA) filtering, fold back into the Nyquist band. This process looks scrambled as the harmonic components are distributed across many different frequencies and the result approximates a uniformly distributed 'noise floor'[4]. An additional assumption will therefore be made to allow summing of the clipping and quantization components:

- Distortion as a result of clipping will be assumed to be accurately modelled with uniformly distributed distortion over the Nyquist bandwidth.

When considering *only* the contributions of quantization and clipping to the distortion of the signal, a specified BER can be achieved by increasing the resolution of the DAC and/or decreasing the clipped portion of the signal until the combination of clipping and quantization distortion drops below a particular level. This level is typically specified as a minimal SNR[5] which follows from analyzing the impact of Additive White Gaussian Noise (AWGN) on the BER of M-QAM signals such as in [26], refer to Equation 3.5.

$$SNR = \frac{P_{signal}}{P_{clipping} + P_{quantization}} = \frac{1}{(x_c^2+1)\cdot erfc\left(\frac{x_c}{\sqrt{2}}\right) - \frac{\sqrt{2}\cdot x_c \cdot e^{\frac{x_c^2}{2}}}{\sqrt{\pi}} + \left(\frac{2\cdot x_c}{2^N\cdot\sqrt{12}}\right)^2} \tag{3.5}$$

The results in Equations 3.2 and 3.4 can be plotted to visualize the distortion components, see Figure 3.9. Following the result in Equation 3.5, to achieve a particular SNR, the inverted combination of clipping and quantization distortion should be lower than the desired SNR.

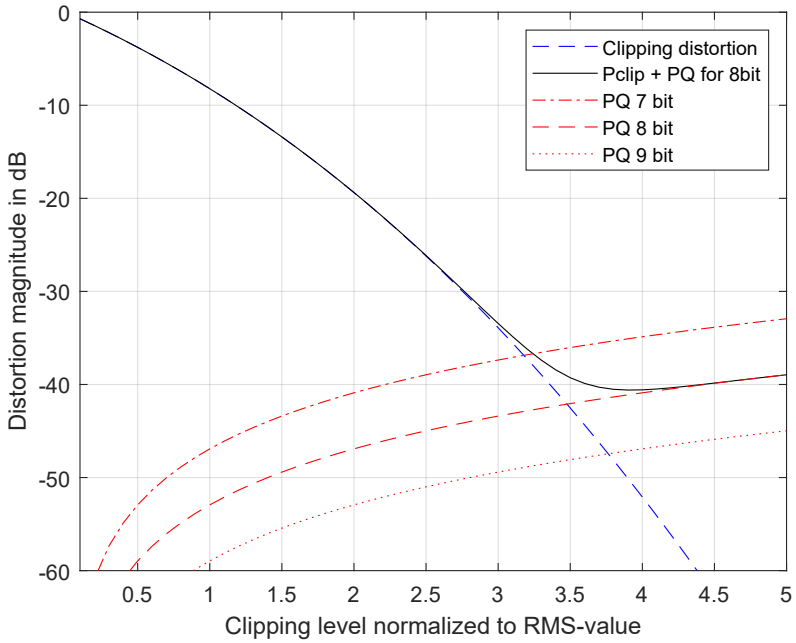


Figure 3.9: Distortion with respect to a normalized signal as a result of clipping (Pclip) and quantization (PQ). An example of a sum of both components has been included (black line).

[4] In a way this is very similar to quantization distortion, where the many harmonic tones also appear scrambled and fold back to form a noise floor.

[5] Note that the term SNR is used to reflect on the uniformly distributed distortion for both the clipping and quantization components as if both form a noise floor. More technically correct would be the use of the Signal-to-Noise-and-Distortion Ratio (SINAD).

Verification in Simulink

A Simulink model has been designed to verify the analysis presented in the previous Section, refer to Appendix A. The basis of this Simulink verification is replication of the theoretical curve of the combined clipping and quantization distortion as included in Figure 3.9 (for 8 bit resolution). Note that for this curve a Gaussian distribution was used, which is asymptotically approached through increasing the number of subcarriers. To keep simulation time within feasible limits, it has been assumed that 64 subcarriers sufficiently approach the Gaussian distribution. The results can be found in Figure 3.10.

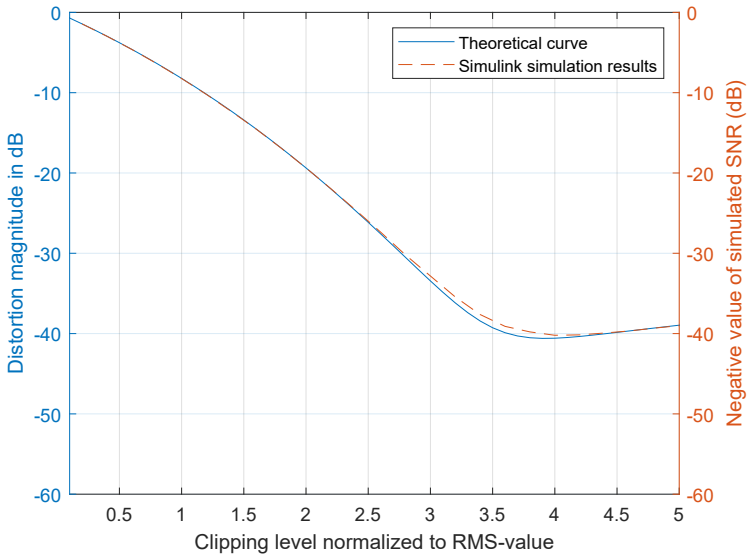


Figure 3.10: Comparison of the theoretical distortion components and Simulink simulation results. The light blue line is identical to the black line as visualized in Figure 3.9.

Several remarks regarding Figure 3.10 are in place:

- For low normalized clipping levels (e.g. up to $2.5 \cdot RMS$), the theoretically predicted distortion closely matches the simulated SNR.
- For high normalized clipping levels (e.g. from $4.5 \cdot RMS$ upwards), the same is true.
- In between these regions, the theoretical curve and simulated result differ up to 2 dB around $3.5 \cdot RMS$. Noting that this discrepancy occurs at the point where the distortion components due to clipping and due to quantization are expected to be roughly of equal magnitude, the assumption that both components could be linearly added does not accurately hold. Considering that the results are accurate within 2 dB nonetheless, no extensive analysis into the correct summation of both distortion components was deemed necessary.

Peak-to-average power ratio of OFDM signals

The results as plotted in Figure 3.10 also provide information on the PAPR of a typical OFDM signal. As an example, based on the settings as used to plot the results in Figure 3.10, the optimum amplitude level to apply clipping would be equal to $3.8 \cdot x_{RMS}$ as here the combined distortion is lowest. This value enables determining the PAPR through the relation as included in Equation 2.5:

$$PAPR = \frac{|x_{peak}|^2}{x_{RMS}^2} = \frac{(3.8 \cdot x_{RMS})^2}{x_{RMS}^2} = 3.8^2 \tag{3.6}$$

Typically, the PAPR is specified in decibels, in which case the result would be equal to 11.6 dB.

3.2.3 Quantization and clipping distortion in split-up OFDM signals

This Section marks a turning point as the theoretical analysis will proceed to explore the novel system concept as presented in this thesis.

Both the I- and Q-components of an OFDM signal can be split-up in multiple signals to reduce their PAPR. As a result, the mathematical analysis as presented in Section 3.2.2 is no longer valid; the assumption that a Gaussian distribution was applicable has been violated. To be able to compute the reduction in PAPR and the resolution that is required for optimal digital-to-analog conversion, another method has to be used to determine the pdf.

As the aim of this section is to determine the optimal resolution for a varying number of subcarriers converted per DAC, it is instructive to have a look at the pdf of a single subcarrier first and analyze what happens when multiple are combined.

The pdf of an OFDM signal with a small number of subcarriers

Mathematically, a subcarrier is a sine-wave running at its subcarrier frequency with an amplitude determined by its QAM symbol value. Figure 3.11 approximates the pdf of a sine-wave with amplitude 1 in histogram format with a limited number of bins.

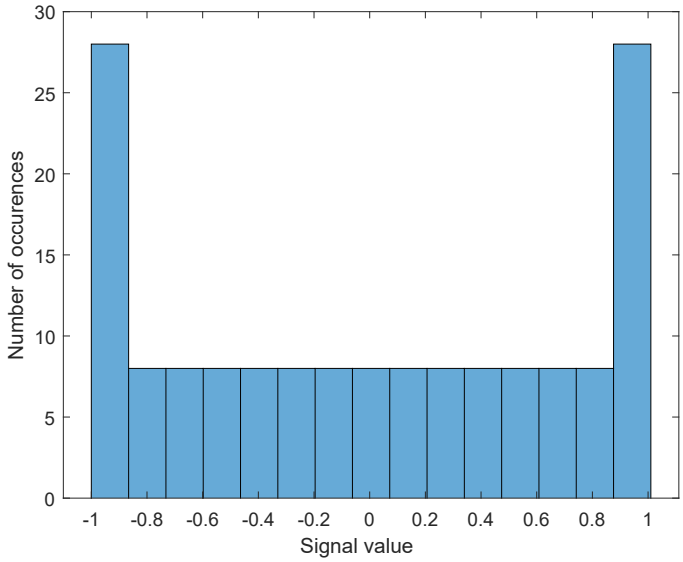


Figure 3.11: A histogram of a single subcarrier (a sine wave) with an amplitude equal to 1. The distribution of a sine wave has the shape of a bathtub.

A closed-form expression for the pdf as approximated in Figure 3.11 is [27]:

$$p(x) = \frac{1}{\pi \cdot \sqrt{1 - (x/A)^2}} \tag{3.7}$$

in which A is the amplitude of the sine-wave, which in case of QAM modulated signals is determined by the QAM symbol value.

To find a pdf for multiple subcarriers combined, convolution in the amplitude-domain can be performed as each subcarrier has identical statistical properties and is independent [7]:

$$p_{combined}(x) = p_1(x) \circledast p_2(x) \circledast ... \circledast p_n(x) \tag{3.8}$$

In order to simplify the analysis, it will now be assumed that the subcarrier can also be sufficiently accurately modelled with a uniform distribution stretching up to the sine-wave's amplitude:

$$p(x) = H(x + A) - H(x - A) \tag{3.9}$$

in which H is the Heaviside step function.

To perform the convolution, the result in Equation 3.9 will now be transformed to the LaPlace domain using the LaPlace transformation:

$$P(s) = \int_0^\infty e^{-st} \cdot p(x) \cdot dx \tag{3.10}$$

Resulting in:

$$P(s) = \left(\frac{1}{s} - \frac{-e^{As}}{s}\right) \tag{3.11}$$

If all subcarriers would take on the same amplitude (QAM symbol), finding the pdf of the combined signal would be a matter of raising the result in Equation 3.11 to the power of the number of subcarriers and transforming back to the time-domain. However, the amplitude of each carrier is a statistical process in its own. As the QAM amplitude values of the subcarriers are discrete samples in a set sized $\sqrt{M}$ elements, it is difficult to substitute this distribution into the result of Equation 3.11. Instead, the combined pdf will be computed numerically for a large set size and the resulting pdfs will be normalized and combined to closely approximate the final pdf of the OFDM signal. Mathematically, this results in the following steps:

1. Determine values for the QAM amplitude of each subcarrier by randomly picking, with replacement, from a set $\{-\frac{\sqrt{M}}{2}, ..., -2, -1, 1, 2, ..., \frac{\sqrt{M}}{2}\}$.
2. Compute the combined pdf for this run by evaluating:

$$P_{combined}(s) = \prod_{i=1}^{n} \left(\frac{1}{s} - \frac{-e^{A(i)\cdot s}}{s}\right) \tag{3.12}$$

3. Transform the result of Equation 3.12 back to the amplitude-domain using the inverse Laplace transform.
4. For a set number of symbols, redo steps 1, 2 and 3. Since the results are multiple runs for the same probability variable, they can now be averaged to find a close approximation to the pdf of a subset of the OFDM signal.

Using the numerically approximated pdf, the clipping and quantization distortion components can now be determined.

Using the numerically approximated pdf to determine distortion components

The results of the previous section have been determined numerically using MATLAB, resulting in a discrete (sampled) expression for the pdf. To find the clipping distortion using Equation 3.2, integration will have to be performed on these (sampled) numerical results of the previous section. An option to do so is through the use of the midpoint rule. It estimates the area under a sampled curve by considering two subsequent sample moments a and b by fitting a rectangle:

$$\int_a^b f(x)dx \approx (b - a)\, f\left(\frac{a + b}{2}\right) \tag{3.13}$$

After applying the midpoint rule to Equation 3.2 and using Equation 3.4 to calculate the quantization distortion of the remaining dynamic range, the optimum resolution for the DAC can be determined.

The distortion curves for a varying number of subcarriers per DAC have been visualized, see Figure 3.12.

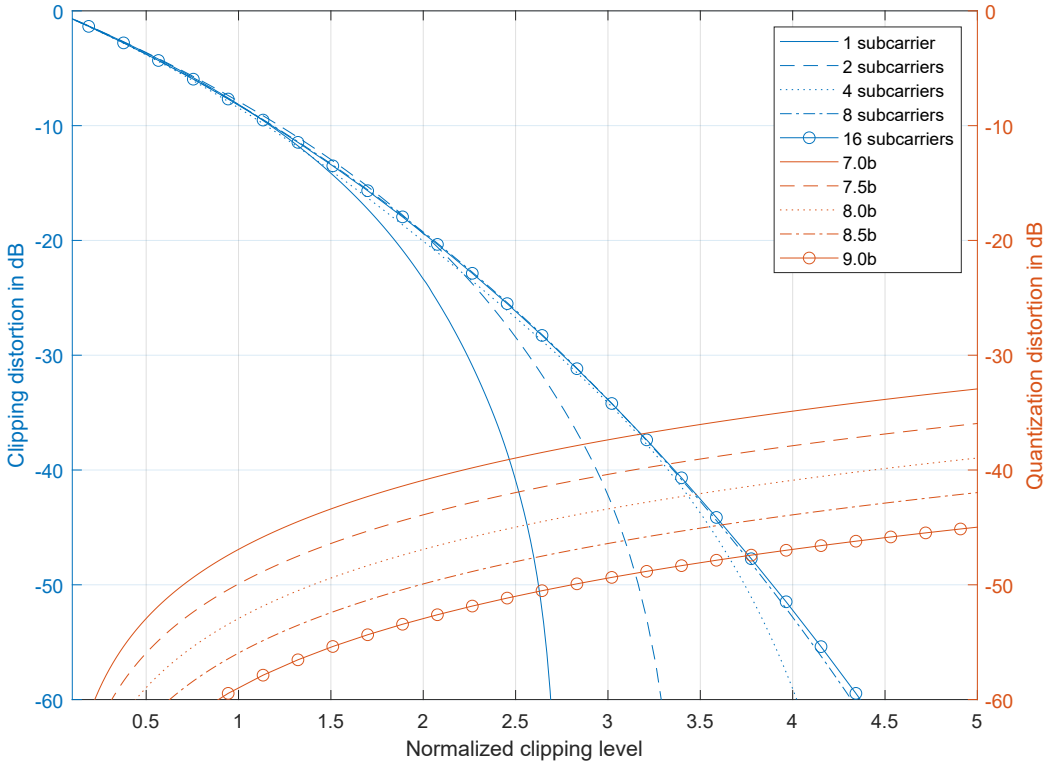


Figure 3.12: Distortion components for a varying number of subcarriers per DAC. Quantization distortion curves have also been included, but do not depend on the number of subcarriers, only on the resulting dynamic range.

Several remarks regarding the plot in Figure 3.12 are in place:

- With an increasing number of subcarriers combined, the clipping distortion for the same clipping level increases and eventually saturates. Eventually, enough subcarriers have been combined to approach the Gaussian distribution as visualized in Figure 3.9. The increment in clipping distortion for the same clipping level between 8 and 16 subcarriers is at maximum smaller than a few dB.
- For a lower number of subcarriers combined, the clipping distortion drops faster as the clipping level is increased. This is in line with the reduced PAPR, more signal power is concentrated closer to the normalized RMS value of the signal.
- Note that the analysis implicitly assumed that the quantization distortion does not depend on the number of subcarriers, but is directly related to the clipping level and the resulting dynamic range.

Optimum resolution for converting a limited number of subcarriers with a DAC

Similar to the OFDM case, when considering only the in-band signal linearity, for each of the distortion curves, the optimum number of bits in terms of distortion can now be determined by considering the combination of clipping and quantization distortion.

As an example, the optimum resolution for 1024-QAM modulated signals that should have a BER less than 10^{-4} has been plotted in Figure 3.13 for a varying number of subcarriers. Through analysis on AWGN and related SNR and BER as in [26], the required SNR for a typical BER is known. Depending on the number of subcarriers desired, the correct distortion curve from Figure 3.12 is selected and combined with the correct quantization curve until the combined distortion drops so that the desired SNR has been achieved.

Figure 3.13 indirectly provides information on both the PAPR reduction and the gain in resolution required for each DAC. Table 3.1 provides an overview of these improvements. The PAPR values as included

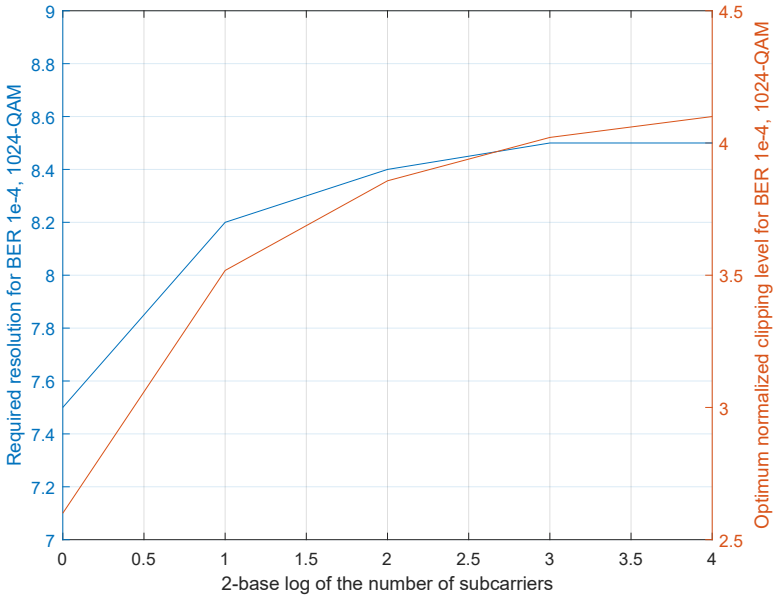


Figure 3.13: Optimum resolution and normalized clipping level to reach a BER of 10^{-4} with 1024-QAM for a varying number of subcarriers.

in the table have been calculated using Equation 2.5 from Chapter 2 in a similar way as the result in Equation 3.6 had been established.

Table 3.1: Overview of PAPR improvement and gain in DAC resolution when splitting up an OFDM signal into smaller signals. Resolution optimum is to achieve a BER of 10^{-4} with 1024-QAM.

Number of subcarriers per DAC	Optimum PAPR (Figure 3.13)	Optimum DAC resolution (Figure 3.13)	Change in PAPR wrt Gaussian case	Gain in DAC resolution wrt Gaussian case
1	$2.60^2 = 6.8$	7.5 bits	2.49x lower (-4.0 dB)	1.0 bit
2	$3.52^2 = 12.4$	8.2 bits	1.36x lower (-1.3 dB)	0.3 bit
4	$3.86^2 = 14.9$	8.4 bits	1.13x lower (-0.5 dB)	0.1 bit
8	$4.02^2 = 16.2$	8.5 bits	1.04x lower (-0.2 dB)	*NA*

Verification in Simulink

Reusing the model that verified the analysis of the optimum in clipping and quantization distortion for OFDM signals, the results in Figure 3.12 have been verified in Simulink. Figure 3.14 compares the estimated results as plotted in Figure 3.12 to Simulink simulation results. Note that the distortion components in Figure 3.12 have been visualized separately. In the Simulink model, one value for the SNR can be determined after applying both clipping and quantization, refer to the model details in Appendix A. Therefore, the combination of clipping and quantization distortion will be verified. The quantization resolution has been set to 8 bit.

Several remarks regarding Figure 3.14 are in place:

- In general, the discrepancy between the Simulation results and the theoretical approximation is larger than in the Gaussian analysis as visualized in Figure 3.10.
- Comparing the Simulink simulation results and the estimated distortion components, systematically lower clipping distortion for a constant clipping level can be identified in the Simulink simulation results. The contributions from quantization distortion that can be identified in the Figure match up to a few dB accuracy.

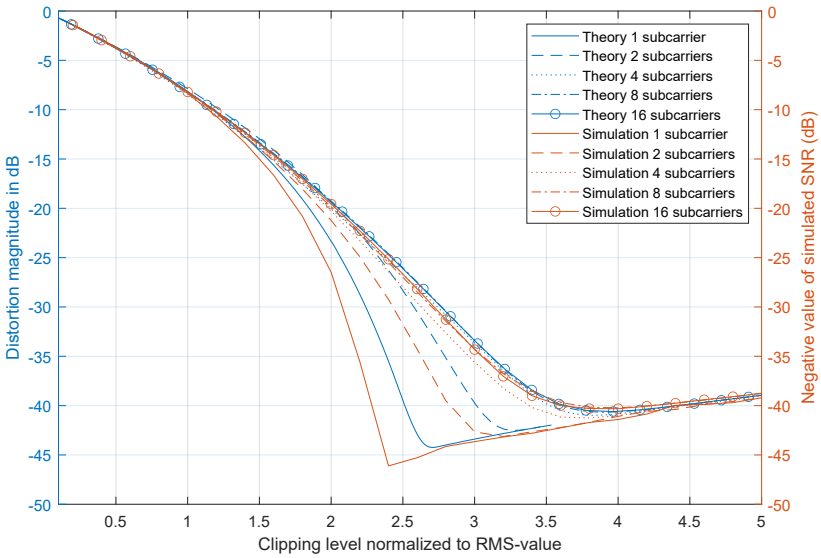


Figure 3.14: Comparison of the estimated distortion components to Simulink simulation results. Quantization resolution has been fixed for all curves to 8 bit. The number of subcarriers converted per DAC is varied, but note that the total number of subcarriers processed is always equal to 64.

- The reason for this difference is oversimplification of the mathematical analysis. Since the pdf of the subcarrier sine waves was approximated with a uniformly distributed pdf, there exists mismatch between the RMS values used in the theoretical approximation and those valid for sine-waves. The RMS value of a uniformly distributed pdf is equal to that of a triangular wave and thus $x_{peak}/\sqrt{3}$, leading to a mismatch factor in the x-axis values of roughly $\sqrt{3}/\sqrt{2} = 1.2$ for the clipping distortion components, which can indeed be identified in Figure 3.14.
- As a result, the gain in both PAPR and resolution has been slightly underestimated in Table 3.1.

The PAPR reduction as a result of splitting up the OFDM signal along the frequency-domain has now been determined and verified using Simulink simulations. Before proceeding to exploring how the reduced PAPR can be exploited in a transmitter, a short discussion on an alternative frequency-domain splitting approach will be included.

3.2.4 Directly converting QAM amplitude levels

An additional implementation alternative to splitting up the OFDM signals into groups of subcarriers is to perform direct conversion of the subcarrier amplitudes, e.g. having the DAC output square waves. Through conversion of amplitude levels only, both the DAC's sample rate and resolution can drop. In case of a 1024-QAM signal, the number of bits required for conversion of both the I- or Q-component of a single subcarrier would be equal to $log_2(\sqrt{1024}) = 5$, which is a significant reduction to the rough 9 bit required in conversion of the full OFDM signal.

There are, however, two major problems with this approach which is why it will not be considered further in this Chapter. First, similar to arguments used in the discussion on integration of subcarrier mixing in Section 3.1.1, to prevent data from overlapping, subcarrier spacing would have to be integrated digitally if multiple subcarriers should be converted per DAC. This would again increase the number of different amplitude levels required and hence the DAC's resolution requirement.

Second, to meet the SEM as defined in Chapter 2, pulse shaping is typically applied in transmitters. Through pulse shaping, abrupt transitions between samples are softened to avoid 'jumps' in the signal which cause harmonic distortion, leading to violation of the SEM. Pulse shaping is applied in the digital domain; if the DAC is optimized to convert only QAM amplitude levels, the number of quantization levels might not be sufficient to apply pulse shaping with enough accuracy.

3.2.5 Overview of PAPR reduction analysis as a result of frequency-domain signal splitting

Section 3.2 analyzed the impact of splitting up an OFDM signal along the frequency domain on the PAPR of the resulting signals. The optimum in clipping and quantization distortion served as basis for this theoretical analysis. For the 'traditional' OFDM case, the combination of the calculated clipping distortion in Equation 3.2 and the quantization distortion in Equation 3.4 resulted in the distortion components of Figure 3.9.

Subsequent analysis on the effect of converting fewer subcarriers in each DAC resulted in the distortion components as plotted in Figure 3.12. To provide an indication, the relative improvement of both the PAPR and required conversion resolution for a 1024-QAM signal with a BER of 10^{-4} had been provided in Table 3.1. This specific, yet realistic, example indicated that to obtain significant reduction (more than a few dB) in both PAPR and required resolution, no more than 1 subcarrier should be converted per DAC.

Along the requirements for 5G transmitters as dictated in the 3GPP requirements [2], hundreds of DACs would be required to obtain the desired output signal if only 1 subcarrier would be converted with a single DAC. Even if the marginal PAPR decrease of a factor 1.1 is accepted and a maximum of 4 subcarriers are converted per DAC, the number of DACs required is excessive. These results provide a first indication that the added benefit of frequency-domain splitting within the boundary of 5G transmitter requirements is marginal. Putting aside the reservations on the amount of hardware required for the moment, the benefit of the reduced PAPR as a result of signal splitting on the transmitter's implementation will now be analyzed.

3.3 Exploiting alleviated PAPR characteristics in transmitters

The overview presented in Section 3.2.5 has shown that the PAPR can be reduced with respect to the 'traditional' OFDM case by converting only a few subcarriers per DAC. This Section will discuss how the transmitter should be designed to benefit from this PAPR reduction. Two distinct areas will be considered. First, design considerations related to the digital-to-analog conversion process will be considered and the impact of the lower PAPR on the DAC's specifications will be analyzed. Second, it will be investigated if the inefficient PA can be omitted by making use of the parallel DAC outputs.

3.3.1 Benefiting from signal splitting in the digital-to-analog conversion process

Within the boundary of the 5G transmitter requirements as presented in Chapter 2, the digital-to-analog conversion process can be given shape. The subsequent sections will analyze DAC designs possible within the design space and how the designs suitable for use with 5G NR benefit from reduced PAPR and a reduction in resolution.

DAC architectures

DACs convert some sampled digital signal into a time-continuous waveform that has to be delivered to a load impedance at some power level. A physical reference quantity is used to derive accurate output levels that can be selected depending on the incoming digital code. In some cases, when the designer can also implement the hardware preceding the DAC, there is an additional freedom in interpreting the digital code representation.

The reference quantity from the physical domain on which the DAC is based is limited to one of four options [25]: voltage, current, charge or time. Multiple output levels can be created from this reference quantity through the use of a unary, binary or segmented structure. A full analysis on all possible DAC implementations is beyond the scope of this thesis. However, some general remarks on each domain will be provided and their compatibility with the 5G design space will be discussed.

Voltage-domain DACs
Voltage-conversion DACs make use of resistor strings to divide a reference voltage into smaller portions. Due to fundamental limitations in matching of the resistors, resolution is roughly limited to 10 bit [25]. In a unary voltage-architecture, a resistor string is typically used, while binary architectures make use of so-called R-2R resistor strings. Several techniques exist to further improve linearity and accuracy, mostly

based on calibration techniques. Voltage-domain DACs require a buffer to drive a load and are most widely used for low-performance conversion.

Current-domain DACs
Multiple weighted copies of a 'golden' current-source reference together comprise a typical current-domain DAC. Current-domain DACs can be classified into two distinct architectures: current-switching and current-steering devices. In current-switching devices, unused currents are switched off to limit dissipation. However, charging and discharging the current sources in switching is slow, limiting the bandwidth of such designs. In addition, to reduce the code-dependent output impedance, a buffer is required for this type of DAC, limiting the output swing and increasing power dissipation.

Current-steering DACs solve these typical problems by only steering the currents to either the supply and the output in a single-ended architecture or between the positive and negative output in a differential output. Consequently, all the current flows all the time (class A conduction), leading to relatively poor efficiency. As no large capacitance has to be charged or discharged when changing codes, current-steering DACs provide fast performance and are the topology of choice for high bandwidth, high performance DACs.

Charge-domain DACs
In charge-domain DACs, charge is the information carrying quantity. Through a combination of charging of select capacitors in a capacitor bank and subsequent amplification of the resulting voltage, digital codes can be converted to an output current. Charge-domain DACs are very efficient due to the small charge packages required in the conversion. Although a buffer is required, slightly limiting the efficiency, these DACs are the topology of choice for efficient implementations. Similar to voltage-domain DACs, because a buffer is required, the bandwidth these DACs can achieve is limited.

Time-domain DACs
Signal information is contained in the succession of switching moments. Most of the time-domain DACs make use of 1-bit systems, which benefits linearity as there are no inherent errors in the transfer curve. However, because information is contained in the succession of switching moments, the bandwidth that can be achieved is lower. Time-domain DACs are the architecture of choice for superior linearity at low frequencies.

Compatibility with the 5G design space
The exact bandwidth of 5G baseband signals varies, but is in the range of hundreds of MHz [2]. In realistic systems, the DAC has to be able to process some overhead in bandwidth to allow for feasible AA-filters[6]. In many cases, oversampling and interpolation are applied to limit spectral leakage, further increasing the bandwidth demands. As a result, bandwidth requirements close to a GHz are to be expected.

Such bandwidth requirements make the design of a buffered DAC very challenging as the buffer adds an additional pole to the DAC's transfer function, limiting its bandwidth. This demand rules out the use of buffered DACs and dictates the use of current-steering DACs.

It should be noted that the bandwidth requirement can be alleviated by performing subcarrier mixing partly in the analog domain along the lines of the technique as proposed in [24] and explained in Figure 3.2. However, this technique results in a significant increase in complexity as additional mixing has to be performed in the analog domain.

Effects of resolution reduction in current-steering DACs

To analyze what the benefits of the reduction in resolution as presented in Table 3.1 would be to current-steering DACs, typical architectures used for current-steering digital-to-analog conversion will be analyzed first.

[6]Typically, oversampling and subsequent FIR interpolation are used to place the first aliases as a result of digital-to-analog conversion sufficiently far away to make anti-alias filtering feasible. For more detailed information on AA-filtering, the reader is referred to treatment in literature, such as in [5].

As with most DACs, current-steering devices can be designed to be either single-ended or differential. For both implementations, the reference quantity can be copied using one of three options: a unary architecture, a binary architecture or a combination of both called the segmented architecture. Figure 3.15 combines each of these options into a single figure. Shown in the Figure is a segmented DAC architecture, which consists of a unary current-source section (here in red) and a binary section (in green). Depending on the implementation, either the single-ended load is connected where only useful current flows to the load and the rest is dumped in the supply rail, or the differential load is selected, where all current flows to the load all the time and the voltage difference between the two sides provides the output quantity.

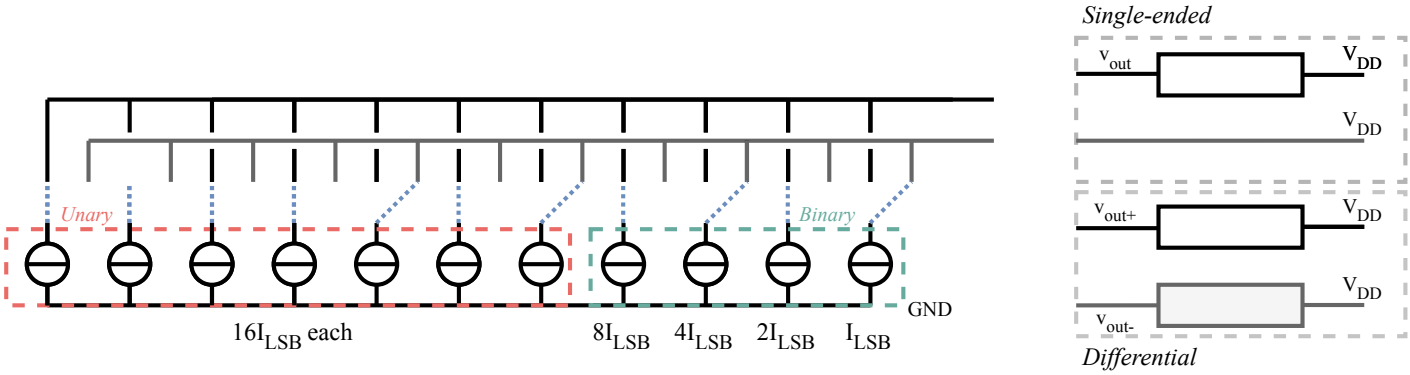


Figure 3.15: A schematic overview of a 7-bit segmented current-steering DAC. Depending on the implementation, either the single-ended load or the differential load is to be connected to the current-source array.

The use of segmentation is a compromise between excessive amounts of hardware and large Differential Non-Linearity (DNL) and Integral Non-Linearity (INL) errors as a result of transistor mismatch [25]. The DNL describes the deviation between two analog values corresponding to adjacent digital values, while the INL describes the deviation between the actual analog output measured for a code and the ideal output value. On the one hand, a full binary implementation would reduce the amount of current sources from 2^N to N but the large 'jumps' in current when sources are steered lead to severe DNL [25]. On the other hand, a full unary implementation requires the maximum number of sources, but has relatively good linearity performance. Basically, the trade-off between the size of the binary and unary partitions is related to the linearity demands of the architecture in which the converter is used.

To be able to evaluate the impact of the resolution reduction on the required number of current sources, its impact on the segmentation trade-off has to be known.

Analysis of optimal segmentation in current-steering DACs
The largest DNL error occurs when the current-steering DAC steers from a binary section to the next unary source, because these sources are largest. The variance of this DNL error has been characterized in [28] as:

$$\sigma^2_{DNL,\ B\dashrightarrow U} = \frac{\sigma^2_{I_{unary}} + \sigma^2_{I_{all\ binary}}}{I^2_{LSB}} = 2^{2N_{binary}} \cdot \frac{\sigma^2_{I_{unary}}}{I^2_{unary}} + \left(2^{N_{binary}} - 1\right)^2 \cdot \frac{\sigma^2_{I_{all\ binary}}}{I^2_{all\ binary}} \tag{3.14}$$

in which $\sigma_{I_{unary}}$, $\sigma_{I_{allbinary}}$ represent the standard deviation of a unary source and of the sum of binary sources respectively, I_{LSB}, I_{unary}, $I_{allbinary}$ are the LSB current source size, unary source size and the sum of the binary source sizes respectively and N_{binary} is the resolution of the binary section.

The physical origin of the mismatch effects can be substituted to characterize the design parameters that impact the DNL [25]. If the assumption is made that the gate area of the binary section is of equal size as the gate area for a single unary source (which is typically roughly true as the binary sources summed provide current of almost equal magnitude as a single unary source, refer to Figure 3.15), then:

$$\sigma_{DNL,\ B\dashrightarrow U} \approx 2^{N_{binary}+\frac{1}{2}} \cdot \frac{\sigma_{I_{unary}}}{I_{unary}} \approx \frac{2 \cdot A_{V_{th}} \cdot 2^{N_{binary}}}{\sqrt{(WL)_{unary}} \cdot (V_{GS} - V_{th})} \tag{3.15}$$

in which $A_{V_{th}}$ is a technology conversion constant, WL is the current source's active area and $(V_{GS} - V_{th})$ is the current source's overdrive voltage.

Equation 3.15 allows estimating the division between unary and binary sources in the 'traditional' OFDM case of 8.5 bit and in the individual DACs after splitting where 7.5 bits would suffice (Table 3.1). The spread in DNL as calculated by Equation 3.15 is not dependent on the DAC's resolution. Since the limit to this spread is set by the linearity specifications, the binary resolution will not scale when the DAC's resolution does. A global decrease in resolution therefore scales the number of unary sources. To provide an indication, using a current source area of $25\ \mu m^2$, an overdrive voltage of $300\ mV$ and a mismatch factor $A_{V_{th}} = 3\ mV\mu m$, the resolution for the binary partition can be estimated for both cases to be a maximum of 5 bit [25].

If the binary partition is set to a size of 5 bit, then the reduction in area as a result of the reduction in resolution from 8.5 to 7.5 bit can be estimated. In making $2^{8.5} = 362$ levels, $2^5 = 32$ of these can be binary. For the other levels, $\frac{(362-32)}{32} = 10.3$ unary sources of $32I_{LSB}$ (hence 11) are required. The total occupied area is then equal to 12 unary sources. A similar analysis for the 7.5 bit case results in an area of 6 unary sources. Assuming the current sources are of equal size, the area reduction per DAC is therefore equal to a factor 2.

Considering that, within the design space of 5G transmitters, the number of subcarriers required is in the hundreds or thousands, the required area becomes excessive. Already with converting 2 subcarriers in a single DAC, the area of the smaller DACs combined would be equal to the original area of the DAC converting the complete I- and Q-components of an OFDM signal. The use of signal splitting along the frequency-domain to benefit the DAC requirements is therefore unattractive.

The next section will explore the use of signal splitting to generate the required output power and benefit efficiency through omission of the inefficient PA stage.

3.3.2 Omitting the power amplifier

A major transmitter requirement is to maintain sufficient output power to allow the receiver to decode the incoming signal at the other end of the wireless channel under presence of noise and distortion. In mmWave applications, providing sufficient amplification with acceptable efficiency is difficult due to the stringent linearity requirements and high PAPR. After splitting in the frequency-domain, multiple DACs concurrently provide a fraction of the required power and at reduced PAPR. This section will explore if these DACs can replace the inefficient PA stage.

Following from the specification for the minimum output power as provided in Chapter 2, an average output power of 10 dBm and 10 dB PAPR will be assumed, leading to a peak power of 20 dBm.

First, the required output current for the DACs in case of a 'traditional' OFDM transmitter will be calculated. After that the impact of splitting up the signal along the frequency-domain on the required output current for each DAC will be evaluated. Then, the need for a PA will be evaluated.

Output current of the DAC in an OFDM transmitter

An estimate of the required output current of a DAC will be provided when no PA is present in the signal chain of a 'traditional' OFDM transmitter. As the peak power requirement is most stringent, it will be used to perform the subsequent analysis.

To generate 20 dBm peak output power in the $50\ \Omega$ antenna load, assuming a sinusoidal signal, the peak antenna current should be equal to roughly:

$$I_{peak-ant} = I_{RMS} \cdot \sqrt{2} = \sqrt{\frac{P_{ant-peak}(W)}{Z_{load}}} \cdot \sqrt{2} \approx 63\ mA \tag{3.16}$$

Typically, to match the antenna's impedance to the preceding hardware, impedance matching is applied, further increasing this current. Here, impedance matching will not be considered as its exact implementation details are dependent on factors related to the PA's exact implementation.

In an OFDM transmitter, as visualized in Figure 3.1, two DACs are generally used; one for the I-phase component and one for the Q-phase component. Quadrature-separating the peak current as provided in Equation 3.16 yields:

$$I_{peak-DAC-OFDM} = \frac{63\ mA}{\sqrt{2}} = 44.5\ mA \tag{3.17}$$

The result in Equation 3.17 holds for each of the DACs. Note that incorporating impedance matching and losses of the other circuits into this equation would further increase the required current.

Next, the effect of signal splitting on the required output current will be determined.

Output current of the DACs after signal splitting

Similar analysis on the required output current after signal splitting for each DAC will now be conducted.

On average, the sum of the DAC output currents should amount to an output power of 10 dBm (refer to the specifications in Chapter 2), resulting in an average output current of 14 mA in a $50\ \Omega$ load. With every doubling of the number of DACs because of signal splitting, the average output current of each DAC drops with a factor $\sqrt{2}$ as a result of power conservation.

The resulting *maximum* current can be determined using the PAPR analysis as conducted in the beginning of this Chapter; if 16 or more subcarriers are combined in a single DAC, the PAPR of the combined signal would be equal to 10 dB. For fewer subcarriers converted per DAC, it drops along the results in Table 3.2.5. For 10 dB PAPR, the peak current at peak output power is a factor $\sqrt{10} \cdot \sqrt{2}$ higher than the average RMS value, due to the 10 dB PAPR and the conversion factor for the RMS value of a sine-wave.

Equation 3.16 relates the number of DACs to their peak current. Note that this Equation assumes that the PAPR at the input of each DAC is equal to 10 dB. For a lower number of subcarriers converted per DAC, the $\sqrt{10}$ factor will have to be adapted according to the result in Table 3.2.5.

$$I_{peak-DAC-splitting} = \sqrt{\frac{P_{ant-avg}(W)}{Z_{load}}} \cdot \left(\sqrt{2}\right)^{-log_2(n)} \cdot \sqrt{10} \cdot \sqrt{2} \tag{3.18}$$

in which the power '$-log_2(n)$' indicates that the power per DAC halves and hence the RMS current decreases with a factor $\sqrt{2}$ with every doubling of the number of DACs and the total number of DACs (sum of I- and Q-DACs) is given by n.

Evaluating the need for PAs after signal splitting

Equations 3.17 and 3.18 provide the peak output currents of the DACs for the traditional OFDM transmitter and transmitter based on signal splitting respectively. Now, in evaluating the need for a PA, it is not so much the ability of the DAC to provide the required output current (it can be scaled by combining multiple DACs in parallel for instance) that is limiting the ability to omit the PA, but the linearity of the mixers following it [5]. In an OFDM system, the mixer handles a relatively small signal and the PA provides the required power gain.

Performing upconversion on signals with excessively large swings will introduce strong non-linearity mainly as a result of gain compression, violating the strict linearity requirements for 5G NR transmitters as discussed in Section 2.5. Appendix B provides an overview of non-linearity mechanisms in mixers. The upper limit to the input current's magnitude with which the mixer can still upconvert the signal with sufficient linearity has to be known to be able to evaluate the need for PAs. More detailed analysis into this limit will now be provided.

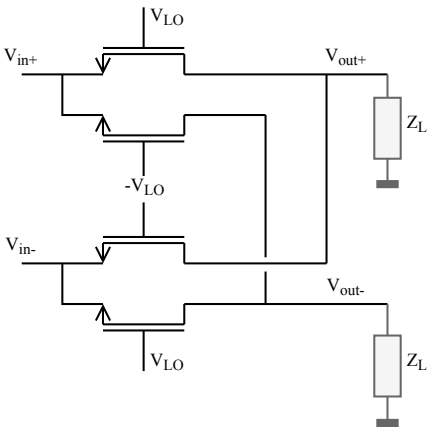


Figure 3.16: Schematic of a fully differential current-driven passive mixer.

Due to the fact that currents are readily available at the DAC outputs and due to their superior linearity performance, the use of current-driven passive mixers has been investigated in more detail. Figure 3.16 provides a schematic of a typical current-driven passive mixer.

Extensive simulations into the linearity of current-driven passive mixer implementations in 22FDX technology have been conducted in a parallel project within the Integrated Circuit Design group at the University of Twente [29]. Results show that the implementation as depicted in Figure 3.16 based on N-type Metal Oxide Semiconductor (Field-Effect Transistor) (NMOS) devices has its 1 dB compression point at -8 dBm (referred to 50 Ω), which would come down to a peak current of 2.5 mA. Even without considering any back-off from the 1 dB compression point, this would require 630 DACs according to the result in Equation 3.18[7], which is unacceptable. Therefore, two implementations with improved linearity have been considered as well, refer to Figure 3.17.

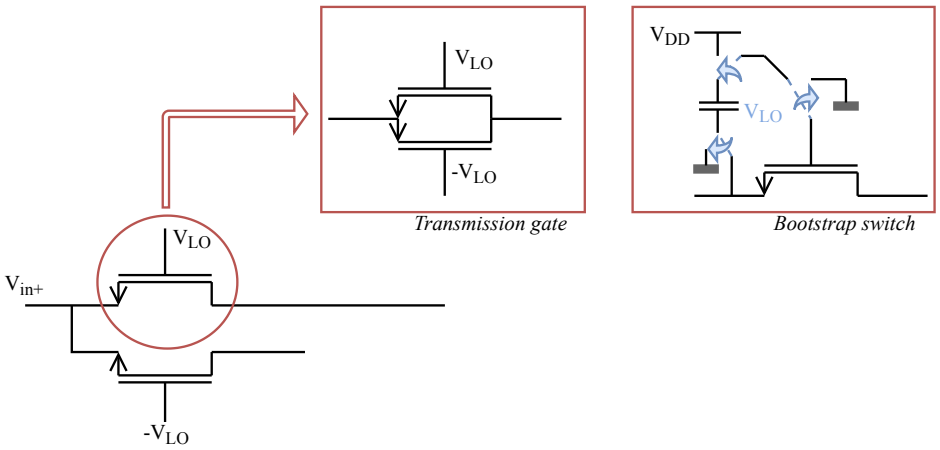


Figure 3.17: Schematic overview of a typical transmission gate and bootstrap implementation.

Simulations show that the transmission gate implementation achieves a 1 dB compression point of -4 dBm [29]. Again, without considering back-off from this compression point, the maximum peak current would be equal to 4.0 mA, requiring roughly 250 DACs, still surmounting to infeasible amounts of hardware.

At the time of publication of this thesis, no results on a bootstrap implementation were available for the 22FDX technology. However, the results on an NMOS or transmission gate implementation show that excessive amounts of hardware are required to upconvert the signal with sufficient linearity and a bootstrap implementation is not expected to boost the -4 dBm compression point of the transmission gate

[7] Solving Equation 3.18 for a maximum peak DAC current after splitting of 2.5 mA yields $n = 630$ with an average power of 10 dBm (is 10 mW).

implementation to acceptable values in the range of 5 - 10 dBm. Implementing a PA in each chain would allow for reducing the magnitude of the current the mixers have to handle while still meeting the requirement on transmitted signal power. Unfortunately, since the analysis in the beginning of this Chapter has shown that only when converting a single subcarrier per DAC the PAPR of each signal would be reduced by -4 dB, each of these amplifiers would be just as inefficient as those presented in Chapter 2 as the PAPR at the input of each PA would still be equal to roughly 10 dB, refer to Table 3.1.

Nonetheless, there might be applications where the linearity requirements are less strict or fewer subcarriers are required. To fully evaluate the effectiveness of splitting the signal for these applications, the design considerations related to signal combination in the analog-domain also have been explored. The next section will cover these considerations.

3.4 Signal combination in the analog domain

To complete the feasibility study on a transmitter design based on frequency-domain signal splitting, design considerations related to combination of the parallel DAC outputs into the antenna load will be treated in this Section. In line with the analysis in the previous section, a situation will be considered where the PA may be omitted. To avoid having to upconvert signals with a large magnitude, each DAC will be followed by its own mixer. Figure 3.18 provides the architecture of the signal combination in the analog domain. Together with the digital hardware as included in Figure 3.4, these provide a schematic overview of the complete transmitter.

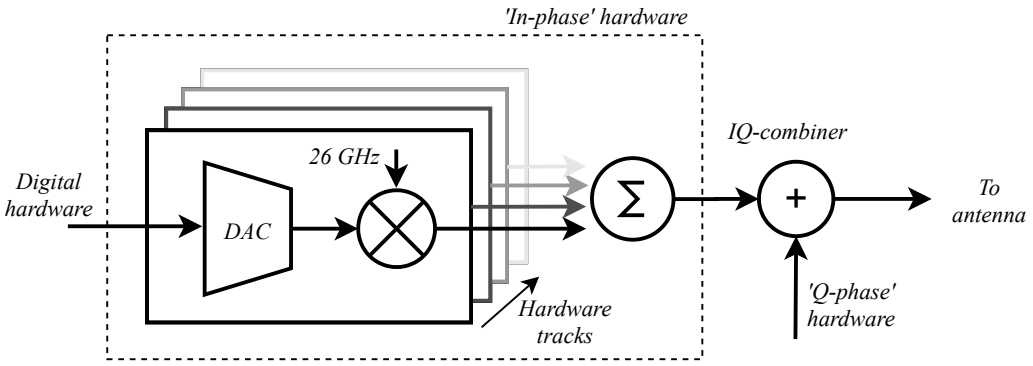


Figure 3.18: Situation sketch of signal combination in the analog hardware. Visualized are the DAC outputs of the in-phase subcarriers which are upconverted to the desired carrier frequency of 26 GHz in each hardware track and subsequently combined. Similar hardware is required for the quadrature-phase subcarriers. Both summed signals are combined in the I/Q-combiner before going to the antenna.

As discussed in Section 3.3.1, each DAC will be implemented using a current-steering architecture. To benefit linearity, differential signals will be used allowing cancellation of even-order harmonic distortion. It will be assumed that the mixers will be implemented as current-driven passive mixers, allowing connecting the 'positive' outputs to sum their current and the 'negative' outputs, similar to the situation as depicted for the differential load as in Figure 3.15.

The goal of omitting the PA was to improve upon its inefficient implementation. However, implementing the transmitter using the hardware as depicted in Figure 3.18 will introduce three new mechanisms that degrade efficiency. To be able to evaluate whether the implementation as in Figure 3.18 would be more efficient than the PA design examples provided in Chapter 2, these mechanisms will be discussed in this Section.

First, current-steering DACs have finite output impedance [25]. In a traditional transmitter, the output impedance of the current-steering DAC is typically much higher than that of its load, resulting in negligible power loss. However, after signal splitting, the number of DACs in parallel has increased significantly. Section 3.4.1 will evaluate the impact of the finite output impedance of the DACs when multiple are connected in parallel on the efficiency.

Second, since the full output power is now continuously generated by the DACs[8], the mixers have to process full signal power continuously. In addition to the dynamic switching dissipation that is normally present in traditional OFDM transmitters, this introduces additional dissipation in the on-resistance of the mixers. Section 3.4.2 will provide an estimate of the additional dissipation as a result of the larger signal.

Third, each of the DACs is fundamentally independent as it treats a subset of subcarriers which are per definition orthogonal to the subcarriers processed by other DACs. One DAC may try to pull the load, while the other may want to push it. As a result, part of the power generated by the DACs will cancel out in the load. This degrades efficiency, Section 3.4.3 will investigate this phenomenon.

3.4.1 Current leakage due to finite DAC output impedance

In this section, the impact of the finite DAC output impedance on the transmitter's efficiency will be analyzed. Consider the schematic as visualized in Figure 3.19. Each current-steering DAC has been modelled as a single current source. The mixer on-resistance was assumed negligible for this analysis.

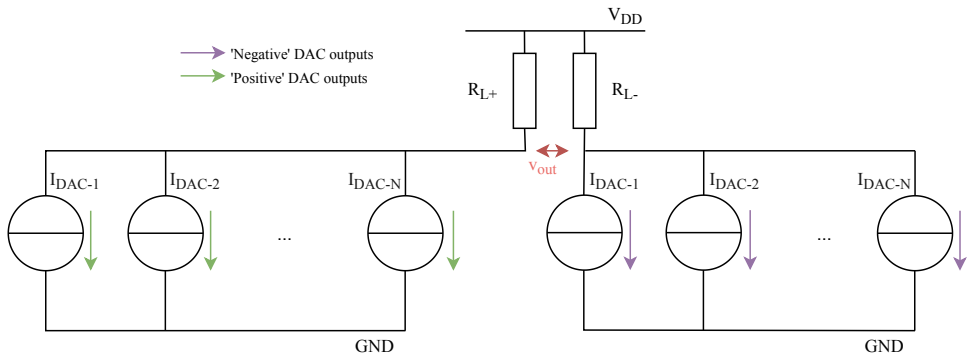


Figure 3.19: Schematic situation sketch for the driving conflicts. Mixer on-resistance was assumed negligible.

The ideal current sources as visualized in Figure 3.19 are typically implemented in CMOS using a cascode MOSFET stage [25], [30]. The limited output impedance of such a topology will lead to current leakage.

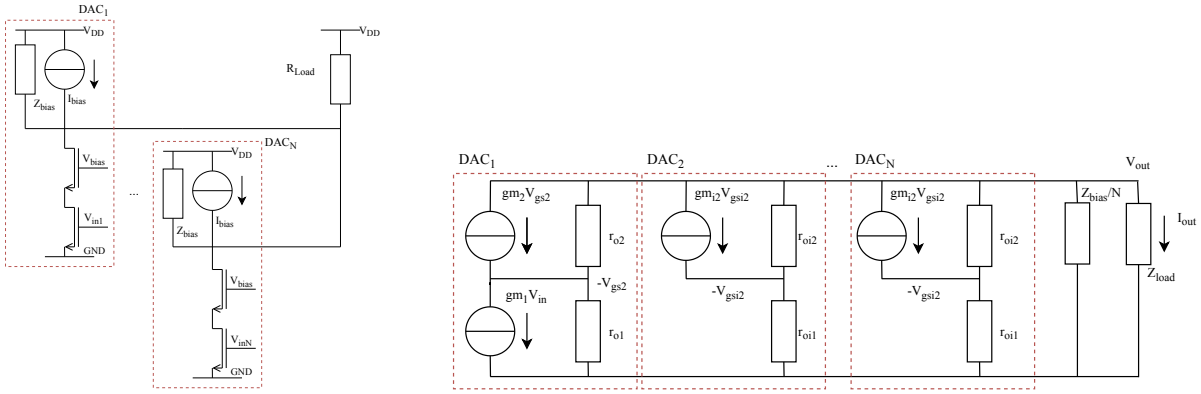
Several assumptions will be made to allow for an analysis that will increase insight into the magnitude of these leakage contributions:

- To benefit simplicity, no capacitive elements will be modelled.
- The mixer on-resistance was assumed to be negligible.
- The DACs will be assumed to be sufficiently accurately modelled using a single cascode stage.
- The DACs will be assumed sufficiently linear to allow for the use of linear signal analysis techniques.

Applying these assumptions leads to the situation sketch as included in Figure 3.20a. Note that the current source providing the bias current has been modelled as a perfect source and not as a transistor implementation, because it provides a DC current. The limited output impedance of this source will also impact the output impedance of the DACs and should be considered. The Small-Signal Equivalent Circuit (SSEC) of this circuit has been included in Figure 3.20b and allows determining the transfer function from a single DAC to the load. Through the superposition principle, the complete transfer function can subsequently be derived.

Solving the SSEC in Figure 3.20b through current balance in the v_{out}-node and substitution of the remaining expressions yields the result in Equation 3.19. The full derivation can be found in Appendix C.

[8]Note that this is a direct result of the use of current-steering DACs; full current flows continuously, even if the voltages to be generated are close to zero.



(a) Modelling of the individual DACs as cascode sources.

(b) Small-signal equivalent circuit for analyzing the transfer function of a single DAC to the output.

Figure 3.20: Circuit schematics for the analysis of current leakage when driving a single load with multiple DACs.

$$\frac{i_{out}}{v_{in}} = \frac{-gm_1}{1 + \sum_{i=2}^{N} \left(\frac{Z_{load}}{r_{oi2}} \left(\frac{1}{1 + r_{oi1}\left(\frac{1}{r_{oi2}} + gm_{i2}\right)} \right) \right) + \frac{Z_{load}}{r_{o2}} \left(\frac{1}{1 + gm_2 r_{o1} + \frac{r_{o1}}{r_{o2}}} \right) + N \cdot \frac{Z_{load}}{Z_{bias}}} \tag{3.19}$$

Through the superposition principle, the result in Equation 3.19 holds for each of the DACs as depicted in Figure 3.19. Several remarks regarding Equation 3.19 are in place:

- The impedance transformation as a result of the cascode transistor desensitises the impact of current leakage on the transconductance gain.
- If the transconductance of the transistors converting the input voltages is increased, the leakage decreases.
- A higher load impedance will increase the current leakage. Keeping the load impedance as low as possible (e.g. through impedance matching) will benefit the efficiency.
- The output impedance of the bias current sources should be as high as possible to limit leakage.

Verification of Equation 3.19 with a simple quadratic MOSFET model in LTSpice (see Appendix A) confirms the dependence on the number of DACs, see Figure 3.21 for an example with the transconductance set to $4\ mS$, the load impedance set to $50\ \Omega$ and Z_{bias} set to $100\ k\Omega$ (the output impedance of the transistors follows from the bias current required to set the right transconductance for each device). The magnitude is slightly off due to the omission of modelling of the body effect.

The result in Figure 3.21 can be replotted for more sources in terms of the relative degradation in efficiency with respect to a single source, see Figure 3.22. If more DACs are placed in parallel, the leakage into the finite output impedances increases.

Figure 3.22 indicates that the leakage becomes severe if the number of DACs is increased heavily, but is relatively small if only a few DACs are used.

3.4.2 Dissipation in the mixers

As a result of the 360 degrees conduction angle (class A) of current-steering DACs and the fact that full power is now being generated before the mixers, the dissipation in the mixer on-resistance increases. Similar considerations have been evaluated for the 22FDX technology in recent research on 'mixer last' transmitters [31].

Aside from linearity considerations, when it comes to dissipation in the mixer transistors, there is a trade-off in choosing a large transistor to limit the on-resistance (less dissipation in this resistance due to the signal current) and choosing a small transistor to limit the dynamic switching losses (less energy required

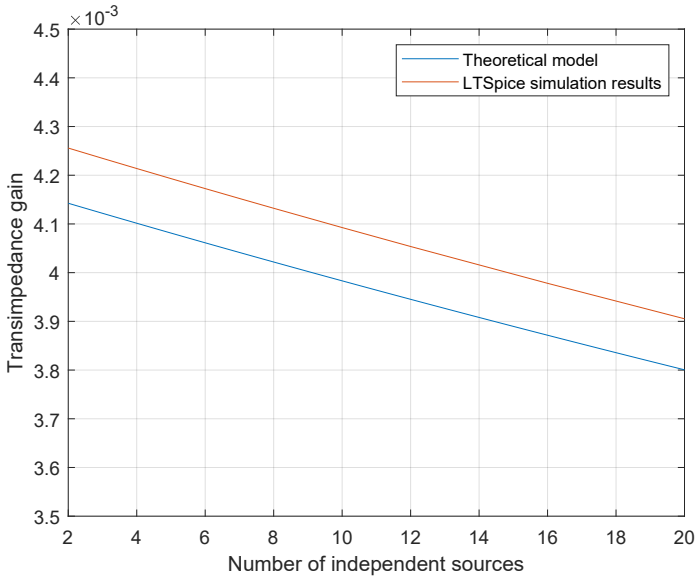


Figure 3.21: Evaluation of Equation 3.19 to visualize the dependence of the transconductance gain on the number of DACs.

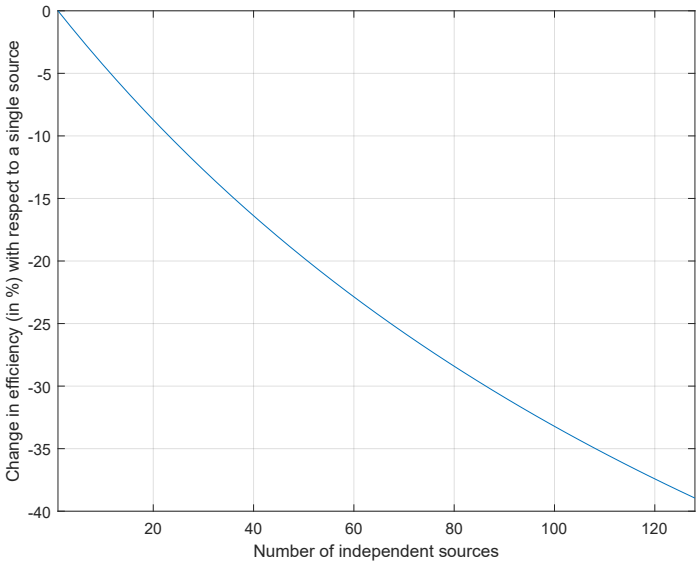


Figure 3.22: Change in efficiency with respect to driving the load with a single source as a result of current leakage into other sources.

to charge and discharge the parasitic capacitances).

The first contribution as a result of the mixer on-resistance can be determined by considering that the DACs collectively provide peak power all the time[9]. Based on the specification in Chapter 2, peak power is 20 dBm. Given the $50\ \Omega$ load, the peak current flowing through all mixers combined is equal to 63 mA. Now, modelling the on-resistance as a function of transistor size (through the number of transistor fingers, as in [31]), the total dissipation in the on-resistance of the mixer transistors can be defined as:

$$P_{on-resistance} = \left(\frac{63\ mA}{\sqrt{2}}\right)^2 \cdot \frac{R_{on-single-finger}}{nf} \tag{3.20}$$

in which $R_{on-single-finger}$ is the on-resistance of a mixer transistor with only 1 finger and nf is the number of transistor fingers[10].

Note that the result in Equation 3.20 is independent of the number of DACs used, as the output current of the DACs is scaled such that the output power requirement is met.

In case of the dynamic switching losses, the number of DACs does play a role, as it scales the number of mixer transistors of which the parasitic capacitances have to be charged/discharged. In [31] an approximation of the dynamic switching losses for a single mixer has been provided. Adapting this result to take into account the number of hardware tracks yields:

$$P_{switching} = f_{LO} \cdot nf \cdot C_{effective} \cdot V_{LO}^2 \cdot N \tag{3.21}$$

in which f_{LO} is the LO frequency (here 26 GHz), $C_{effective}$ is the effective parasitic capacitance of the mixer transistor, V_{LO} is the LO voltage swing and N is the number of mixers.

The power losses as in Equations 3.20 and 3.21 can be balanced to find the optimum number of fingers for the mixer transistors when the number of mixers is known. As an example, assume that 8 hardware tracks are used, thus $N = 8$. Figure 3.23 shows an example of the dissipation components using some 22FDX parameters available to the author.

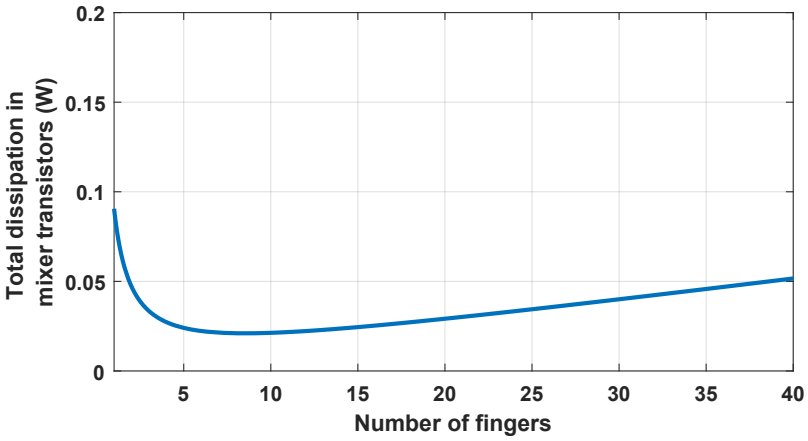


Figure 3.23: Total dissipation in the example of 8 mixer transistors under a sweep of transistor size (number of fingers). An optimum between low on-resistance and low switching energy can be distinguished at roughly 8 transistor fingers.

Figure 3.23 shows that the dynamic switching losses are relatively large as it increases with the number of mixer transistors used. The dissipation as plotted in the Figure is constant, irregardless of the useful power in the load due to the class A DAC architectures. Considering the average power of 10 mW,

[9]The largest swing in the load occurs when the differential current-steering DACs either direct all their current to the positive load resistor or fully to the negative load resistor, refer to Figure 3.19. For all output power levels that are lower than peak power, the DACs still output the same amount of current, but a part of it cancels in the load.

[10]Assuming an inversely proportional relation between the number of transistor fingers and the transistor's on-resistance as in [31].

the dissipation in the mixers would be larger than the useful power in the load, which is detrimental to efficiency. In addition to the dissipation in the mixer transistors, also cancellation of power in the load as a result of the class A DAC architecture will degrade efficiency. The next section addresses this issue.

3.4.3 Efficiency considerations regarding signal combination in the load

As a result of the use of current-steering DACs, the current required to generate the maximum output power flows to the load continuously, even if the power to be transmitted is lower, refer back to Figure 3.15.

The ratio of the useful output power in the load to the peak output power is therefore a figure for the efficiency in the load, as the rest of the power dissipates in the load resistors and cancels:

$$\eta = \frac{\frac{|v_{out}|^2}{R_L}}{\frac{v_{out-max}^2}{R_L}} \tag{3.22}$$

To find average efficiency over time, the distribution of v_{out} can be incorporated into this result, refer to Equation 3.23. Note that, since the DAC outputs have been combined in the load, a Gaussian distribution should again be used.

$$\eta_{load} = \int_{-\infty}^{\infty} \eta(v_{out}) \cdot p(v_{out}) \cdot dv_{out} \tag{3.23}$$

in which $p(v_{out})$ is the pdf of v_{out}.

The combination of Equations 3.22 and 3.23 shows a strong dependence of the average efficiency on the ratio of maximum output voltage to the average output value, e.g. its PAPR. A plot can be made evaluating the efficiency degradation in the load based on the PAPR, see Figure 3.24.

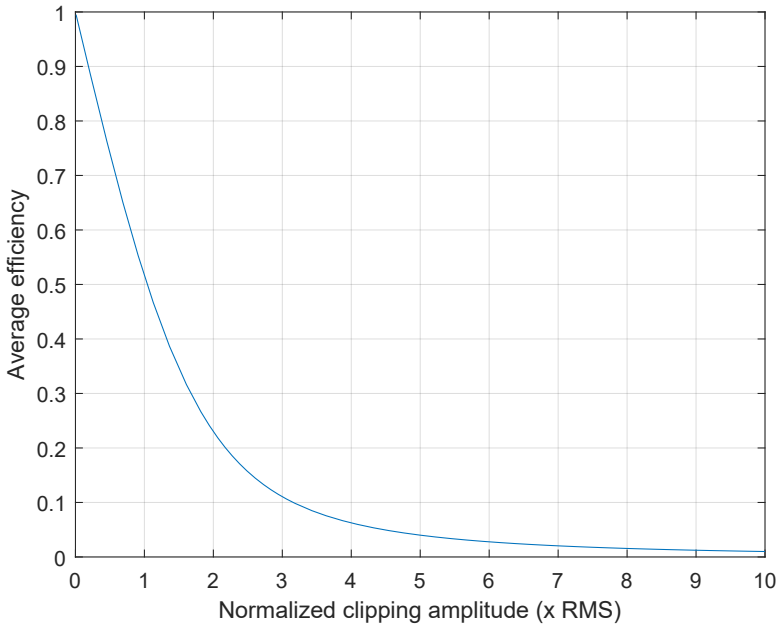


Figure 3.24: Average efficiency of signal combination in the load.

Figure 3.24 shows a strong degradation in efficiency as a result of the late signal combination. In case of the PAPR of 10 dB as specified in Section 2.5, it is only 9 % efficient. Early combining of the signal and subsequent amplification of the result would benefit efficiency when making use of current-steering DACs. In other words, converting as many subcarriers as possible in a single DAC will improve the efficiency, but this is limited by the linearity constraints imposed on the upconversion process.

3.4.4 Signal combination overview

The previous three sections have shown that predominantly the class A conduction angle of the current-steering DACs degrades efficiency. Also, especially unfortunate is the fact that the three efficiency degrading mechanisms stack.

Consider as an example 8 hardware tracks used to transmit a signal with the requirements as in Chapter 2. It will be assumed that the mixers can handle this signal with sufficient linearity. The dissipation in the mixer transistors would be equal to 20 mW (assuming the optimum as presented in Figure 3.23). Cancellation of power in the load as visualized in Figure 3.24 would come down to 90 mW for a signal with 10 dB PAPR. Combining these figures, one would generate an average of 10 mW with only 8 % efficiency, without considering any other losses in the transmitter.

It is clear that the combination of a class A conduction angle for the DACs and generation of the required power before performing upconversion is an inefficient combination. The Discussion chapter provides some suggestions on how to overcome some of these problems.

3.5 Comparison to existing techniques and chapter conclusion

The final Section in this Chapter will evaluate the effectiveness of the proposed transmitter concept for both PAPR reduction and efficiency improvement through the omission of the PA.

3.5.1 Using frequency-domain splitting for PAPR reduction

The effect of splitting up the OFDM signal along the frequency domain has been investigated. By reducing the number of subcarriers converted per DAC, the PAPR of the resulting signals could be improved with respect to the OFDM case. Converting a single subcarrier per DAC yields the highest PAPR reduction, amounting to a factor of 2.5 times lower PAPR, see Table 3.1. As with any PAPR reduction technique, this improvement comes at a cost.

Along the lines of the performance parameters as defined by Tao and Jiang in [3] and already introduced in Chapter 2, the impact of signal splitting on other system parameters has been determined. The effects of reducing the number of subcarriers converted per DAC to 1 on these parameters have been evaluated.

When it comes to bandwidth expansion and spectral pollution, signal splitting transmitters perform well. As the ideal iFFT signal is fully reconstructed, albeit now using a combination of expanded digital and analog hardware, there is no increase in used bandwidth. Also, as no changes are made to the output signal with respect to OFDM waveforms, there is also no increase in spectral pollution.

The other parameters, however, do show degradation. Implementation complexity, for instance, is high. The additional hardware that is required to replace the functionality of the iFFT block is significant. This includes both digital hardware as well as analog hardware, see Sections 3.3.1 and 3.4. Note that extensive analysis into the required digital hardware has not been conducted in this Thesis, some remarks on this decision have been included in the Discussion chapter, refer to Chapter 5. When it comes to analog hardware, integration of signal combination was found to be especially challenging considering the stringent linearity requirements and the relatively high magnitude of the analog signals.

The increase in power consumption required for frequency-domain signal splitting is tremendous, predominantly because the amount of DACs required is equal to the number of subcarriers, refer to Section 3.3.1. Within the 5G NR design boundaries, this comes down to hundreds of DACs in addition to the analog hardware to combine their outputs. In addition, conflicting sources in driving the load and the class A conduction angle of the required current-steering DACs result in additional efficiency degradation, refer to Section 3.4.4.

BER degradation is difficult to characterize. In the DACs, as a result of signal splitting, the distortion due to clipping has reduced significantly, leading to improved linearity performance. However, the effect of frequency-domain splitting on the linearity in the analog hardware following the DAC is difficult to characterize as it depends on the made design trade-off between linearity and other factors such as area and

power consumption. More research is required to determine if (bootstrapped) mixers can be designed that can handle relatively large signals with sufficient linearity.

In the work of [3], a generalized overview has been included of the magnitude of the PAPR reduction that can be achieved using the various techniques as discussed in Chapter 2. Carefully recognizing all implementation uncertainties, a PAPR reduction of a factor 2.5 as achieved in this chapter (amounting to -4 dB) is roughly in line with the performance obtained with clipping, PTS and TR techniques.

Table 2.2 from Chapter 2 can be expanded with an additional row, see Table 3.2.

Table 3.2: Comparison of PAPR reduction techniques, including frequency-domain signal splitting.

PAPR reduction technique	**PAPR improvement**	**Implement. complexity**	**Bandwidth expansion**	**BER degradation**	**Power consumption**	**Spectral spillage**
Clipping and filtering	$+/-$	$+$	$++$	$-$	$++$	$--$
Coding schemes	Varies	$+/-$	$-$	$+/-$	$+$	$++$
PTS and SLM	$+/-$	$--$	$++$	$+$	$-$	$++$
Nonlinear companding schemes	$++$	$+/-$	$++$	$+/-$	$+$	$+$
TR and TI	$+/-$	$-$	$--$	$+$	$+/-$	$+$
Signal splitting (f-domain)	$+/-$	$--$	$++$	$+/-$	$--$	$++$

The two major drawbacks for the frequency-splitting technique are the increased implementation complexity and power consumption. Changes required to the analog hardware are significant and the increase in hardware needed to process all signals in parallel is excessive. As a result of independent hardware acting concurrently, the efficiency further degrades due to driving conflicts. These downsides make the use of frequency-splitting to reduce the PAPR less attractive than using any of the other techniques as identified in Table 3.2.

3.5.2 Using frequency-domain splitting for the omission of the PA

There is another area of application for the frequency-splitting concept. Following from the analysis in Section 3.3.2, using multiple hardware tracks to generate the required output power could render an inefficient PA and its associated problems unnecessary. Also, this design strategy does allow for the combination of multiple subcarriers before conversion with a DAC.

However, challenges related to driving conflicts were identified and the possible gain in efficiency has to be balanced to the losses as a result of multiple independent sources driving the load and signal dissipation in the mixers. An example calculation for the efficiency in Section 3.4.4 based on the signal parameters as provided in Chapter 2 showed that when considering only the losses related to the class A conduction angle of the DACs and the dissipation in the mixers, the efficiency already dropped to only 8 %.

Considering that the PA designs as presented in Chapter 2 achieve considerably higher efficiency figures, the use of frequency-domain signal splitting to improve the efficiency of a transmitter for 5G NR applications is not recommended.

Chapter 4 will address another signal splitting concept aimed at solving the problems related to concurrently active hardware.

4 PAPR REDUCTION BASED ON TIME-DOMAIN SPLITTING

Chapter 3 has treated transmitter concepts based on frequency-domain signal splitting. This chapter marks the second half of this thesis and covers signal splitting in the time-domain to alleviate the effects of a high PAPR in OFDM modulated transmitters.

First, the system concept for time-domain signal splitting will be introduced. Then, based on an analysis of the signal statistics of OFDM signals, the achievable PAPR reduction will be determined. Challenges related to hardware implementation will be addressed and suitable parameters to evaluate performance will be determined. This chapter will be concluded with an overview of the obtained results.

4.1 System overview of time-domain signal splitting in transmitters

The research in Chapter 3 has surfaced several efficiency related problems caused by concurrently active hardware. Because of splitting the OFDM signal into groups of multiple subcarriers, multiple DACs were generating orthogonal signals that had to be combined into a single load.

As an alternative to frequency-domain splitting, a transmitter has been designed in this chapter which is based on processing of the OFDM signal depending on its value at a certain time-instant, avoiding conflicting sources. To understand how this could benefit PAPR reduction, refer to Figure 4.1.

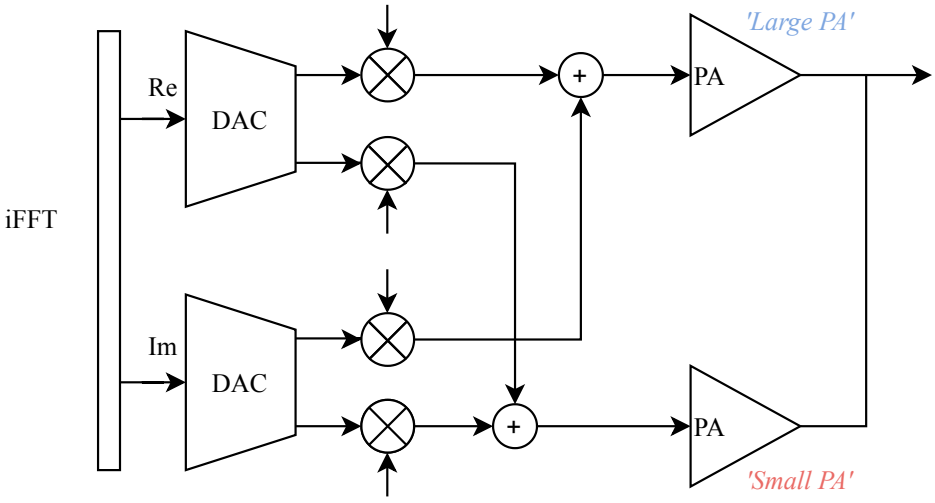


Figure 4.1: Schematic representation of time-domain splitting with two hardware tracks applied in a transmitter. The large magnitude samples are dedicated to the upper hardware track, small magnitude samples to the lower hardware track. Only one PA is active at a certain time instant.

Figure 4.1 shows an example of a transmitter with two 'hardware tracks', e.g. two chains of a typical DAC - Mixer - PA combination. Incoming samples from the iFFT are dedicated by the DACs to either the upper or lower hardware track depending on their magnitude; samples larger than a predefined threshold are dedicated to the hardware track with the 'large' PA and samples smaller than the threshold are allocated

to the 'small' PA. As a result, the maximum amplitude for the smaller PA drops significantly, while the average signal value of the larger PA increases. This leads to a reduction in PAPR for both hardware tracks. Note that this concept can be scaled as well; increasing the number of hardware tracks will further benefit PAPR reduction for each of them. For simplification and illustration purposes, only two hardware tracks will be considered in the rest of this section.

While either of the PAs is active, the idea is that the other can be switched off as it is not processing any samples at that time instance. This avoids the problems with concurrently active hardware as identified in the PA topologies that were discussed in Chapter 2; there is no need for a power combining network at the output as in an outphasing amplifier or for quarter-wave transmission line sections as in a Doherty structure. It will also limit the increase in power consumption as a result of the introduction of additional hardware.

Figure 4.2 includes a Simulink model that visualizes the most important baseband signal flows. Model details can be found in Appendix A.

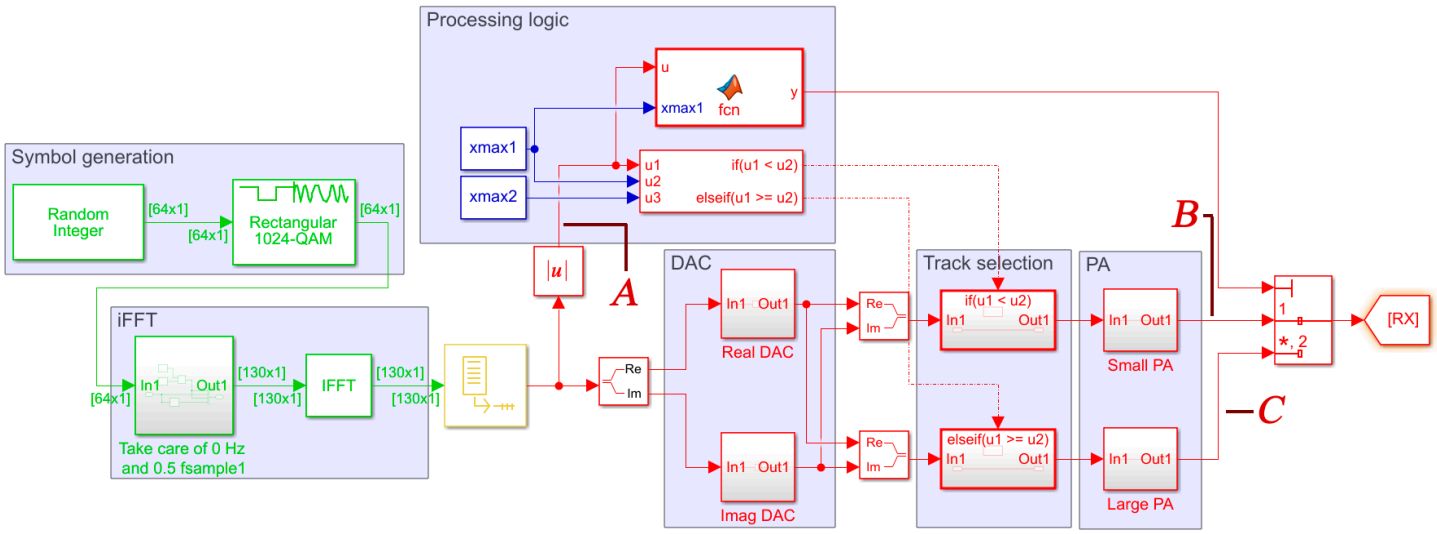
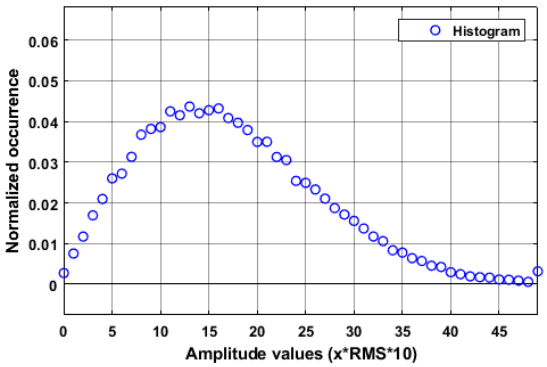


Figure 4.2: Baseband Simulink model of the time-domain splitting transmitter topology, based on the idea as presented in Figure 4.1. The block 'Baseband processing' can be replaced with an upconversion block to implement the transfer to the Radio Frequency (RF). The characters A, B and C refer to the signals as plotted in Figure 4.3.

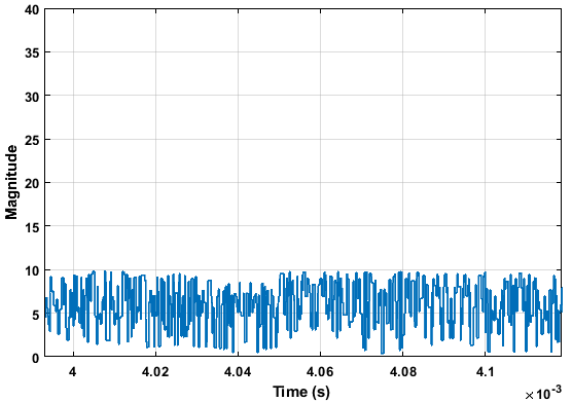
The implementation of symbol generation and the iFFT are similar to those used in a typical model based on OFDM modulation. The model in Figure 4.2 deviates from a 'traditional' OFDM modulated transmitter in the two outputs for both the 'real' and 'imaginary' DACs and the additional power amplifier.

Several signals have been visualized in Figure 4.3 to increase insight into the model of Figure 4.2. The magnitude of the incoming samples has been plotted in a histogram in Figure 4.3a. The histogram shows that the magnitude of the baseband samples adheres to a Rayleigh distribution as was discussed in Chapter 2. Based on a pre-defined threshold, subsequently, one of the two PAs is selected. This results in two waveforms with reduced PAPR, refer to Figure 4.3b and Figure 4.3c.

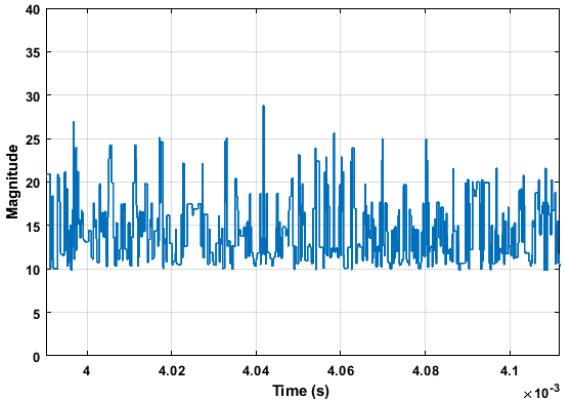
The rest of this chapter will analyze the achievable reduction in PAPR and characterize design challenges and opportunities related to the topology as included in Figure 4.1. First, the optimum strategy for signal splitting will be covered and an analysis into the obtainable PAPR reduction will be provided.



(a) A histogram plot of the magnitude of the baseband samples as generated by the iFFT block. As discussed in Chapter 2, the magnitude of these samples adheres to a Rayleigh distribution.



(b) Time-domain plot of the waveform at the output of the 'small' PA.



(c) Time-domain plot of the waveform at the output of the 'large' PA.

Figure 4.3: Visualization of the three signals as designated with the characters A, B and C in the Figure of the Simulink model used to model baseband signal flows in the time-domain signal splitting transmitter, refer to Figure 4.2.

4.2 Analysis of PAPR reduction through time-domain splitting

The value at which to switch to the other PA directly impacts the achievable reduction in PAPR. The goal of this section is to find an optimum for these transition moments in terms of lowest overall PAPR and therefore highest overall transmitter efficiency.

In considering an optimum division of the original dynamic range into two (or more) subranges, also PA usage over time has to be factored in. For instance, heavily optimizing the PAPR for one PA when it is only rarely used will yield a suboptimal implementation in terms of overall PAPR reduction. To be able to provide some generalized analysis into the PAPR reduction that can be achieved, it will be assumed that each of the PAs would profit equally in terms of efficiency from an equal reduction in PAPR of their input signal. Figure 4.4 illustrates the consequences of this assumption to the efficiency of both PAs and hence to overall transmitter efficiency. Chapter 2 already showed that PA efficiency for stand-alone CMOS PAs is typically around 20 % PAE at average output power. As Figure 4.4 shows, a reduction in PAPR will boost this value.

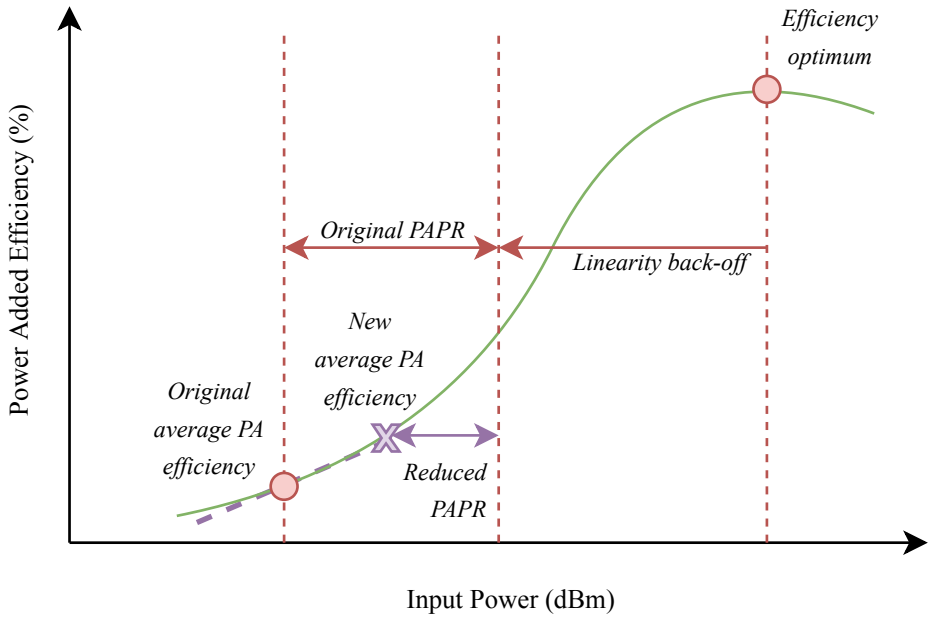


Figure 4.4: A linear relation between PAPR reduction and PAE will be assumed. The reduction in PAPR for both PAs will result in an increase of overall PA efficiency.

In order to calculate the largest possible PAPR reduction to profit as much from the efficiency improvement as possible, OFDM signal statistics will be analyzed. The goal is to calculate the lowest weighted average of the PAPR possible, to take into account the aforementioned usage of both PAs over time as well as their respective PAPR reduction.

Chapter 2 already discussed many of the signal properties of transmitters based on OFDM modulation; the real- and imaginary-components of an OFDM waveform were shown to adhere to a normal distribution with zero mean. As discussed, orthogonally combining both results yields a Rayleigh distributed magnitude. This Rayleigh-distributed complex signal is at the input of each PA and will therefore be the basis of the subsequent PAPR reduction analysis. Note that the pdf of a Rayleigh distribution has been introduced in 2.4 and visualized in Figure 2.3b.

Before proceeding, it should be noted that the RMS value of the normalized Rayleigh distribution as visualized in Figure 2.3b is equal to $\sqrt{2}$, which follows directly from the combination of two orthogonally spaced normalized Gaussian distributions with an RMS value of 1, yielding a magnitude of $\sqrt{1^2+1^2}$. Therefore, for normalized Rayleigh plots used in this chapter, the x-axis will be plotted in terms of $x \cdot \frac{RMS}{\sqrt{2}}$, where RMS indicates the RMS value of the *complex* time-domain waveform (e.g. *not* the RMS value of

the separate in-phase or quadrature-phase components). Note that it would also be possible to use the RMS value of the I- or Q-component of the OFDM signal, in which case the x-axis could be normalized in terms of $xRMS$. However, $x\frac{RMS}{\sqrt{2}}$ is used as this makes it easier to compare the theoretical results to the Simulink simulation results in which the RMS value of the complex waveform is typically readily available.

4.2.1 Defining PAPR for the subranges

As a first step in determining the optimum subranges, the impact of splitting up the original Rayleigh-distributed signal into smaller subranges on the PAPR of each subrange will be determined. A definition for the PAPR was provided in Chapter 2 and remains valid here, refer to Equation 2.5. After splitting, the maximum amplitude and the RMS value of each subrange may change, resulting in a change in PAPR. As can be seen in Equation 2.5, both need to be known to determine the subrange PAPR.

The new maximum value is equal to the upper limit of the new amplitude subrange:

$$x_{max-i} = max\left(x_i[x_{max-(i-1)}, x_{max-i}]\right) \tag{4.1}$$

in which i designates the i-th hardware track and x_i indicates the possible amplitude values of the i-th track.

The squared RMS value of the subrange can be found by determining the expected value of the range of squared amplitude values for the smaller amplitude range after renormalization of the probability density function to the subrange boundaries:

$$x^2_{RMS-i} = \frac{\int_{x_{max-(i-1)}}^{x_{max-i}} \left(x^2 \cdot p(x) \cdot dx\right)}{\int_{x_{max-(i-1)}}^{x_{max-i}} \left(p(x) \cdot dx\right)} \tag{4.2}$$

in which $p(x)$ is the pdf of the amplitudes of the baseband OFDM signal.

Taking the ratio of the square of Equation 4.1 and Equation 4.2 yields the PAPR for each of the subranges.

Note that the obtained PAPR *reduction* will strongly depend on the PAPR of the original signal; if its high, the hardware will benefit more from signal splitting than when its low. Following from the requirements in Section 2.5, the PAPR of the original signal is 10 dB, which would come down to $x_{max} = 4.4\ \frac{RMS}{\sqrt{2}}$. To aid calculation simplicity, this figure will be rounded in this chapter to the nearest integer ($4\ \frac{RMS}{\sqrt{2}}$), yielding a PAPR of 9.0 dB.

4.2.2 Analysis of optimum subrange boundaries

The second step in the analysis is to determine the optimum location of the subrange boundaries to yield the lowest PAPR over time. To that end, the weighted average of the combined PAPR for the subranges should be minimum[1]. The pdf of the amplitudes as used in Equation 4.2 can be used to calculate how often a PA would be used with a particular set of subrange boundaries as integrating over the pdf along the subrange boundaries provides relative use of that PA over time[2].

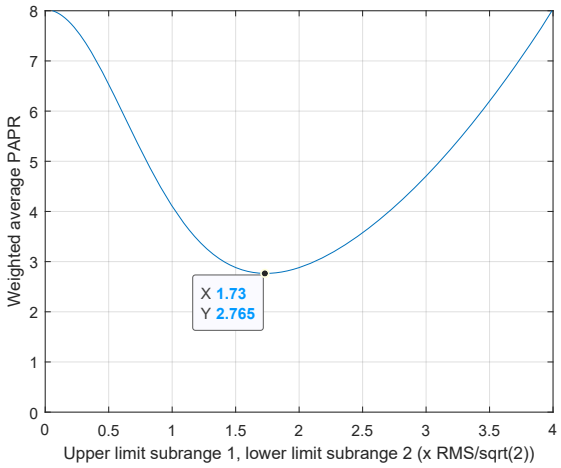
Mathematically, the weighted average has been defined as:

$$PAPR_{combined} = \sum_{i=1}^{N} \left(\int_{x_{min-i}}^{x_{max-i}} p(x) \cdot dx \cdot PAPR_i \right) \tag{4.3}$$

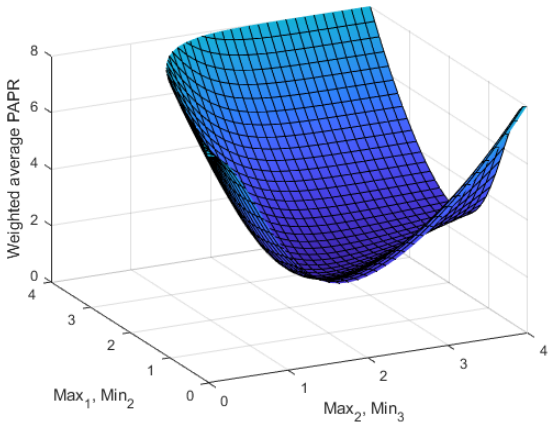
in which $PAPR_i$ is the PAPR of the i-th hardware track according to the ratio of Equation 4.1 squared and Equation 4.2.

[1] It is important to realize that the weighted average is important, because of the assumption introduced while explaining Figure 4.4 that each hardware track will profit equally from PAPR reduction in terms of efficiency.

[2] The total area under the pdf is equal to 1 as it is normalized. Integration of a section of it will therefore yield relative use over time.



(a) Weighted average of the combined PAPR under a sweep of the boundary between two subranges in the case with 2 hardware tracks.



(b) Weighted average of the combined PAPR under a sweep of the boundaries between three subranges in the case with 3 hardware tracks.

Figure 4.5: Visualization of the impact of the boundary borders on the weighted average of the PAPR.

Equation 4.3 has been generalized to an arbitrary number of hardware tracks (N). Also note that the subranges border each other, hence $x_{max-i} = x_{min-(i+1)}$, except for the outer two where $x_{min-1} = 0$ and $x_{max-N} = x_{max}$. The remaining subrange borders can be solved for lowest overall PAPR. To that end, the minimum of Equation 4.3 will be found for 2, 3 and 4 hardware tracks. Figure 4.5 has plotted the weighted average of the PAPR for the relevant subrange boundaries in case of 2 or 3 hardware tracks.

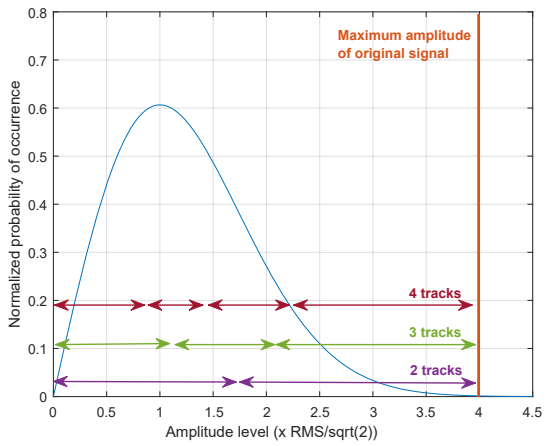
Figure 4.5 visualizes that there is an optimum in the location of the subrange boundaries to yield lowest overall PAPR. These optima as visualized in Figure 4.5 (the minima) can also be analytically determined by solving the (partial) derivative(s) to be equal to zero. This yields the optimum subranges as included in Table 4.1.

Table 4.1: Overview of amplitude subranges after signal splitting to yield overall lowest PAPR.

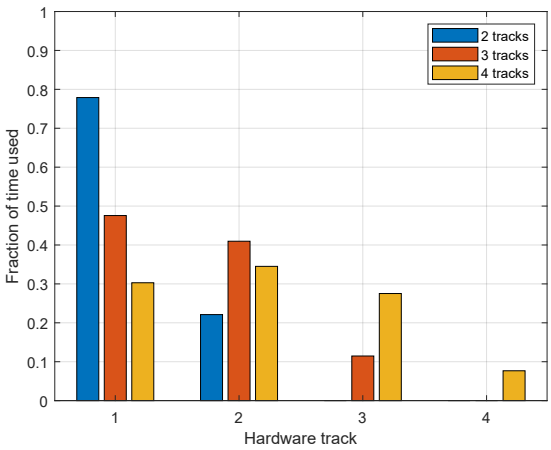
Number of tracks	**Hardware track**	**Amplitude subrange ($\frac{RMS}{\sqrt{2}}$)**
2	1	[0.00, 1.74)
	2	[1.74, 4.00]
3	1	[0.00, 1.13)
	2	[1.13, 2.08)
	3	[2.08, 4.00]
4	1	[0.00, 0.85)
	2	[0.85, 1.45)
	3	[1.45, 2.27)
	4	[2.27, 4.00]

By providing an overlay of the subrange boundaries as included in Table 4.1 on the Rayleigh-distribution, visual inspection of the usage of each hardware track becomes possible, refer to Figure 4.6a. The area of the Rayleigh distribution bounded by a particular subrange relative to the distribution's total area provides a figure for the usage of that subrange over time similar to how it was used to determine the weighted average, see Figure 4.6b.

The hardware usage as plotted in Figure 4.6b allows for making an estimate of the magnitude of the transient effects as a result of switching between hardware tracks. For example, when considering the use of 2 hardware tracks, 78 % of the time it is 22 % likely that the next sample will be processed by the other track while it is 22 % of the time 78 % likely that there will be a switch. This combines to a 34 % switching rate. The consequences of this result will be covered later on in this Chapter, when transient effects are analyzed.



(a) Distribution of the original signal into subranges that yield lowest overall PAPR for 2, 3 and 4 hardware tracks.



(b) Fraction of time that each hardware track is used when the pdf is divided into subranges that yield lowest overall PAPR.

Figure 4.6: Distribution of the original signal into subranges that yield lowest overall PAPR and hardware usage over time.

4.2.3 PAPR reduction after signal splitting in the time-domain

The acquired subrange boundaries can be substituted into the PAPR definition for each subrange as introduced in Equations 4.1 and 4.2. Table 4.2 provides an overview of the obtained results. Note that the weighted average has been calculated using Equation 4.3 and therefore also takes into account the usage of each hardware track over time (refer to plot 4.6b).

Table 4.2: Overview of PAPR reduction through signal splitting into subranges that yield lowest overall PAPR. The original signal has a PAPR of 9.0 dB.

Number of tracks	Hardware track	PAPR	Reduction wrt to original
1 (original)	*NA*	8.00	*NA*
2	1	2.64	3.03x
	2	3.21	2.49x
	Weighted aver.	2.77	2.89x
3	1	2.24	3.57x
	2	1.77	4.52x
	3	2.53	3.16x
	Weighted aver.	2.08	3.85x
4	1	2.13	3.76x
	2	1.57	5.10x
	3	1.59	5.03x
	4	2.25	3.56x
	Weighted aver.	1.80	4.44x

Figure 4.7 provides an overview of the PAPR reduction in decibels that can be achieved with 2, 3 and 4 hardware tracks based on the results in Table 4.2. Three things should be observed regarding Figure 4.7:

- Because of the assumption that each hardware track would profit equally from PAPR reduction in terms of efficiency, overall PAPR reduction through the weighted average provides an estimate of total efficiency improvement.
- Already with 2 hardware tracks, the PAPR of each track is reduced by over a factor 2 with respect to the original signal.
- Further reducing the PAPR of each track by increasing the number of hardware tracks has limited impact; the PAPR reduction starts to saturate.

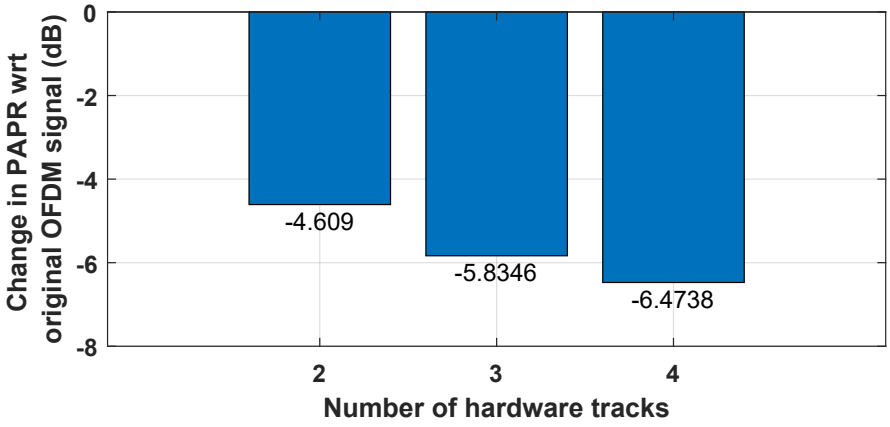


Figure 4.7: PAPR reduction as a result of signal splitting in decibels with respect to an OFDM signal with 9 dB PAPR.

The biggest step in PAPR reduction for adding an additional hardware track occurs when moving from 1 to 2 hardware tracks. Hence, to benefit the PAPR reduction versus simplicity trade-off as much as possible, the rest of this chapter will be dedicated to the implementation of a transmitter based on 2 hardware tracks.

4.2.4 Verification in Simulink

The PAPR reduction results have been verified in Simulink, refer to Appendix A for model details. Table 4.3 provides an overview of the results. The results from Table 4.2 match within 3 % accuracy to the Simulink simulation results.

Table 4.3: Simulink verification on the impact of time-domain signal splitting on PAPR.

Number of tracks	Hardware track	Theor. PAPR	Simulated PAPR	Percentual deviation
2	1	2.64	2.64	$+0.0$ %
	2	3.21	3.23	$+0.6$ %
3	1	2.24	2.23	-0.4 %
	2	1.77	1.79	$+1.1$ %
	3	2.53	2.58	$+2.0$ %
4	1	2.13	2.12	-0.5 %
	2	1.57	1.58	$+0.6$ %
	3	1.59	1.58	-0.6 %
	4	2.25	2.31	$+2.7$ %

Now that the analysis in this Chapter has been verified, the subsequent sections will explore the required hardware to fully benefit from the 4.6 dB PAPR reduction through the use of 2 hardware tracks. Challenges related to the implementation will also be covered to be able to draw conclusions on the effectiveness of this approach.

4.3 Exploiting alleviated PAPR characteristics in transmitters

This section covers the hardware required to implement a transmitter based on time-domain signal splitting. Dedicated subsections will be devoted to each of the three analog transmitter components as included in Figure 4.1: the DACs, the mixers and the PAs. After that, an additional section will be dedicated to the hardware overhead required to select the appropriate hardware track and disable the unused PA.

4.3.1 Extending the DACs with additional outputs

Similar to the DACs treated in Chapter 3, the DACs discussed in this Chapter have to adhere to the requirements on 5G NR transmitters. Mainly as a result of the strict bandwidth requirements, this dictates the use of current-steering digital-to-analog conversion [25].

Here, in addition to the design considerations as treated in Chapter 3, the DAC has to be extended with one (or more) additional outputs to be able to feed the multiple PAs. Considering that only one output is used at a particular time-instance, there are in general two options to extend the DAC with additional outputs:

1. Extend the DAC with additional current sources to be able to sustain multiple output currents.
2. Extend the DAC with multiplexing hardware to be able to route its current to one of multiple outputs; the unused output(s) will Return-to-Zero (RZ).

A short discussion on the use of each option will now be provided.

Sustaining multiple output currents

The first option extends the DAC with additional current sources such that multiple currents can be sustained. The advantage of sustaining multiple output currents is that there are no sudden changes in the signals appearing at the inputs of the PAs, which would degrade in-band signal quality[3]. This comes at the cost of the additional current sources that are required to generate the additional current and dissipation in the mixers of the hardware tracks that are unused. Alternatively, an RZ approach would limit dissipation in the mixers of the hardware tracks that are unused to dynamic switching losses, as there is no current flowing through the unused mixer transistors.

To illustrate how many additional sources would be required, consider the case with 2 hardware tracks. Assume that there is a typical current-steering DAC with a single output (as depicted in Figure 3.15) which handles the larger output current. If the additional assumption is made that the DACs processing the real and imaginary signals both have to output maximum current to reach the maximum complex magnitude of $4 \cdot RMS/\sqrt{2}$ (for 9 dB PAPR, refer to Section 4.2.1), the maximum output current for each of these DACs would be $2 \cdot RMS$.

In case of the 7 bit DAC as depicted in Figure 3.15 and used as example here, this results in an LSB size of $2/128 \cdot RMS$. Now, to be able to sustain the smaller current as well, additional sources are required. Following from the analysis at the start of this chapter, the complex magnitude of the smaller signal is at most $1.74 \cdot RMS/\sqrt{2}$. This is the complex magnitude given by the combination of the Re- and Im-DAC output currents through $\sqrt{I_{Re}^2 + I_{Im}^2}$. This complex magnitude could also be generated fully by the real or imaginary DAC if the other outputs a very small current (I_{Re} or I_{Im} is near zero).

Therefore, when considering this worst-case scenario, another $\frac{1.74 \cdot RMS/\sqrt{2}}{2/128 \cdot RMS} \approx 79\ I_{LSB}$ are required for both the Re- and the Im-DAC. Along the segmentation requirements as discussed in Chapter 3, this would require another binary section (to make the smaller values between the larger steps in unary sources) and several sources of $16 \cdot I_{LSB}$. In case of the additional $79\ I_{LSB}$, the additional binary section provides a total of $15\ I_{LSB}$, leaving $64\ I_{LSB}$ which can be generated with 4 unary sources of $16 \cdot I_{LSB}$, refer to Figure 4.8.

This example shows that a substantial increase in the number of current sources (1.5x) is required to sustain one additional output. In combination with the class A conduction angle of 360 degrees, this introduces additional dissipation, also in the mixer transistors. To avoid these problems, an RZ-approach would be preferred, the next section addresses this option.

[3] The use of oversampling and interpolation to reduce the strength of aliases after conversion makes sure that the change from one PA to the other is smooth and occurs at the subrange boundaries.

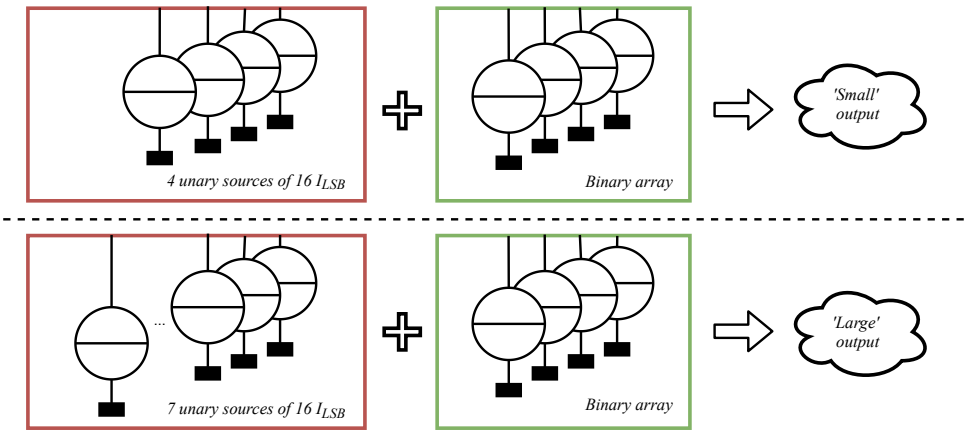


Figure 4.8: Extending the DAC with current sources allows sustaining multiple output currents, in this case two. Since the extra sources have to sustain a current that does not reach full scale, less sources are required than when two full-scale DACs would be used.

RZ for unused outputs

Instead of sustaining two or more output currents, integration of multiplexing hardware will allow for routing the DAC's output current to the desired hardware track. An example of the use of a multiplexing structure in a DAC can be found in [32]. The multiplexer enables one DAC to serve multiple hardware tracks without requiring additional current sources. Figure 4.9 provides a schematic of an RZ approach with 2 hardware tracks.

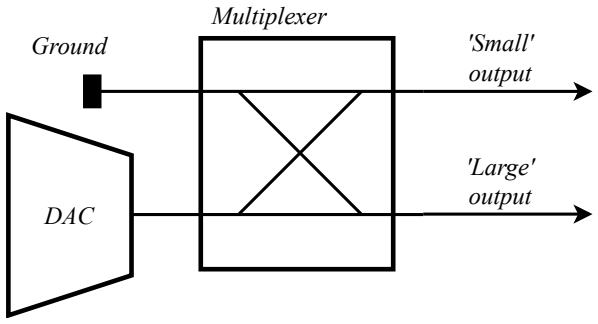


Figure 4.9: A multiplexing structure can be used to serve multiple hardware tracks with a single DAC.

An added benefit of using the structure as depicted in Figure 4.9 is that there is no current flowing through the mixer transistors of unused hardware tracks, limiting dissipation to dynamic switching losses only.

The downside of an RZ approach is the negative impact of the signal changes on the quality of the transmitted waveform. As a result of bandwidth limitations, it will take time to settle to either the desired output value or to zero. An extensive analysis into the impact of such transient effects will be provided in Section 4.4.1.

Considering the significant increase in dissipation as a result of using additional current sources, the RZ-approach as presented is preferred provided that the transient effects are sufficiently limited to allow for signal transmission with acceptable quality.

4.3.2 Integration of upconversion in the transmitter hardware

To reach the targeted carrier frequency of 26 GHz as specified in the requirements in Section 2.5, mixers will have to be implemented. Implementation is relatively straightforward and deviates from implementing upconversion in a typical transmitter only in the number of hardware tracks required. One could use current-driven passive mixers for example, to make use of the current-mode outputs of the DACs.

It should be noted however, that, depending on the upconversion implementation, the PAPR of the signals at the inputs of the PAs may change. To see how, first consider upconversion using a current-driven passive mixer. In a passive mixer, the incoming baseband signal is upconverted to the RF carrier frequency using multiplication with +1 and -1, effectively mixing with a square wave running at the carrier frequency. Assuming infinite bandwidth, no changes are made to either the peak value after upconversion or to the RMS value.

Now consider upconversion using a sine-wave running at the carrier frequency. The resulting RF signal will have an RMS value that is a factor $\sqrt{2}$ lower than that of the original baseband signal while the peak value remains unchanged, effectively increasing the PAPR. The impact on the PAPR *reduction* is a matter of perspective; comparing the PAPR after time-domain splitting to the PAPR of a transmitter also based on sine-wave upconversion, but without time-domain splitting, results in equal PAPR reduction. Would the PAPR after time-domain splitting be compared to the baseband PAPR or to a transmitter based on square-wave upconversion, the reduction would be lower.

Considering the higher PAPR *reduction*, mixing with a square-wave is the preferred solution, given that the spurious tones are kept in check. The reader should, however, keep in mind that, in any real system, the inherently limited bandwidth of the electronics will round of the square-wave corners and approximate a sinusoidal signal, limiting the obtained PAPR reduction.

4.3.3 Design considerations related to power amplification

As the final component considered here in the signal chain, the PAs have to be implemented. Their design will be more complicated than that of the DAC and mixers as they have to be designed such that they profit from the reduced PAPR at the input of both PAs. Figure 4.10 provides a situation sketch based on the results in the beginning of this Chapter. Note that predrivers may be required in front of the PAs to achieve sufficient amplification, refer to for instance Figure 4.17.

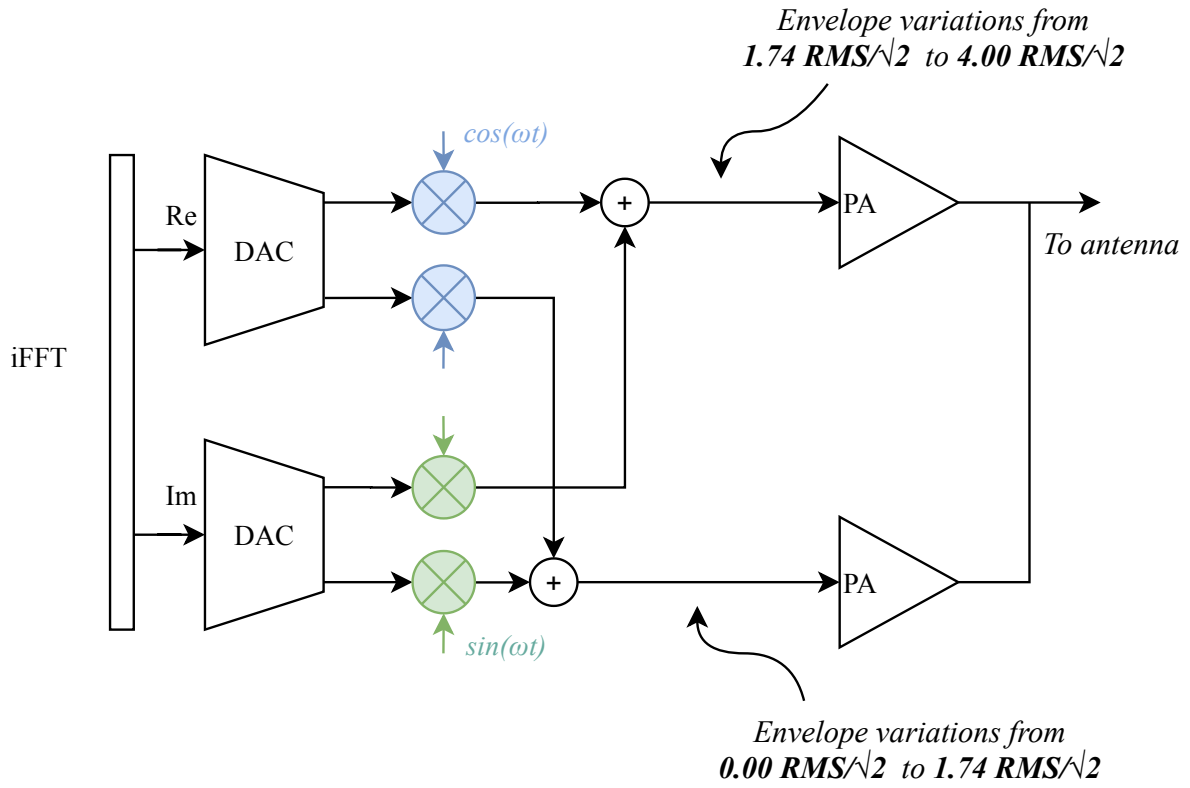


Figure 4.10: Situation sketch with PA locations and some signal properties highlighted. Both PAs have to be designed such that they fully profit from the PAPR reduction while adhering to the transmitter requirements.

The main question that will be answered in this section is:

> How can both PAs profit in terms of efficiency from the PAPR reduction at their input whilst satisfying 5G NR transmitter requirements and allowing for the disabling of the unused PA?

The fact that both PAs have to be designed to be efficient at different output (and input) powers introduces design challenges. Furthermore, the use of amplitude modulation, the strict linearity requirements for 5G NR and the use of the PA at mmWave frequencies dictate the use of so-called linear amplifiers, restricting the design options to amplifier classes A, AB, B and C (and possible improvements to these designs, such as a class G PA) [18].

Since the use of class C amplifiers introduces considerable distortion [18] and class AB amplifiers are a combination of class A and class B amplifiers, only the latter two will be treated in this section. A class A amplifier will be treated first.

Class A amplifiers

A class A amplifier is a linear PA that has a 360 degrees conduction angle, meaning that it is always on [18]. To explain how it works, a typical schematic for a class A amplifier has been included in Figure 4.11.

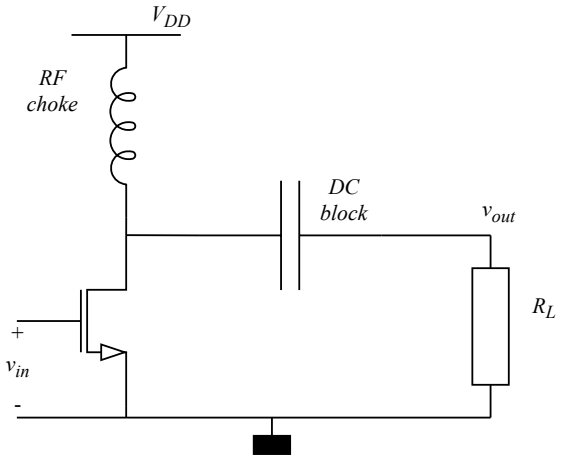


Figure 4.11: Schematic for a typical class A amplifier.

There are two key mechanisms that are important to understand this PA's efficiency performance.

First, the RF choke presents a short at DC between the supply rail and the NMOS drain terminal. Any AC voltage will therefore swing around V_{DD}, allowing an amplitude of almost the supply voltage before the NMOS device is pushed out of the desired operating region.

Second, the DC supply current flowing through the RF choke is fixed and therefore relatively constant (because the inductor is relatively large). For low input voltages, the NMOS conducts little current and the DC current flows into the load, increasing the output voltage. To have the output voltage decrease, the NMOS has to sink both the DC current as well as current from the load resistor [5].

As a result of the combination of these two mechanisms, the efficiency of a class A PA is maximum when both the voltage and current swings are maximized, where it hits an efficiency figure of 50 %. When moving away from this efficiency optimum as the input power decreases (referred to as power back-off) efficiency falls rapidly as useful output power is decreased while DC voltage and current remain unchanged [18]. Matching networks can be deployed to adapt the load resistance to the available supply voltage such that the efficiency optimum can be achieved at the desired peak output power.

To understand how this works, the larger PA from Figure 4.10 will be designed as a class A PA. Considering the maximum output power of 20 dBm in the 50 Ω-load, the peak-to-peak sinusoidal voltage swing over the load would be $6.3\ V_{pp}$. Considering that the swing available at the drain terminal of the NMOS

device is equal to twice the supply voltage of 0.8 V in 22FDX technology, a matching network is required that increases the drain terminal voltage swing by a factor of 4.0. Following from power conservation, the current swing at the output will be reduced by the same factor with respect to the drain terminal. Figure 4.12 illustrates this situation for the voltage swings.

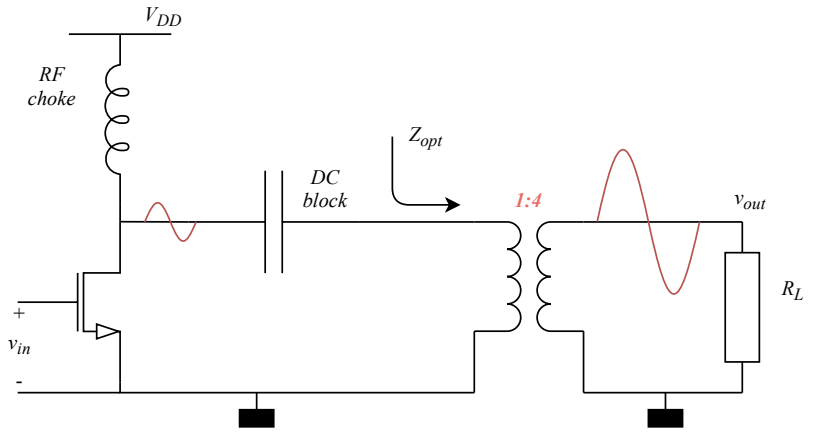


Figure 4.12: Matching of the large PA from Figure 4.10 for optimum efficiency at the peak output power of 20 dBm in 22FDX technology. Voltage swings have been annotated.

Now, in an attempt to also design the smaller PA, fundamental limitations to the class A topology start surfacing. The first important observation in this regard is that it is impossible to cut-off the bias current. To design two PAs and turn off the one that is unused, one would have to pull the supply voltage to zero Volt, pulling along the NMOS drain voltage. This would introduce significant transient effects at the output that would not be present if only the bias current would be cut-off. As an alternative, one could attempt to reuse at least the RF choke, but this has limitations as well. Consider the case in Figure 4.13.

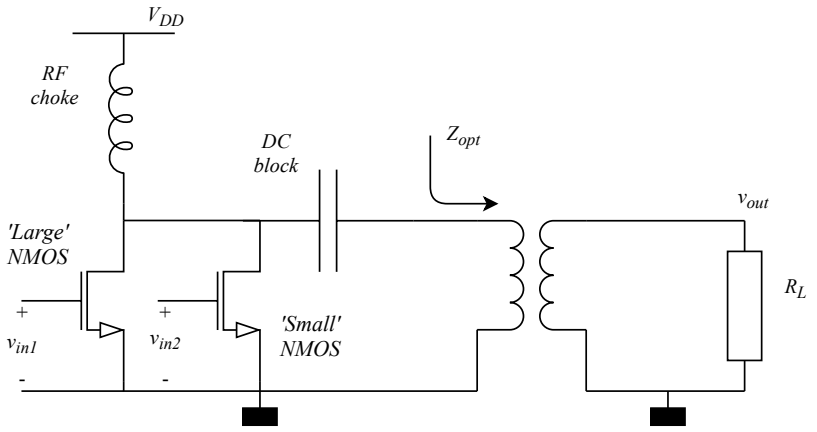


Figure 4.13: A class A PA extended with an additional NMOS device.

To be able to fully profit from the reduction in PAPR for both the large and small NMOS device, each should have full swing at its maximum output power. The only way to realize this with a fixed DC current (the RF choke is shared) and fixed supply voltage is to change the optimum load and have it optimized for both NMOS devices separately. However, changing the optimum load will lead to a discontinuity in the power transmitted through the load resistor when switching from the large to the small device and vice-versa. In other words, changing the optimum load to restore full voltage swing when switching to the smaller device would require the current swing to decrease in order to maintain continuous output power in the load, leading to low efficiency.

Considering the fact that designing two separate PAs cannot be made efficient as the unused PA cannot be disabled efficiently and the fact that attempting to re-use parts of the large PA in the smaller one to circumvent these issues results in fundamental efficiency limitations, a class A implementation for the

time-domain splitting transmitter is unattractive.

The reader should note that letting go of some of the properties of a typical class A PA may provide workable solutions, such as allowing variations in its supply voltage, e.g. a class G approach. Solutions based on variations in supply voltage exist and are typically referred to as Envelope Tracking (ET) amplifiers [18], they are closely related to the polar amplifiers as discussed in Chapter 2.

Since the main aim of this thesis is to explore the benefits of using parallel hardware, no explicit attention will be paid to such an implementation in this chapter. A small discussion of this option will be provided in the Discussion chapter, refer to chapter 5.

Another alternative that does work with parallel hardware is a class B implementation, which will now be considered.

Class B amplifiers

A class B PA differs from a class A PA in its conduction angle which is 180 degrees. Although several implementation exist, in this section the focus will be on an implementation based on a set of differential NMOS devices that together provide the desired voltage swing, refer to Figure 4.14.

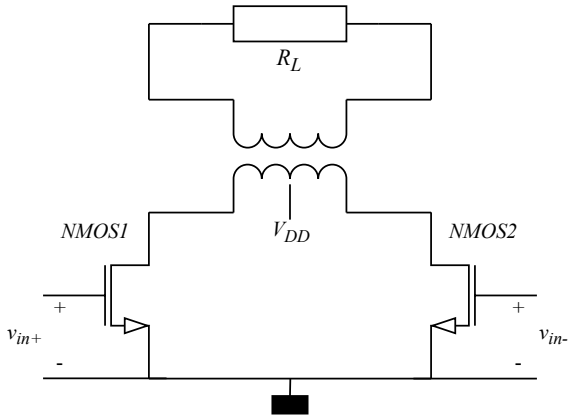


Figure 4.14: A schematic of a differential class B amplifier based on NMOS devices.

The advantage of this differential implementation is that the bias current of the complete PA can be cut-off by disabling both NMOS devices, while maintaining the DC-voltage at the drain terminals. Class B PAs have several other properties that serve the current application [18]:

- There is no quiescent[4] current flowing, as both NMOS devices are biased directly below cut-off for zero AC input voltage.
- Similar to class A, matching is based on resistive load matching and transistor parasitics absorption.
- The DC power consumption scales with input power.

Class B amplifiers also reach peak efficiency when AC swings are maximized [18]. Therefore, to design both class B PAs such that they can optimally profit from the reduced PAPR at their input, similar design considerations for the matching network are required as in the case of a class A device. This means that for the large PA, again one finds the output swing in a 50 Ω-load to be equal to 6.3 V_{pp}, which has to be transformed to the maximum voltage swing of twice the supply voltage. This yields a transformation ratio of 1 in 4 (primary to secondary). As a result, the peak current that has to be delivered to the transformer by the NMOS devices is equal to 0.26 A_p.

Also the required transconductance for the NMOS devices can be derived if an additional assumption is made that the voltage swing at the NMOS gates is at most the supply voltage of 0.8 Volt. Considering that the input amplitude is then 0.4 V and the required peak drain current is 0.26 A_p, a transconductance

[4] Current flowing with no input voltage applied.

of 0.65 Siemens is required.

Now, as calculated in the beginning of this chapter, ideal PA transition occurs when the magnitude of the input voltage is equal to $1.74 \cdot RMS/\sqrt{2}$, which is $4.00/1.74$ times smaller than the assumed maximum input amplitude of 0.4 V, hence 0.17 V. The power in the load of the large PA is then equal to 19 mW following from the transconductance and matching ratio. As output power has to be continuous when switching to the smaller PA, peak power in the load for the small PA is also 19 mW. Restoring full supply voltage swing in the small PA therefore requires a transformation ratio of 1:1.7.

Following similar reasoning for the transconductance as in case of the large PA, the required transconductance for the small PA is equal to 0.27 Siemens. Note that the ratio of transconductance to matching ratio has to be equal for the large and small PA to guarantee continuous output power, which is the case.

Figure 4.15 summarizes the required matching and transconductance values.

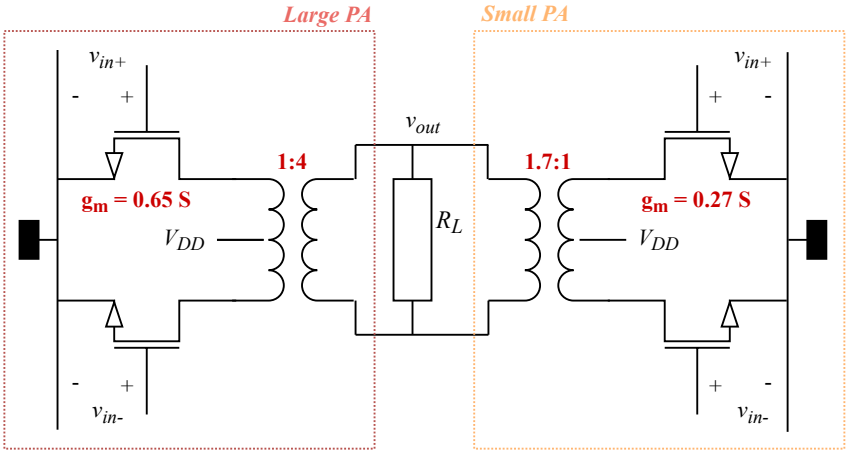


Figure 4.15: Dual differential class B implementation for time-domain signal splitting transmitters with matching ratios and transconductance values annotated.

Figure 4.15 also reveals two other challenges related to this topology.

First, transformers work bidirectional. This means that if one of the PAs has been disabled through pulling the NMOS gates to ground, there will still be a large voltage swing at their drains. Especially the smaller PA suffers from this problem. Consider peak output power for the large PA, yielding a peak-to-peak voltage swing of 6.3 V_{pp} across the load. Through the transformer for the small PA, this yields a voltage swing of 3.7 V_{pp} at the drain terminals of the NMOS devices (also at a very high frequency). Such a large voltage may cause breakdown and shortens the lifetime of the transistors. Depending on the exact implementation and breakdown voltage specs it may therefore be required to adjust the transformation ratio for the smaller PA. Note that because current is fully cut-off at the NMOS devices, no current leakage occurs in the 'unused' transformer.

Second, the relatively large transformer inductance limits a sudden change in current. It will take time for the bias current to die out or to reach its desired bias value whenever the gates of the NMOS devices are adjusted. This will introduce transient effects in the output voltage whenever a PA handover takes place. Because of its importance to system performance, this phenomenon will be analyzed in Section 4.4.1.

Efficiency improvement with reduced PAPR in a class B PA topology

The previous section has described how two class B PAs could be used to realize a transmitter based on time-domain signal splitting. Impact on PA efficiency based on this dual class B PA implementation will now be determined[5].

[5]To present a general result, typical linearity back-off will be ignored. If desired, analysis can be redone using the equations in this Section to determine efficiency improvement when required linearity back-off is known.

First, consider the typical efficiency curve of a single class B PA. Assuming that the efficiency of the amplifier achieves its theoretical maximum of $\frac{\pi}{4} \rightarrow 78$ % at full voltage swing, then its efficiency will decline proportionally with input voltage [18], refer to Equation 4.4.

$$\eta_{classB} = \frac{v_{in}}{v_{in-max}} \left(\frac{\pi}{4}\right), 0 < v_{in} < v_{in-max} \tag{4.4}$$

Now, referring the result in Equation 4.4 to the dual class B implementation, both PAs were designed to profit from maximum efficiency at their own respective maximum input voltage. This effectively means that two distinct typical class B PA efficiency curves can be identified when sweeping the input (or output) power over its range that both adhere to Equation 4.4 albeit with different maximum voltages, refer to Figure 4.16.

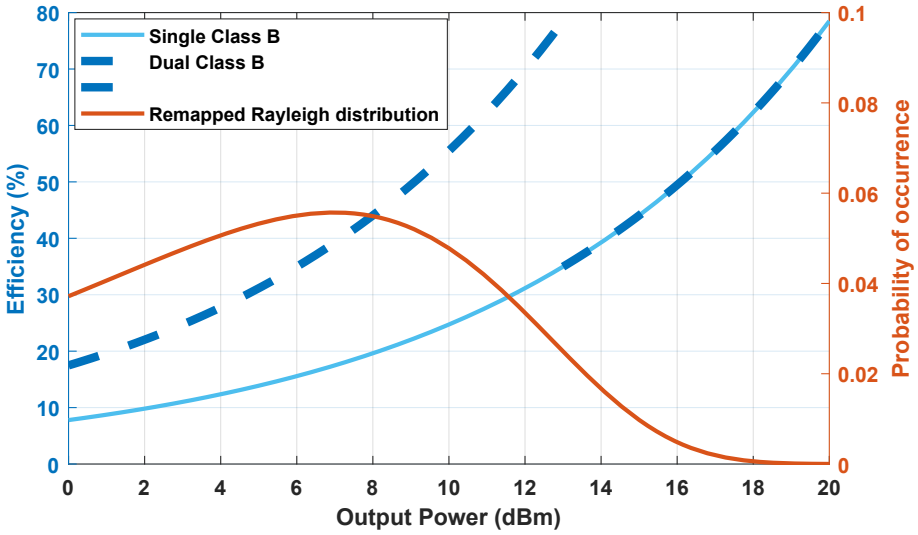


Figure 4.16: A plot of typical efficiency curves for class B PA implementations. In case a single PA would be used, the light blue curve provides the efficiency. The dual class B implementation based on time-domain signal splitting boosts efficiency and has been visualized using the dashed dark blue curve. The distribution of the output power levels in a transmitter based on OFDM modulation has been included.

Figure 4.16 visualizes that efficiency has improved at power BO as a result of the dual class B implementation. Because the PAPR of the signals at the inputs of both PAs has been reduced, overall efficiency improves:

- In case of the larger PA, samples that are much smaller than the input maximum (and hence also efficiency maximum) no longer have to be processed as the RMS value of the input value increased.
- In case of the smaller PA, the maximum input voltage is much smaller, allowing re-optimizing this smaller range in terms of efficiency.

A quantitative number for the efficiency improvement depends strongly on hardware implementation details such as the required back-off to satisfy linearity demands (to avoid gain compression). Nonetheless, an estimate of the efficiency gain would be instructive. To that end, the average efficiency for both the single and dual case will be determined by taking into account how the output power levels are distributed. Therefore, the Rayleigh distribution as presented in Equation 2.4 has to be remapped to the output power levels in dBm. This can be done by replacing x in Equation 2.4 by the inverse relation between dBm and Volts, yielding the curve as plotted in Figure 4.16 after renormalization.

For both cases, the average efficiency can then be determined by calculating the product of the efficiency curve and the pdf, yielding 19 % and 39 % average efficiency for the single and dual case respectively. Hence, without linearity back-off taken into account, an efficiency improvement of a rough factor 2.0 has been achieved. However, this improvement comes at a cost. Section 4.4 will evaluate the impact of signal splitting on other transmitter parameters.

4.3.4 Hardware overhead

To make time-domain splitting work, also some additional hardware overhead is required. Figure 4.17 provides a schematic overview of a full transmitter based on time-domain signal splitting.

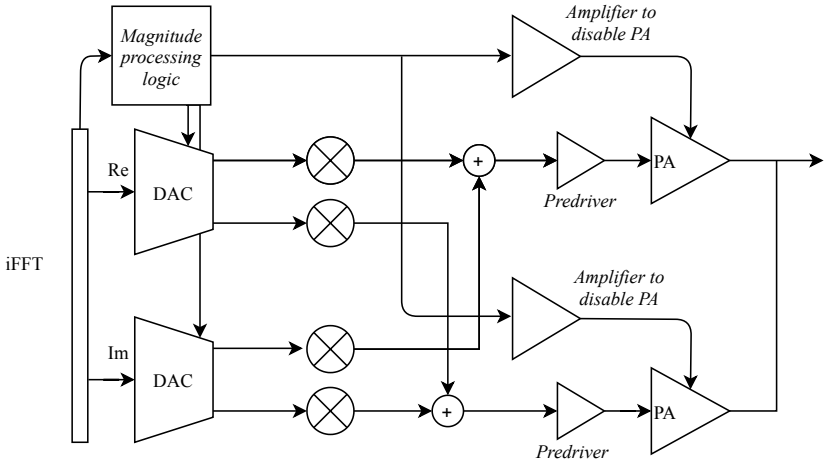


Figure 4.17: A complete overview of the required functional blocks of a transmitter based on time-domain signal splitting using two hardware blocks.

Two additional blocks have not yet been treated in this Section and require further attention: the magnitude processing logic and the amplifier required to disable the unused PA.

The implementation of the digital logic is relatively straightforward. It could consist of a digital comparator and memory storing the pre-defined switching thresholds. Information on the signal's magnitude can be provided by the iFFT block. Some kind of amplifier would furthermore be required to transform the digital output after comparison to a signal that can control the amplifiers required to disable the unused PA.

Implementing these amplifiers is much more difficult. To be able to disable the unused PA, the additional amplifier should pull the NMOS gates of the unused devices to ground. As Section 4.4.1 will show, this additional amplifier/driver needs to be fast to avoid excessive transient effects. Unfortunately, splitting up the OFDM signal into a magnitude and phase part to determine if the PA has to be disabled will yield increase the bandwidth for each of these parts that is in the order of 3 to 5 times the original modulated signal bandwidth [33]. Considering that the signal bandwidth in 5G NR transmitters can go up to 1 GHz and assuming that the 3rd and 5th harmonic are required to disable the main PA fast enough, the additional amplifier would have to support up to 25 GHz bandwidth worst-case, which is large.

In combination with the large input capacitance of the main PA that has to be disabled, the required signal bandwidth poses major problems to this amplifier's implementation. For instance, interconnect inductance is considerable at mmWave frequencies, which, combined with the large input capacitance, imposes bandwidth limitations that would also negatively impact the amplifier's efficiency.

Similar implementation challenges can be recognized in ET PAs [18], which are closely related to the polar amplifiers as discussed in Chapter 2. In ET, the supply voltage of the RF PA is varied according to the baseband signal's magnitude instead of the load, leading to a similar efficiency gain as discussed in Section 4.3.3. An ET PA also experiences similar bandwidth limitations as the time-domain signal splitting transmitter discussed in this Chapter [18].

A brief review of recent ET PAs for telecommunication applications confirms the difficulty to implement PAs that support the required 5G NR bandwidths. An overview of ET PAs has for example been presented in [34] which support bandwidths of at most 20 MHz [35], [36]. ET PAs at mmWave frequencies do exist [37], but also here, the bandwidth is limited to tens of MHz. Also, no polar or ET PAs have been included in the PA overview on 5G NR mmWave applications as presented by Camarchia et al. in [9], which is a strong indication that no designs have been presented yet that can cope with the strict bandwidth requirements.

In light of the limited time frame available for this thesis, no extensive analysis has been conducted to determine alternative solutions to this problem[6]. For now, analysis into transient effects will be performed as if an amplifier/driver is available that can disable the main PAs fast enough.

4.4 Challenges related to hardware switching

The previous Sections have shown that splitting up the signal in the time-domain benefits PA efficiency. However, because of the use of different hardware tracks and the switching between them, transient effects are introduced that will degrade the signal's linearity. Section 4.4.1 will address this issue. Furthermore, to allow for switching between the hardware tracks, additional hardware is required that degrades part of the gained efficiency. Considerations related to this phenomenon will be treated in Section 4.4.2.

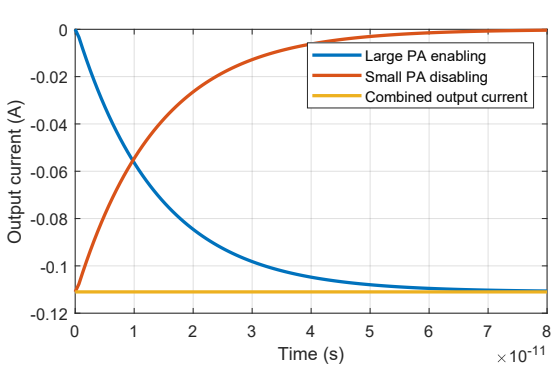
4.4.1 Transient effects and their impact on performance parameters

Previous sections in this Chapter have shown that making use of time-domain splitting in transmitters can enhance the transmitter's efficiency at the cost of, amongst other things, the introduction of additional transient effects. The goal of this section is to provide an estimate of the impact of these transient effects on the in-band signal quality and out-of-band pollution. Before proceeding, the reader should note that in the rest of this section, transient effects introduced because of hardware switching will be referred to as PA handover effects to avoid confusion with typical crossover distortion that also originates in class B PAs.

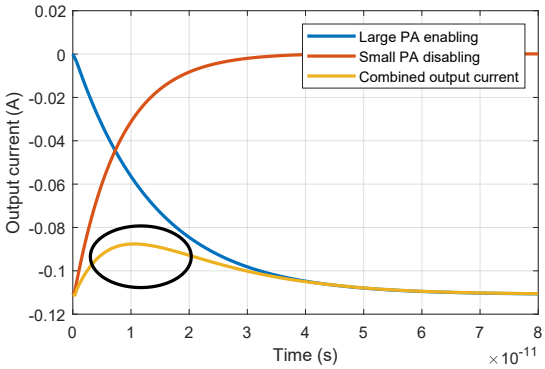
This section will contain three subsections. First, some words on modelling details will be provided, such as assumptions made and analysis strategies. Then, a theoretical analysis will be performed and the results will be discussed. Finally, a Simulink model will be used to evaluate the impact on the EVM and ACPR.

Modelling transient effects due to PA handover

The main aim of this analysis on transient effects is to determine the impact of PA handover with the two different PAs in the configuration of Figure 4.15 on the signal's quality in the load. Ideally, when both PAs adhere to the same Transfer Function (TF), the signal in the load remains stable during handover, even if both PAs are band-limited. However, differences between both TFs will show up as distortion in the load, refer to Figure 4.18. The Figure shows the output current under a step in the gate voltage of the PA (e.g. pulling it to ground or re-enabling the PA).



(a) With an equal TF for both PAs, ideally no distortion shows up in the output current. As an example two bandlimited, but equal, TFs in response to an equal step in their input and the sum have been plotted.



(b) With a different TF for each PA, distortion shows up in the summed output current.

Figure 4.18: The impact of different transfer functions on PA handover illustrated.

[6]There might be a way for instance to make the RZ of the DAC output fast and strong enough to disable the class B PA fast enough to avoid excessive distortion, refer to the Discussion chapter.

To model this phenomenon and evaluate its impact on the EVM and ACPR, the TF from input voltage to output current will be derived for both PAs in Figure 4.15 using a SSEC. Note that the use of linear circuit analysis techniques has major impact on modelling accuracy as SSEC analysis requires linearization around the transistor's operating point. The benefit is that linear analysis provides significant insight without requiring complicated analysis techniques, but one has to be careful not to violate the operating region boundaries.

In the push-pull class B PA as depicted in Figure 4.15 there is always 1 NMOS device operating in the linear region, except at the crossover point, where the Common-Mode (CM) level is exactly at the threshold for both devices. The impact of this brief moment in time where the operating region does not hold (because both devices are 'off'), will be assumed to be small enough.

Several additional assumptions will be introduced now to benefit simplicity and increase insight into the handover behaviour.

- Along the lines of linear analysis, any used RF chokes providing the bias current to the NMOS devices will be AC-modelled as a short in subsequent analysis[7]. A discussion on the impact of this modelling decision will be provided separately, after evaluation of the simulation results.
- Considering that the matching networks can be implemented in various different ways that impact the SSEC analysis, it will be implemented using ideal transformers to generalize the results. The generality with which the SSEC analysis will be conducted will allow readers to substitute a matching network later, to their own liking.
- Since the focus in this thesis is on 22FDX technology, parasitic transistor modelling will be limited to the gate-drain and gate-source capacitances. The other parasitic capacitances will be assumed to be negligible in Fully Depleted Silicon-on-Insulator (FDSOI) technology.
- Using input voltage and transconductance ratios as depicted in Figure 4.15, the smaller PA has to provide a maximum output current that is $4.00/1.74 \cdot 0.65/0.27 = 5.5$x smaller than what the large PA has to provide. The smaller PA will therefore be dimensioned to have an area of a factor 5.5x less than the larger PA, easing the requirements on the predriver (which has to drive less input capacitance). It will be assumed that this will result in similarly scaled parasitic capacitances.
- It will be assumed that the amplifiers that have to disable/enable the PAs are infinitely fast, e.g. the input voltage can be changed instantly. This will allow for evaluating the transfer of the amplifier itself separate from any artefacts as a result of the limited bandwidth of the driving network.

Figure 4.19 combines the above assumptions into an SSEC for either of the PAs after making use of the symmetry plane in the class B amplifier.

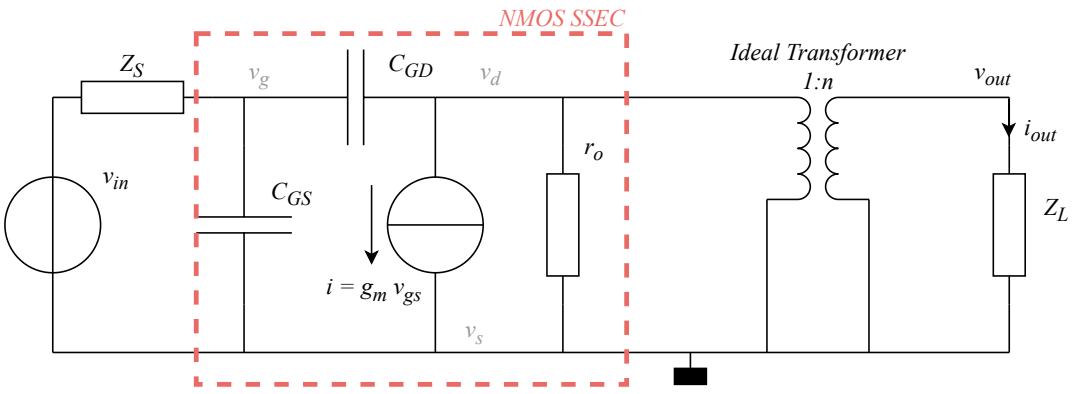


Figure 4.19: Small signal equivalent circuit of one symmetric half of the class B amplifier in the configuration as included in Figure 4.15.

A transfer function will be derived for the SSEC as depicted in Figure 4.19 in the next section.

[7] In Figure 4.15 no RF chokes have been visualized. Alternative implementations exist where an RF choke is used for both NMOS devices and where these have been separated from the typical fundamental matching network [18].

Transfer function analysis for PA handover

The SSEC in Figure 4.19 will be the starting point for the TF analysis in this section to determine the impact of the transient effects as a result of PA handover. That is, the expression for $\frac{i_{out}}{v_{in}}$ will be determined.

Based on current balance in the drain node and substitution, this expression can be solved to yield the result in Equation 4.5. The full derivation can be found in Appendix C.

$$\frac{i_{out}}{v_{in}} = \frac{-g_m + j\omega C_{GD}}{n\left(j\omega C_{GD} Z_S + j\omega C_{GS} Z_S + 1\right)\left(\frac{Z_L}{n^2 r_o} + \frac{j\omega C_{GD} Z_L}{n^2} + \frac{j\omega C_{GD} Z_L Z_S (g_m - j\omega C_{GD})}{n^2 (j\omega C_{GD} Z_S + j\omega C_{GS} Z_S + 1)} + 1\right)} \tag{4.5}$$

To quantitatively evaluate the result in Equation 4.5, some typical values for the NMOS output resistance and parasitic capacitances will now be substituted into the expression. Also the results from the previous sections for the transconductance and matching ratio will be substituted.

Based on a characterization of an NMOS design model for the 22FDX technology available to the author, a representative value of 10 fF will be chosen for the parasitic capacitances. An analysis of the impact of this decision on the final results will be provided. The input voltage source resistance will be set to 50 Ω and the load resistor will be set to 25 Ω which is half of the typical 50 Ω due to the use of the symmetry plane to capture the transition from differential to single-ended [18].

The choice for the output resistance of the NMOS transistors is more difficult. Typically, this resistance is modelled to be inversely proportional to the drain current multiplied by a factor λ that represents the relative length of the NMOS channel [38], but in a class B PA the devices are not biased with a constant current. To benefit model simplicity, the output resistance for the NMOS devices will still be modelled as a constant resistor of which the value will be set to $1.0\ k\Omega$. A more precise value for the output resistance depends strongly on the exact amplifier implementation and can be substituted in the results later if more model accuracy is desired.

Table 4.4 provides an overview of the used parameter values for the large and small PA.

Table 4.4: Parameter values used for the evaluation of the TF in Equation 4.5.

Large/small PA	Parameter	Parameter value
Both PAs	Z_L	$25\ \Omega$
	Z_S	$50\ \Omega$
	r_o	$1.0\ k\Omega$
Large PA	g_m	0.65 S
	n	4.0
	C_{GD}	$10\ fF$
	C_{GS}	$10\ fF$
Small PA	g_m	0.27 S
	n	1.7
	C_{GD}	$10/5.5\ fF$
	C_{GS}	$10/5.5\ fF$

After substitution of the parameters in Table 4.4 into Equation 4.5, a bode plot of the magnitude and phase transfer for both PAs can be made, refer to Figure 4.20.

Several things can be discerned from Figure 4.20 that are worth highlighting:

- The TF magnitude for low frequencies (e.g. roughly at frequencies lower than 10 GHz) is equal for both PAs where it was designed to be $20 \cdot log_{10}(g_m/n) \approx -15.8\ dB$.
- For frequencies higher than 10 GHz, low-pass behaviour can be recognized in the plot. Note that the fall-off for the large PA starts earlier. The larger parasitic capacitances for the large PA limit the bandwidth stronger than in the case of the smaller PA. At the RF carrier frequency, the difference in transfer magnitude is roughly 0.2 dB with the parameters from Table 4.4.

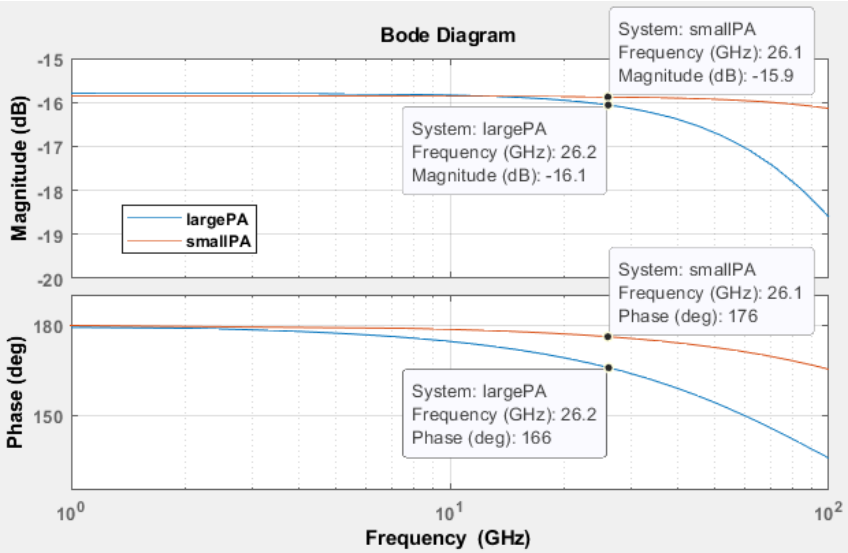


Figure 4.20: Bode plof of the magnitude and phase transfer for the TF in Equation 4.5 using the parameters in Table 4.4. Values at the targetted RF carrier frequency of 26 GHz have been annotated.

- Due to the same bandwidth limitations as in the previous point, there is also a difference in phase transfer for both TFs. This difference is equal to 10 degrees at 26 GHz.

The mentioned differences impact PA settling times. Additional analysis into this effect has been conducted in Appendix D and shows that the small PA settles a rough factor 4 faster than the large PA, based on the results in Figure 4.20. As a result, if the small PA would be enabled while the large PA has to be disabled, the smaller PA would settle faster, leading to an increase in the combined output current, because the large PA is slower in turning off. Similar transient effects occur when disabling the small PA and enabling the large one.

Appendix D also contains additional analysis into the impact of the parameter values in Table 4.4 on the result in Figure 4.20. The results show that especially the combination of the input voltage source impedance and the gate-source capacitance exerts low-pass behaviour. This also confirms the difficulties that were described in Section 4.3.4 regarding the implementation of the amplifier/driver that has to disable the PA. Impact of variations in the NMOS parasitic output resistance was found to be small.

The next Section will determine the actual distortion as a result of PA handover, using the TFs as derived in Equation 4.5.

Impact of PA handover effects on EVM and ACPR

Using the TF in Equation 4.5 and the parameters in Table 4.4, the impact of the transient effects due to PA handover on the EVM and ACPR of a 5G NR transmitter will be determined. To that end, a Simulink model has been designed to perform a measurement of both parameters.

The Simulink model has been kept as simple as possible to benefit insight into model behaviour and to separate PA handover effects from other distortion mechanisms. In addition to the baseband functionality as visualized in the Simulink model in the beginning of this chapter (refer to Figure 4.2), the model will include the following functionality:

- Upsampling and Finite Impulse Response (FIR) interpolation and corresponding blocks in the receiver. Interpolation is typically used in 5G transmitters to ease the requirements on the AA-filter and since it directly impacts the magnitude of the 'steps' between subsequent samples it will also directly impact the distortion generated during PA handover. To keep the distortion measurements realistic, it was decided to also integrate upconversion and FIR interpolation into the Simulink model.
- Upconversion and downconversion to evaluate the impact of PA handover at the targeted RF frequency of 26 GHz.

- Synchronization and demodulation blocks in the receiver to allow for an EVM measurement along the requirements as dictated in the 3GPP 5G NR specifications [2].

Additional model details for the used Simulink model can be found in Appendix A.

Now, to be able to integrate the continuous time TF into a discrete time Simulink model, the TF in Equation 4.5 has to be made discrete numerically at the Simulink sampling rate. Using zero-order hold functions at the targeted sampling rate, the continuous time TF can be converted to a discrete time TF. For the TF in Equation 4.5 and the parameters in Table 4.4, this yields the following discrete TF for the large PA:

$$\frac{i_{out}}{v_{in}}(z) = \frac{-0.1319 \cdot z^2 + 0.009748 \cdot z + 3.129 \cdot 10^{-5}}{z^3 - 0.2616 \cdot z^2 + 0.01421 \cdot z - 1.972 \cdot 10^{-19}} \tag{4.6}$$

Using a similar approach, a discrete TF can also be numerically determined for the small PA.

Both PA TFs can now be incorporated into the Simulink model using the discrete TF Simulink function block. Together with the other model functionality, this completes the basic model implementation.

Before proceeding with a discussion on the simulation results, the baseline EVM and ACPR values will be determined in Simulink. The EVM is expected to be zero as no other distortion mechanisms have been incorporated into the simulation. Simulated EVM was equal to $2.784 \cdot 10^{-14}$ % which is indeed practically zero. A plot of the upconverted spectrum allows determining the ACPR, refer to Figure 4.21.

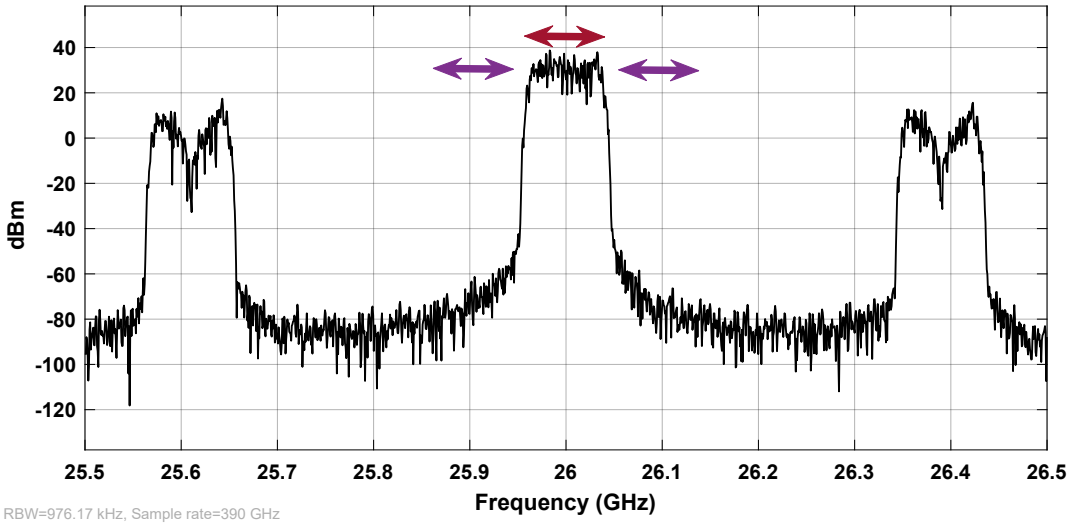


Figure 4.21: Transmitted spectrum zoomed around the RF carrier frequency of 26 GHz to determine the adjacent channel power ratio. No transient effects have been implemented. The red arrow indicates the channel power, the purple arrows the adjacent channels. ACPR in this baseline case is equal to -49 dB.

The ACPR before introduction of PA handover effects is equal to -49 dB. Notice that Figure 4.21 also shows the aliases as a result of the digital-to-analog conversion process.

Running the simulation after implementation of the discrete TFs for the large (as in Equation 4.6) and small PAs yields the spectrum in Figure 4.22.

Figure 4.22 shows that the ACPR has increased from -49 dB to -37 dB which is a significant increase of 12 dB. Note that the ACPR after implementing transient effects is still below the requirement of -30 dB (refer to Chapter 2).

Now, looking at the in-band signal quality, the constellation diagram of the received QAM symbols has been plotted in Figure 4.23.

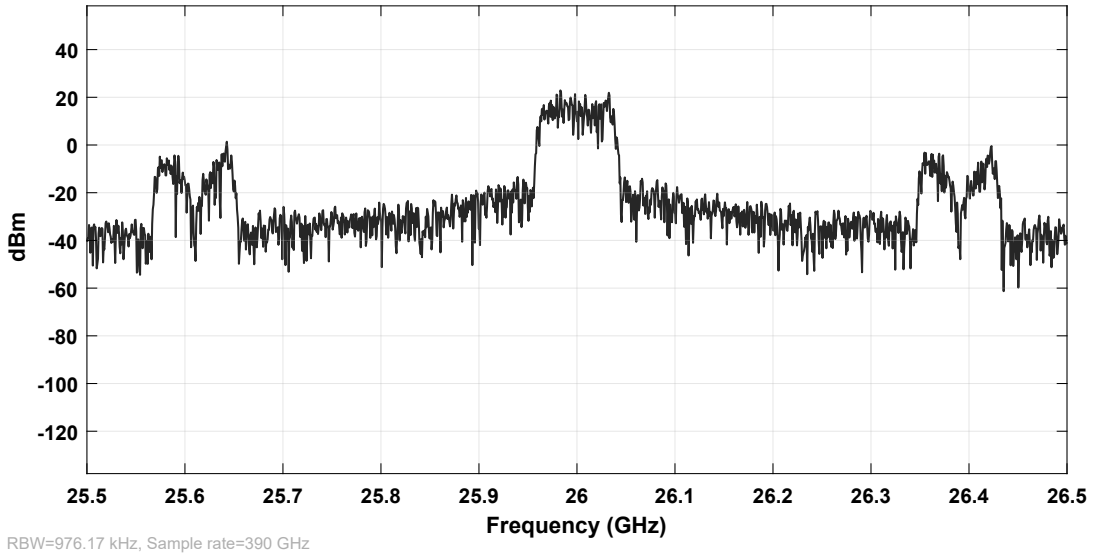


Figure 4.22: Transmitted spectrum zoomed around the RF carrier frequency of 26 GHz to determine the adjacent channel power ratio after implementing transient effects due to PA handover. ACPR has increased to -37 dB.

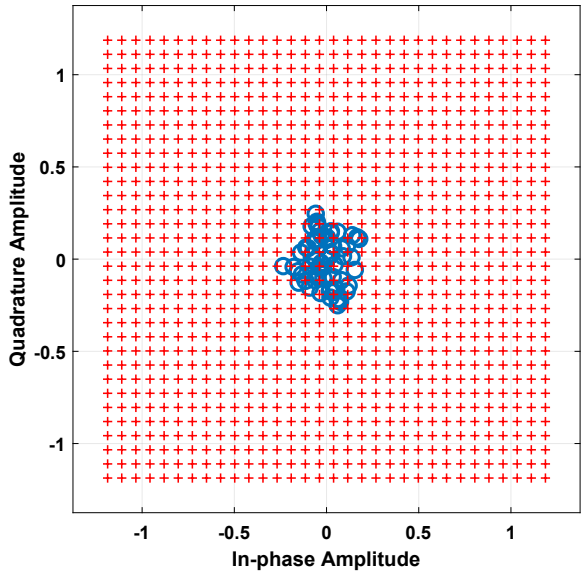
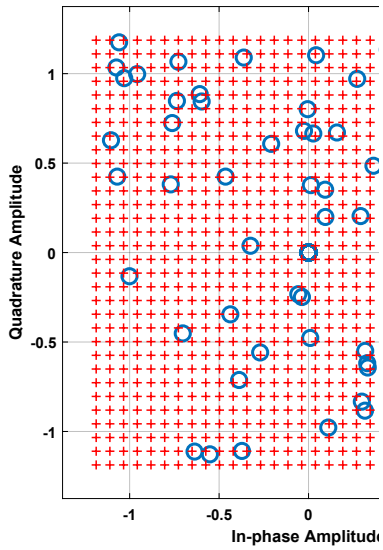
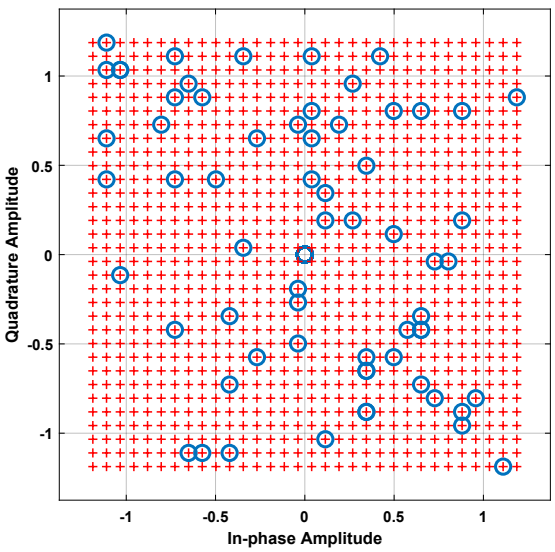


Figure 4.23: Constellation diagram of the received QAM-symbols after implementing transient effects due to PA handover. The reference constellation has been added in red. Received symbols have a phase rotation and a gain mismatch.

The plotted constellation points in Figure 4.23 clearly show that both a phase and a gain mismatch are present in the received symbols. Referring back to the bode plot for both PAs in Figure 4.20, gain and phase changes are to be expected at 26 GHz. Along the 3GPP requirements for 5G NR transmitters, frequency (and therefore also phase) and gain corrections may be made in the receiver before performing an EVM measurement [2]. Therefore, in the Simulink model, both the gain and phase mismatches will be compensated for as much as possible. The constellation diagram after correction has been plotted in Figure 4.24a. For reference, the original intended constellation has also been plotted, refer to Figure 4.24b.



(a) Received QAM symbols plotted on a constellation diagram as depicted in Figure 4.23 after gain and phase synchronization. Note that the differences in the TFs for both PAs make that the phase cannot be fully synchronized.



(b) Constellation diagram of QAM symbols as transmitted in the Simulink model, can be used as a reference constellation.

Figure 4.24: Visualization of the received QAM symbols after phase and gain correction and the intended symbols plotted in a constellation diagram.

After correction, the simulated EVM RMS value was found to be 4.2 %, which is far over the in Chapter 2 specified maximum EVM of 1.0 %. The root cause of this problem is a type of 'discrete' AM/AM and AM/PM conversion; when the large PA is active the samples experience a different overall phase and magnitude transfer then when the small PA is active. As a result, when increasing the signal's amplitude from a value below the threshold to a value above the threshold for PA handover, an additional step in phase and gain occur (which is why one might refer to it as 'discrete' AM/AM or AM/PM conversion). A similar effect can be witnessed when decreasing the magnitude from a value above the threshold to a value below the threshold.

Because there are two different phase and gain transfers for a subset of the samples, it is impossible to fully correct these discrepancies in the receiver with a single phase and gain correction. However, since the error is predictable, it may be possible to add pre-distortion in the *transmitter* to correct for the error, trading system complexity for signal linearity. Note that this only corrects the in-band signal quality and that the differences in PA TF still persist, leading to the transient effects as plotted in Figure 4.18 and degradation of the ACPR.

An overview of the simulation results after implementation of PA handover without any additional error correction in the transmitter has been included in Table 4.5.

Table 4.5: Overview of simulation results on EVM and ACPR performance after integration of PA handover behaviour.

Performance parameter	**Baseline model**	**No gain/phase corr. in RX**	**Gain/phase corr. applied in RX**
ACPR	-49 dB	-37 dB	-37 dB
EVM	0.0 %	31 %	4.2 %

Considering that the simulations as performed show violation of the EVM requirement while only considering differences in the PA implementation and not considering differences elsewhere in the transmitter (such as in the matching networks and predrivers), it is clear that the in-band signal quality has to be improved. There are several options to reduce the EVM:

- Use the aforementioned digital pre-distortion in the transmitter to compensate for the difference in gain and phase transfer for both PAs. This is technique that is typically applied in similar transmitter concepts, such as the beforementioned ET PAs [18].
- Design the combination of the predriver and PA for both hardware tracks such that the TFs are as similar as possible (e.g. it might be possible to optimize the output resistance of the predriver for optimum linear signal transfer).
- Implement measures to smoothen the PA handover. It may for instance be possible to have the next hardware track settle *before* it is connected to the output, again trading system complexity for signal linearity.

Even if signal linearity could be improved to satisfy the linearity requirements as dictated in Chapter 2, the additional hardware required to make it work will degrade a significant portion of the gained efficiency. The next section will cover efficiency-related considerations.

4.4.2 Efficiency-related considerations

Section 4.3.3 showed the ideal efficiency improvement using a dual class B setup to benefit from time-domain signal splitting. In a realistic system, the additional hardware required will mitigate (part of) the gained efficiency. This section will evaluate the additional hardware that is required and how these components impact overall efficiency. Due to the limited time-frame available for this thesis, only proof-of-concept calculations will be provided.

Figure 4.25 provides an overview of a time-domain splitting transmitter. The additional hardware required has been given a shade of red, hardware typically present in a traditional OFDM transmitter has been greyed out.

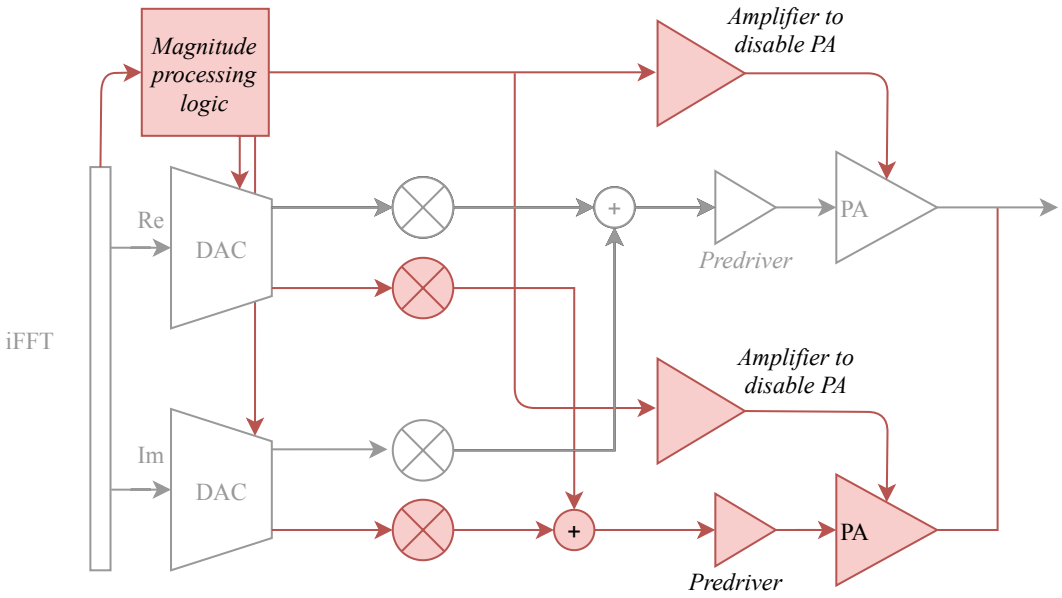


Figure 4.25: Overview of additional hardware that is required to profit from time-domain signal splitting.

Digital logic
The required digital logic consists only of a digital comparator that checks the magnitude of the incoming sample to stored reference values. As the supply voltage is only 0.8 V in 22nm FDSOI, the power consumption for every switch of the digital logic is relatively low (it scales with $P = f \cdot C_{eff} \cdot V_{DD}^2$ [39]). Therefore, the additional power consumption required for the digital processing logic will be assumed negligible in comparison to the power consumed by the amplifiers.

Additional mixers
When it comes to the extra mixers, the RZ for the unused DAC outputs limits the additional power dissipation. Because no current flows through the unused mixers, this current does not cause dissipation

in the mixer's on-resistance. Both 'unused' mixers are, however, still connected to the Local Oscillator (LO) driver leading to additional dynamic power consumption due to switching.

The dynamic power loss in the unused mixer transistors can be approximated using $P_{dynamic} = f_{LO} \cdot C_{eff} \cdot V_{pp-LO}^2$ based on analysis in [31]. Modelling the parasitic capacitance of each mixer with 10 fF, using an LO swing of 0.8 V and an LO frequency of 26 GHz, the additional dissipation for both extra mixers can be estimated to be roughly 0.3 mW. This is approximately 2 % of the power consumed by the PA at average output power[8].

Second predriver
Only after the exact transmitter requirements and PA implementation details are known, the design requirements for the predriver can be determined. For now, using similar reasoning for the design of the class A predriver as Razavi uses in [5] for the transceiver design example, the power consumption of the predriver will be estimated.

The predriver in [5] uses an output inductor to create a resonance tank with the input capacitance of the main PA for higher efficiency. Using the 10 fF input capacitance for the PA from Table 4.4, the required inductance is roughly 3.7 nH for resonance at 26 GHz. Assuming that a Q of 8 as used by Razavi in [5] is also reasonable for inductors in the 22FDX technology, the inductor will have a parallel resistance of roughly 5.0 $k\Omega$. To create a voltage swing of 0.8 V_{pp}, this would require a constant bias current of roughly 0.2 mA[9]. Combined with a supply voltage of 0.8 V, the additional power consumption for the predriver is estimated to be 0.16 mW, which is roughly 1 % of the power consumed by the PA at average output power.

Amplifiers to allow disabling unused PA
To be able to disable the unused PA, an additional amplifier is required to pull the NMOS gates of the unused devices to ground. As Section 4.4.1 has shown, this additional PA needs to be fast to avoid excessive transient effects. Section 4.3.4 already discussed the difficulties related to the implementation of this amplifier/driver.

Due to these implementation difficulties, there is no design example on which an estimation of the required power consumption could be based. In case one would settle with a much smaller bandwidth (e.g. at most 20 MHz), the typical efficiencies of existing ET PAs could serve as basis. Some recent examples achieve an efficiency of the supply amplifier of roughly 60 % [35], [36], also in the mmWave frequency range [37]. However, these efficiency figures are not representative for an amplifier based on a much larger bandwidth.

To be able to provide some indication nonetheless, based on the efficiency gain as presented in Section 4.3.3, the remaining power consumption budget available for disabling the PAs before the power gained because of using 2 tracks is cancelled completely will serve as basis.

Remaining power budget for the amplifiers that allow for disabling the PA
Note that, as for the calculation to determine the gain in efficiency, no linearity back-off will be considered. Without time-domain signal splitting and using only a single class B PA, the average PA efficiency was determined to be 19 % in Section 4.3.3. Using the average output power of 10 dBm, the average power consumption by the PA is 53 mW. After implementation of time-domain signal splitting, the power consumption drops to the ideal case of 26 mW (based on 39 % efficiency), which is a reduction of 27 mW.

Subtracting the additional dissipation in the extra mixers and the second predriver leaves 26.5 mW for any required Digital Pre-Distortion (DPD) hardware and the amplifiers/drivers to allow for disabling the main PAs before completely annulling the gained efficiency.

[8] Using Figure 4.16, at 10 dBm output power, the efficiency is 55 % ignoring required linearity back-off, leading to a PA power consumption of 18 mW.

[9] Note that to satisfy linearity demands and avoid compression, the required bias current is typically higher.

4.5 Comparison to existing techniques and chapter conclusion

The research in this Chapter has shown that splitting an OFDM signal along the time-domain reduces the PAPR of the split signals at the input of the PAs of each hardware track, refer to Figure 4.1. Through dedicating the baseband samples after the iFFT to one of multiple of these hardware tracks based on their magnitude, significant PAPR reduction could be achieved. Figure 4.7 provided a summary of the outcomes of the analysis conducted in this Chapter on achieving lowest overall PAPR for each of the hardware tracks. Already with 1 additional hardware track, average PAPR can be reduced by 4.6 dB, which is considerable in comparison to the digital techniques as presented in Chapter 2, refer to Table 2.1. Analysis in Section 4.2 has shown that the PAPR can be reduced even further by using more hardware tracks. To benefit simplicity, a transmitter based on two hardware tracks was considered in the rest of the chapter.

Further research on an example transmitter with two hardware tracks has shown that to be able to profit from the reduced PAPR at the input of each of the PAs, these PAs must be designed as push-pull class B amplifiers. Class A amplifiers were also considered, but are more inefficient for power BO than class B devices and had to be disabled through pulling their entire supply to ground instead of the NMOS gates, refer to Section 4.3.3. The impact of the 4.6 dB PAPR reduction on the typical efficiency performance of class B PAs was evaluated in Section 4.3.3, which showed the efficiency could be increased from 19 % for a single PA to 39 % for the dual class B set-up (assuming no power back-off for linearity). Given that the unused PA can be disabled, most of the additional hardware required with respect to a traditional OFDM transmitter was shown to consume only a few percent of either PA's power budget, refer to Section 4.4.2.

However, research has also shown that the efficiency improvement comes at a cost. First, because of switching between the two PAs, additional transient effects are introduced that have a negative impact on the signal's linearity and on spectral pollution. A method has been presented to be able to evaluate the impact of transient effects on EVM and ACPR performance in a transmitter model for 5G NR applications. TF analysis based on this method has shown that even minor differences between both PAs lead to considerable linearity degradation, refer to the results in Table 4.5. For instance, the differently sized parasitic capacitances for both PAs already increase the EVM to 4 %, violating the requirements presented in Chapter 2. DPD techniques will be required to bring the linearity performance back to acceptable levels. Any differences between the PA will also increase spectral pollution.

Second, overhead hardware is required to allow for switching between the hardware tracks. Unfortunately, the sensitivity of this design to transient effects requires fast disabling and enabling of the PAs. This demands a driver that is capable of quickly pulling the NMOS gates of the unused PA to ground requiring a strong driver with a large bandwidth, refer to Section 4.3.4. A similar design problem can be identified in ET PAs where a driver is required to provide discrete levels in the PA's supply. An overview of ET PA drivers presented in Section 4.3.4 showed that such devices typically achieve only tens of MHz of bandwidth, falling short of the required hundreds of MHz for the 5G NR application at hand.

Based on this observation, the preliminary conclusion is that time-domain signal splitting can be very attractive as a PAPR reduction technique, but that bandwidth limitations in the design of a driver required to disable unused hardware should be considered a show-stopper. Since the results are promising, recommendations will be made in the Discussion chapter to further explore if different implementations may be possible to realize a workable time-domain signal splitting transmitter. A power budget for this alternative implementation has been provided in Section 4.4.2 for future researchers to evaluate the efficiency of their ideas.

Assuming that such a driver could be developed with additional research, the concept as presented in this Chapter can be compared to existing PAPR reduction techniques, refer to Table 4.6.

The results in Table 4.6 show that only companding techniques achieve higher PAPR reduction when considering a time-domain splitting transmitter with two hardware tracks. However, the additional hardware required for the concept presented in this Chapter is significant (2 additional mixers, extra pre-driver + PA and overhead hardware) which makes that it scores low on implementation complexity[10]. BER performance may be improved by implementing the mentioned DPD techniques. This makes time-domain

[10] Mainly because no straightforward implementation for the driver to disable the PAs has been found.

Table 4.6: Comparison of PAPR reduction techniques, including frequency- and time-domain signal splitting.

PAPR reduction technique	**PAPR improvement**	**Implement. complexity**	**Bandwidth expansion**	**BER degradation**	**Power consumption**	**Spectral spillage**
Clipping and filtering	$+/-$	$+$	$++$	$-$	$++$	$--$
Coding schemes	Varies	$+/-$	$-$	$+/-$	$+$	$++$
PTS and SLM	$+/-$	$--$	$++$	$+$	$-$	$++$
Nonlinear companding schemes	$++$	$+/-$	$++$	$+/-$	$+$	$+$
TR and TI	$+/-$	$-$	$--$	$+$	$+/-$	$+$
Signal splitting (f-domain)	$+/-$	$--$	$++$	$+/-$	$--$	$++$
Signal splitting (t-domain)	$+$	$-$	$++$	$+/-$	$+$	$-$

splitting a serious alternative to companding schemes given the stringent linearity requirements for 5G NR applications, but only if an efficient integration of PA disabling can be realized.

When comparing the results presented in this Chapter to techniques aimed at reducing the negative impact of a high PAPR on the PA, the resemblance to ET PAs should be highlighted. In terms of efficiency improvement, both techniques achieve similar performance. However, in time-domain splitting, the additional hardware presents additional optimization opportunities, as the smaller PA may be optimized further for the smaller output power it has to generate. In ET, only one PA is used, allowing no dedicated optimization. Also, note that in ET, the supply voltage is varied. It may be easier to disable an unused PA through pulling the amplifier's gates to ground, but more research into this issue is required, refer to the Discussion Chapter.

Finally, the resemblance to a Doherty PA deserves some attention. In both cases PAs are disabled or enabled based on the incoming signal's magnitude. A Doherty structure, however, relies on advanced loading techniques to present optimum load impedance during operation; the peaking amplifier will assist the main amplifier for larger amplitudes. In time-domain signal splitting, only 1 PA is active at a certain time instant and ideally no loading occurs. This prevents the need for an advanced output network but does require additional hardware of which the consequences have been discussed in this Chapter.

5 DISCUSSION

The research as presented in Chapters 3 and 4 has provided an answer to the main research question posed in Chapter 1. Sections 3.5 and 4.5 contain an overview of the factual results used to determine if the combination of signal splitting and parallel hardware can be used to effectively alleviate the negative impact of a high PAPR on 5G NR transmitter performance. The goal of this Chapter is to interpret the results and provide a discussion on their implications and limitations. Recommendations for further research will conclude this chapter.

5.1 A discussion on frequency-domain signal splitting

Section 3.5 provided an overview of the findings regarding the use of frequency-domain signal splitting in transmitters. Analysis into the statistical properties of OFDM signals showed that the PAPR can be reduced with at most 4 dB by splitting up the signal into smaller portions. This is a significant reduction and in line with typical figures achieved by other PAPR reduction techniques that were presented in Chapter 2.

However, Section 3.5 has also shown that to be able to profit from this reduction, excessive amounts of hardware would be required within the in Chapter 2 presented 5G NR transmitter design space. For optimal PAPR reduction, only 1 subcarrier should be converted with a single DAC. Considering that the benefits of reduced PAPR yielded only a factor of 2 reduction in the area of the individual DACs, a large number of DACs occupying significant chip area would be required to support the 5G transmitter requirements. The research in Chapter 3 has also shown that combining the many parallel DAC outputs introduced its own efficiency degrading problems.

Now, in interpreting these results, three root causes have been identified that are at the foundation of these problems. These are:

1. As a result of the signal properties of OFDM signals and the QAM signals with a high modulation factor that form them, PAPR increases rapidly when combining a few subcarriers and then starts to saturate. For example, the analysis in Section 3.2.3 has shown that increasing the number of combined subcarriers from 1 to 2 doubles the PAPR, but that it already saturates after combining only 8 subcarriers.
2. The large bandwidths required in 5G NR transmitters dictated the use of current-steering DACs, as discussed in Section 3.3.1. The class A conduction angle of this type of DAC limits efficient implementation of the analog hardware required to combine the parallel DAC outputs.
3. Since signal splitting was carried out along the frequency-domain by splitting up the OFDM signal into groups of one or multiple subcarriers, a situation occurs where multiple independent sources as a result of subcarrier orthogonality concurrently drive the same load.

It is important to highlight the implications of these root causes.

The fact that the PAPR saturates already with a few subcarriers is problematic to the implementation of systems relying on a significant number of subcarriers. However, systems with a number of subcarriers in the order of 10 to 20, might benefit from this approach, as both the required number of DACs to reduce the PAPR and the required amount of analog hardware for signal combination would then be within feasibility limits. Also, because of this steep PAPR increase, the option as discussed in Section 3.2.4 to convert only the QAM amplitude levels in each DAC may be attractive for systems with only a few subcarriers as it offers a significant reduction in required DAC resolution. It should however be noted

that these systems still suffer from the other 2 root causes.

Next to PAPR reduction, a second opportunity for frequency-domain signal splitting that has been explored was aimed at rendering the inefficient PA stage unnecessary. However, this idea suffered specifically from the last two fundamental problems.

Following from the second root cause, a class A conduction angle for the DACs results in significantly larger currents flowing to the load. These currents were found to introduce additional dissipation in the mixer transistors. Now, as discussed in Section 3.1.1, the use of hybrid mixing may ease the bandwidth requirements for the DACs such that alternative, more efficient, DAC implementations become possible. Researchers that see use in frequency-domain splitting to improve efficiency performance are advised to look into this opportunity. Implications of this design decision on the design of the digital hardware will have to be evaluated as only basic attention has been dedicated to digital hardware design in this thesis.

Future researchers should, however, be aware of the third root cause when optimizing the DAC design. Even if each DAC can be designed such that it only provides useful output current, there will still be cancellation of part of these currents in the load as the sources are independent with respect to each other and will balance out in the load. This problem can be minimized by reducing the number of parallel DAC outputs, but, as was shown in Section 3.3.2, it will be challenging to achieve sufficient linearity in the mixers when these larger signals have to be processed[1].

To circumvent these problems with orthogonal sources, instead of splitting the OFDM signal into groups of subcarriers, other splitting methods may be more useful. One could think of the use of 'look-ahead' hardware to synchronize control of the DAC outputs and avoid driving conflicts. Considering the conclusions of the research on frequency-domain signal splitting, these implementations may be useful. However, future researchers looking to make use of such a transmitter structure are advised to solve three problems related to the transmitter concept presented in Chapter 3 first to guarantee feasibility of this system concept. First, one should make sure that the DACs work together, to avoid driving conflicts. Second, the mixers should be linear enough to handle the relatively large signal power. Finally, it should be explored if the use of class A DACs can be avoided.

5.2 A discussion in time-domain signal splitting

Chapter 4 has evaluated the use of time-domain signal splitting to reduce the high PAPR of OFDM signals. Analysis has shown that the PAPR can be reduced by 4.6 dB through the use of an additional hardware track, which is a stronger reduction than most digital techniques that were presented in Chapter 2. To make time-domain signal splitting work, additional hardware is required, both to realize the additional hardware track as well as for some overhead hardware to facilitate switching. The results have been presented in Section 4.5.

Interpreting these results, two key issues were identified in the implementation of time-domain signal splitting transmitters. First, the switching between the two hardware tracks introduces additional transient effects that were shown to degrade in-band signal linearity and out-of-band pollution. Second, the additional hardware required also consumes power and limits the gain in efficiency.

Especially the combination of high bandwidth requirements and a stringent linearity specification for 5G NR transmitters as included in Section 2.5 has severe implications for these key issues. As the TF analysis in Section 4.4.1 analyzed, any differences between the hardware tracks will show up as distortion in the load and DPD techniques will be required to meet the linearity specification[2]. In addition, to minimize this distortion as much as possible, the process of enabling one PA while disabling the other should be fast, to keep the transition period as short as possible. In combination with the large signal bandwidth, this requires fast drivers/amplifiers to disable/enable hardware tracks, as was discussed in Section 4.3.4.

[1] It is important to note here that readers should recognize the limitations of the research on mixer linearity as presented in Section 3.3.2, as it lacks results on bootstrapped mixer implementations.

[2] In the analyzed case only differences in transistor size for both PAs were considered. Even with perfect impedance matching networks and no other effects evaluated, the EVM increased to an unacceptable value of 4.0 %, refer to Section 4.4.1.

Within the timeframe of this thesis, only little research has been conducted into possibilities to disable or enable hardware tracks. The similar increase in efficiency performance and bandwidth demands with respect to ET PAs [18] are strong indicators that a similar problem context can be identified for possible drivers. A short overview of ET PAs has shown that the high signal bandwidth proves problematic to feasible implementation.

Since the PAPR reduction is promising and, aside from the linearity considerations, only few issues with time-domain signal splitting transmitters were identified, future research into possibilities to disable/enable hardware tracks is recommended. One possibility would be to explore if the DAC's RZ connection may be sufficient to disable the unused class B PA, refer to Figure 5.1.

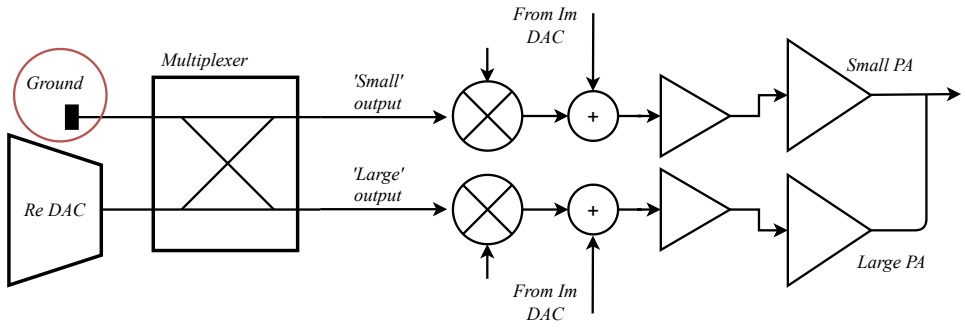


Figure 5.1: An alternative to using a separate driver to disable the unused PA may be to use the RZ connection of the DAC. More research is required to determine the feasibility of this option and to determine the impact of the mixer transistors and the relatively long interconnect.

The concept presented in Figure 5.1 may be sufficient to 'disable' the unused PA as the class B topology does not have a quiescent current flowing, eliminating the need for a dedicated and inefficient driver. However, more research into for instance the combination of the large PA input capacitance and the relatively long interconnect between the ground connection and the PA input (including mixer transistors) is required. Additional research into this problem is recommended.

On a final note, if the impact of transient effects on the transmitter's performance proves too great, an additional option to reduce transient effects would be to switch less often. Currently, usage of both hardware tracks has been optimized to yield lowest overall PAPR, refer to Section 4.2.2. A different optimization strategy is possible that limits the amount of switching between tracks at the cost of higher PAPR. One could for instance only occasionally help the very large peaks with a dedicated PA. If implementation of DPD techniques to bring the distortion to acceptable levels proves cumbersome, researchers are advised to see if reducing switching activity may ease the requirements to allow for a feasible implementation.

5.3 Summary of recommendations for further research

Based on the remarks in this Chapter, there are four recommendations for future research on PAPR reduction through signal splitting.

- Frequency-domain signal splitting may be used to generate the required signal power at baseband, without requiring a PA. To avoid needing class A current-steering DACs that degrade efficiency, the in Section 3.1.1 discussed hybrid subcarrier mixing method may be used to ease the DAC's requirement to allow for different implementations. The implications of this option should be investigated.
- An alternative splitting method to frequency-domain signal splitting may also present an opportunity to omit the inefficient PA stage. Investigation into synchronized DAC outputs to avoid driving conflicts is recommended.
- Time-domain signal splitting proves an interesting alternative to existing digital PAPR reduction techniques. Additional research into disabling/enabling hardware tracks is recommended to see if an efficient transmitter implementation can be realized.
- In case linearity requirements prove problematic in time-domain signal splitting transmitters, additional research into decreasing the switching activity to reduce transient effects is recommended.

6 CONCLUSION

With the introduction of the 5th generation mobile networking technology, mobile communications can be enabled in many new domains. As the usage of telecommunication soars, the demand for more efficient implementations grows. However, one of the major bottlenecks in the development of efficient 5G NR transmitters is the negative impact of a high PAPR on the transmitter's analog hardware. This relatively high PAPR is a result of the combination of many independently modulated subcarriers into an OFDM signal. The main aim of this thesis was to investigate to what degree the negative impact of a high PAPR as a result of the use of OFDM modulation on the transmitter's efficiency and linearity could be alleviated by making use of a combination of signal splitting and parallel hardware.

To that end, two specific system concepts have been explored. The first is based on splitting in the frequency-domain and uses decomposition of the OFDM signal into groups of multiple subcarriers to reduce the PAPR. Each of these signals is processed by its own analog hardware track of which the parallel outputs concurrently deliver power to the load. The fact that these parallel outputs are capable of delivering power concurrently presented an additional opportunity to investigate the possibility of omission of the inefficient PA stage typically present in the transmitter. The second system concept makes use of splitting in the time-domain. Based on the magnitude of the baseband OFDM signal, it is processed by the appropriate hardware track, resulting in PAPR reduction for the signals in each of these tracks.

Analysis on frequency-domain signal splitting has shown that to be able to profit from splitting in terms of PAPR, only a few subcarriers may be combined for each hardware track. This is a result of the signal properties of OFDM signals and the QAM signals with a high modulation factor that form them; the PAPR increases rapidly when combining a few subcarriers and then saturates quickly. These properties prove problematic when attempting to profit from the resulting reduction in PAPR as excessive amounts of hardware would be required to support the number of subcarriers required in 5G NR. Combined with the fact that the parallel outputs drive the same load concurrently with orthogonal signals resulting in partial cancellation of the generated power in the load, the use of frequency-domain splitting to alleviate the high PAPR in 5G NR transmitters is not recommended.

Research has also focused on exploring the feasibility of using frequency-domain splitting to render the inefficient PA stage unnecessary. Here, predominantly the strict linearity requirements for 5G NR transmitters complicate implementation. Analysis has shown that, in case the DACs would be used to collectively provide the required output power, the large signals flowing through the mixers in combination with these linearity requirements lead to problems with gain compression. In an attempt to reduce the currents flowing through the mixers, the use of more DACs has been investigated, but, again, the number of DACs required to meet the 5G linearity requirement in the investigated 22FDX technology is excessive. Also, since full power would be generated before upconversion, the dissipation in the on-resistance of the mixer transistors was shown to increase significantly. After evaluation of these arguments, one should conclude that using frequency-domain splitting in an attempt to omit the inefficient PA stage is not recommended either.

Then, the feasibility of using time-domain signal splitting to reduce the PAPR was investigated. Results from the statistical analysis that was conducted have shown that significant PAPR reduction can be achieved with relatively little additional hardware. Through using one additional combination of a mixer and PA and the required hardware overhead, the PAPR reduction was shown to exceed that of of most of the digital PAPR reduction techniques that were considered in this thesis. Though these results are promising, analysis of other performance parameters has surfaced difficulties in this concept's implementation. Switching between the hardware tracks was shown to introduce transient effects that degraded

both the signal's in-band linearity as well as the ACPR. The combination of the stringent linearity requirements in place for 5G NR transmitters offering only little headroom to linearity degradation and the large signals bandwidths requiring fast switching to limit generation of these transient effects, complicate especially the implementation of functionality to disable or enable the unused hardware tracks. Based on the resemblance of this design problem to considerations in the design of ET PAs, the preliminary conclusion in this thesis is that the large bandwidths required prevent a feasible implementation of this transmitter concept. However, the limitations to this research should be recognized and, as discussed in the Discussion chapter, more research into disabling and enabling the hardware tracks is recommended to come to a well-considered final verdict.

Considering the conclusions as presented in this chapter, the answer to the main research question is: no, the use of signal splitting and parallel hardware tracks does not benefit the negative impact of a high PAPR in 5G NR transmitters without significant degradation of other performance parameters. The root causes for the limitations that have surfaced in this thesis are the strict linearity requirements and the large signal bandwidth in 5G applications. If these specifications could be relaxed, especially the use of time-domain splitting may be an attractive alternative to other PAPR reduction techniques given its advantages and disadvantages as considered in this thesis.

REFERENCES

[1] Qualcomm. Everything you need to know about 5G.

[2] 3GPP. 38.101-2 NR; User Equipment (UE) radio transmission and reception; Part 2: Range 2 Standalone, January 2020.

[3] Tao Jiang and Yiyan Wu. An Overview: Peak-to-Average Power Ratio Reduction Techniques for OFDM Signals. *IEEE Transactions on Broadcasting*, 54(2):257–268, June 2008.

[4] W.H. Doherty. A New High Efficiency Power Amplifier for Modulated Waves. *Proceedings of the IRE*, 24(9):1163–1182, September 1936.

[5] Behzad Razavi. *RF microelectronics*. Prentice Hall, Upper Saddle River, NJ, 2nd ed edition, 2012.

[6] Maarten Ditzel, Ralph H. J. M. Otten, and Wouter A. Serdijn. *Power-aware architecting for data-dominated applications*. Springer, Berlin, 2007. OCLC: 836508155.

[7] Athanasios Papoulis and S. Unnikrishna Pillai. *Probability, random variables, and stochastic processes*. McGraw-Hill, Boston, Mass., 4. ed., internat. ed., nachdr edition, 2009. OCLC: 255904469.

[8] van Beek, P.C.W. *Multi-carrier single-DAC transmitter approach applied to digital cable television*. Technische Universiteit Eindhoven, 2011. OCLC: 8086866678.

[9] Vittorio Camarchia, Roberto Quaglia, Anna Piacibello, Duy P. Nguyen, Hua Wang, and Anh-Vu Pham. A Review of Technologies and Design Techniques of Millimeter-Wave Power Amplifiers. *IEEE Transactions on Microwave Theory and Techniques*, pages 1–1, 2020.

[10] Byungjoon Park, Sangsu Jin, Daechul Jeong, Jooseung Kim, Yunsung Cho, Kyunghoon Moon, and Bumman Kim. Highly Linear mm-Wave CMOS Power Amplifier. *IEEE Transactions on Microwave Theory and Techniques*, 64(12):4535–4544, December 2016.

[11] Fei Wang and Hua Wang. 24.6 An Instantaneously Broadband Ultra-Compact Highly Linear PA with Compensated Distributed-Balun Output Network Achieving >17.8dBm P_{1db} and >36.6% PAE_{p1db} over 24 to 40GHz and Continuously Supporting 64-/256-QAM 5G NR Signals over 24 to 42GHz. In *2020 IEEE International Solid- State Circuits Conference - (ISSCC)*, pages 372–374, San Francisco, CA, USA, February 2020. IEEE.

[12] Ebrahim B. Al-Safadi and Tareq Y. Al-Naffouri. Peak Reduction and Clipping Mitigation in OFDM by Augmented Compressive Sensing. *IEEE Transactions on Signal Processing*, 60(7):3834–3839, July 2012.

[13] R. Cavalcante and I. Yamada. A Flexible Peak-To-Average Power Ratio Reduction Scheme for OFDM Systems by the Adaptive Projected Subgradient Method. *IEEE Transactions on Signal Processing*, 57(4):1456–1468, April 2009.

[14] Martha C. Paredes Paredes and M. Julia Fernandez-Getino Garcia. Energy efficient peak power reduction in OFDM with amplitude predistortion aided by orthogonal pilots. *IEEE Transactions on Consumer Electronics*, 59(1):45–53, February 2013.

[15] Yasir Amer Al-Jawhar, Khairun Nidzam Ramli, Aida Mustapha, Salama A. Mostafa, Nor Shahida Mohd Shah, and Montadar Abas Taher. Reducing PAPR With Low Complexity for 4G and 5G Waveform Designs. *IEEE Access*, 7:97673–97688, 2019.

[16] Kasun Bandara, Atul Sewaiwar, and Yeon-Ho Chung. Efficient nonlinear companding scheme for substantial reduction in peak-to-average power ratio of OFDM. *Journal of Systems Engineering and Electronics*, 26(5):924–931, October 2015.

[17] Hua Zhang, Ying Yuan, and Wei Xu. PAPR Reduction for DCO-OFDM Visible Light Communications via Semidefinite Relaxation. *IEEE Photonics Technology Letters*, 26(17):1718–1721, September 2014.

[18] Steve C. Cripps. *RF power amplifiers for wireless communications*. Artech House microwave library. Artech House, Boston, 2nd ed edition, 2006. OCLC: ocm67728816.

[19] Leonard Kahn. Single-Sideband Transmission by Envelope Elimination and Restoration. *Proceedings of the IRE*, 40(7):803–806, July 1952.

[20] H. Chireix. High Power Outphasing Modulation. *Proceedings of the IRE*, 23(11):1370–1392, November 1935.

[21] David M. Pozar. *Microwave engineering*. Wiley, Hoboken, NJ, 4th ed edition, 2012. OCLC: ocn714728044.

[22] Fei Wang, Tso-Wei Li, and Hua Wang. 4.8 A Highly Linear Super-Resolution Mixed-Signal Doherty Power Amplifier for High-Efficiency mm-Wave 5G Multi-Gb/s Communications. In *2019 IEEE International Solid- State Circuits Conference - (ISSCC)*, pages 88–90, San Francisco, CA, USA, February 2019. IEEE.

[23] Narek Rostomyan, Mustafa Ozen, and Peter Asbeck. 28 GHz Doherty Power Amplifier in CMOS SOI With 28% Back-Off PAE. *IEEE Microwave and Wireless Components Letters*, 28(5):446–448, May 2018.

[24] Christian Schmidt, Christoph Kottke, Vicky H. Tanzil, Ronald Freund, Volker Jungnickel, and Friedel Gerfers. Digital-to-Analog Converters Using Frequency Interleaving: Mathematical Framework and Experimental Verification. *Circuits, Systems, and Signal Processing*, 37(11):4929–4954, November 2018.

[25] Marcel Pelgrom. *Analog-to-digital conversion*. Springer, Cham, third edition edition, 2017. OCLC: 965332185.

[26] Kyongkuk Cho and Dongweon Yoon. On the general BER expression of one- and two-dimensional amplitude modulations. *IEEE Transactions on Communications*, 50(7):1074–1080, July 2002.

[27] R. Esposito and L. Wilson. Statistical properties of two sine waves in Gaussian noise. *IEEE Transactions on Information Theory*, 19(2):176–183, March 1973.

[28] A. van den Bosch, M.A.F. Borremans, M.S.J. Steyaert, and W. Sansen. A 10-bit 1-GSample/s Nyquist current-steering CMOS D/A converter. *IEEE Journal of Solid-State Circuits*, 36(3):315–324, March 2001.

[29] A.R. Khodkhumbe. The design of a current-driven passive mixer in 22nm cmos - slides, May 2020.

[30] Behzad Razavi. The Current-Steering DAC [A Circuit for All Seasons]. *IEEE Solid-State Circuits Magazine*, 10(1):11–15, 2018.

[31] Friso v.d. Boom. Analysis and Design of an Up-Conversion Mixer for Mixer-Last Transmitters Using a 25% Duty-Cycle LO. University of Twente, January 2019.

[32] Erik Olieman, Anne-Johan Annema, and Bram Nauta. An Interleaved Full Nyquist High-Speed DAC Technique. *IEEE Journal of Solid-State Circuits*, 50(3):704–713, March 2015.

[33] Jaco du Preez and Saurabh Sinha. *Millimeter-wave power amplifiers*. Springer, Cham, 2017. OCLC: ocn989966266.

[34] Myoungbo Kwak, Jinseong Jeong, Muhammad Hassan, Jomei J. Yan, Donald F. Kimball, Peter M. Asbeck, and Lawrence E. Larson. High efficiency wideband envelope tracking power amplifier with direct current sensing for LTE applications. In *2012 IEEE Topical Conference on Power Amplifiers for Wireless and Radio Applications*, pages 41–44, Santa Clara, CA, USA, January 2012. IEEE.

[35] Dongsu Kim, Jinsung Choi, Daehyun Kang, and Bumman Kim. High efficiency and wideband envelope tracking power amplifier with sweet spot tracking. In *2010 IEEE Radio Frequency Integrated Circuits Symposium*, pages 255–258, Anaheim, CA, USA, 2010. IEEE.

[36] Muhammad Hassan, Myoungbo Kwak, Vincent W. Leung, Chin Hsia, Jonmei J. Yan, Donald F. Kimball, Lawrence E. Larson, and Peter M. Asbeck. High efficiency envelope tracking power amplifier with very low quiescent power for 20 MHz LTE. In *2011 IEEE Radio Frequency Integrated Circuits Symposium*, pages 1–4, Baltimore, MD, USA, June 2011. IEEE.

[37] Jonmei J. Yan, Calogero D. Presti, Donald F. Kimball, Young-Pyo Hong, Chin Hsia, Peter M. Asbeck, and James Schellenberg. Efficiency Enhancement of mm-Wave Power Amplifiers Using Envelope Tracking. *IEEE Microwave and Wireless Components Letters*, 21(3):157–159, March 2011.

[38] Behzad Razavi. *Design of analog CMOS integrated circuits*. McGraw-Hill, 2 edition, 2017. OCLC: 987434744.

[39] Sabih H. Gerez. System-On-Chip Design - Low-Power IC Design - Slides, September 2018.

A OVERVIEW OF USED SIMULINK AND LTSPICE MODELS

This Appendix covers modelling details for MATLAB/Simulink® and LTSpice models used for theory verification. The first section covers the models used in Chapter 3 on the frequency-splitting hardware, the second section covers the remaining models used in Chapter 4.

A.1 Simulink and LTSpice models used in Chapter 3

The following models will be covered in this Section:

1. Simulink baseband transmitter and receiver combination for signal splitting along the frequency domain
2. Simulink model used for verification of OFDM clipping and quantization distortion
3. Simulink model used for verification of the impact of signal splitting on clipping and quantization distortion
4. Simulink model used for verification of the system impact of spread in DAC current sources
5. LTSpice model used for verification of current leakage

A.1.1 Baseband transmitter/receiver combination

A Simulink model has been designed to cover basic signal flows in the frequency-splitting hardware. The research in this thesis focusses on the design of TX hardware, but to be able to verify data integrity, also a corresponding Receiver (RX) has been designed.

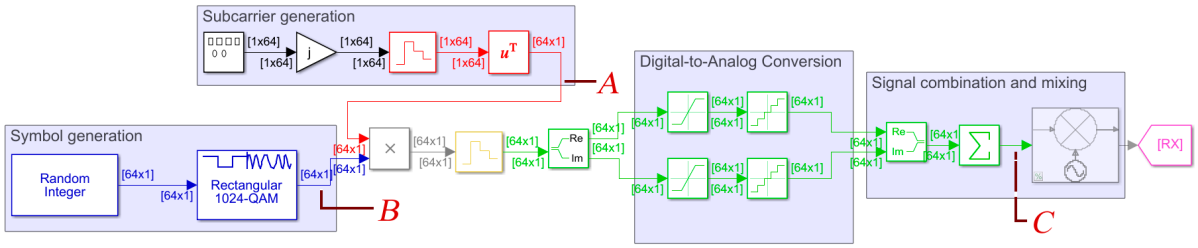
Figures A.1a and A.1b show the Simulink model.

Table A.1 shows the global system parameters that are in line with the typical requirements for OFDM-systems as discussed in Chapter 2.

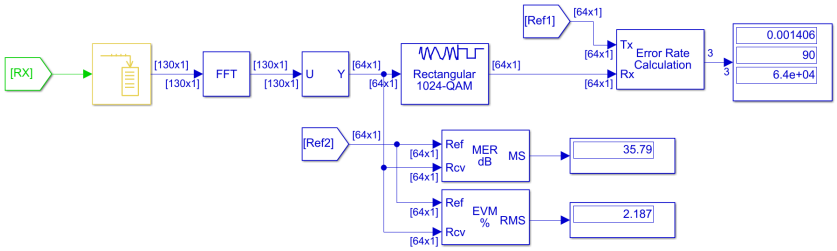
Table A.1: Global parameters for the baseband TX/RX combination.

Parameter	**Simulink variable**	**Value**
Number of subcarriers	n_subcarriers	64
Subcarrier spacing	f_spacing	$60e3$ [Hz]
QAM modulation order	M_QAM	1024
Average power of transmitted constellation	P_average_constellation	1 [W] (referenced to 1 Ω)

The Simulink model blocks have been set according to the values in Table A.1, see Tables A.2 and A.3. For both Tables holds that parameters without a table entry should be set to their default value.



(a) Simulink model of the transmitter.



(b) Simulink model of the receiver.

Figure A.1: Simulink model of the transmitter and receiver combination.

Table A.2: Simulink TX model parameters for Model 1.

TX/RX	Block name	Block parameter	Simulink value
TX	Subcarrier generation/Signal Generator	Amplitude	2
		Frequency	[1:n_subcarriers]*f_spacing
	Subcarrier generation/Gain	Gain	j
	Subcarrier generation/ZOH	Sample time	1/(f_sample_subcarrier)
	Symbol generation/Random integer generator	Set size	M_QAM
		Initial seed	0
		Sample time	1/f_spacing/n_subcarriers
		Samples per frame	n_subcarriers
	Symbol generation/Rectangular QAM Modulator Baseband	M-ary number	M_QAM
		Constellation ordering	Gray
		Normalization method	Average power
		Average power	P_average_constellation
		Phase offset	0
	ZOH	Sample time	1/(n_subcarriers*f_spacing*2 +2*f_spacing)
	DAC/Saturation (2x)	Upper limit	A_c*sqrt(P_average_constellation)
		Lower limit	-A_c*sqrt(P_average_constellation)
	DAC/Quantizer (2x)	Quantization Interval	2*A_c*sqrt(P_average_constellation)/N

Table A.3: Simulink RX model parameters for Model 1.

TX/RX	Block name	Block parameter	Simulink value
RX	Buffer	Output buffer size	n_subcarriers*2+2
	FFT	Checked parameters	'Divide output by FFT length' and 'Inherit FFT length from input dimensions'
	Selector	Index mode	Zero-based
		Index	[1:n_subcarriers]
		Input port size	2*n_subcarriers+2
	Rectangular QAM Demodulator Baseband	M-ary number	M_QAM
		Normalization method	Average Power
		Average power	P_average_constellation
	Error Rate Calculation	Computation delay	n_subcarriers
	EVM Measurement	Measurement interval	Entire history

A few parameters are out of the ordinary and deserve some discussion. Due to limitations in Simulink to

efficiently generate an array with amplitude values of subcarriers running at different frequencies, only the signal generator block could be used. However, it does not offer the option to change the sine-wave properties. To align the incoming subcarriers with the Fast Fourier Transform (FFT) requirements, the subcarriers are therefore multiplied with the complex number j. Two frequency components can be found at the positive and negative frequencies of the desired frequencies. To compensate for the halving of the amplitude, the subcarriers have amplitude value 2. The selector behind the FFT then selects the positive frequencies, which are used for decoding.

In the receiver, the incoming symbols are compared to the intended transmission symbols and several figures for data integrity are calculated: the SNR, EVM and BER. These implementations are relatively straightforward and the resulting values have been visualized using displays.

A.1.2 Model used for verification of OFDM clipping and quantization distortion

The model used to verify the clipping and quantization distortion contributions is very similar to the base-band model as included in Figure A.1. There is one difference.

The subcarrier mixing is done in the 'traditional' way using an iFFT. However, to match the situation as in Model 1 some signal preparation is done to also include signals at the corresponding negative frequencies. This does not change the result, as the selector block after the FFT omits these carriers again. Figure A.2 shows the iFFT section that directly follows the 'Symbol generation' section of Figure A.1a.

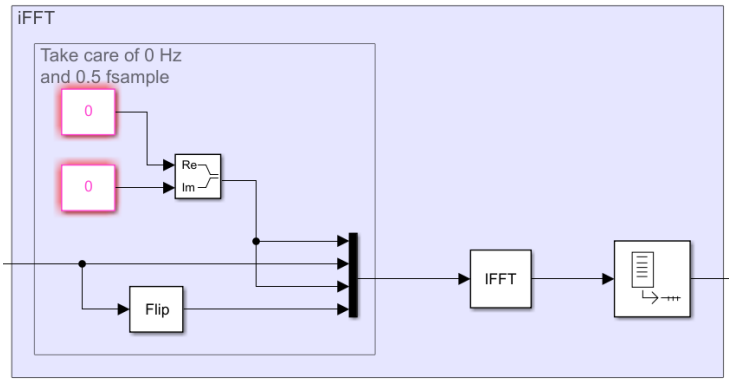


Figure A.2: The iFFT section that replaces the separate subcarrier mixing as in Figure A.1a.

The only box checked in the iFFT block is 'Inherit FFT length from input dimensions'.

A plot of the clipping and quantization distortion can be generated by measuring the SNR (Modulation Error Ratio (MER) in Simulink, but since no other errors are introduced in the model, the MER and SNR are equal) under a sweep of the clipping level. To measure the clipping or quantization distortion separately, the other block can be commented through.

A.1.3 Model used to verify clipping and quantization distortion after signal splitting

The third model is used to verify the clipping and quantization distortion components after signal splitting. There is one change with respect to the model in Figure A.1a; the digital-to-analog conversion section is replaced by the one in Figure A.3.

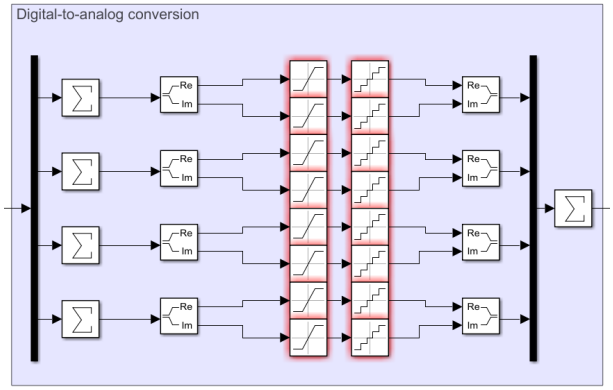


Figure A.3: Digital-to-analog conversion section in Model 3 that replaces the DAC section in Model 1, see Figure A.1a.

If the number of subcarriers in the model is set to those in Table A.1, then the hardware in Figure A.3 can be used to verify the clipping distortion curve of converting $64/4 = 16$ subcarrier per DAC. The selector block divides the incoming array into 4 equal partitions that are then summed and clipped and quantized. Then, the four outputs are summed again and fed to the receiver.

Using an MER-measurement, the distortion curves can then be determined. To measure the curves for the other number of subcarriers per DAC, the global number of subcarriers can be changed.

A.1.4 Model used to analyze the spread in current-sources in the DAC

The fourth Simulink model consists out of two parts. The first part is identical to Model 1, except for 1 change. The DAC area in Model 1 is replaced by a new quantization process as depicted in Figure A.4.

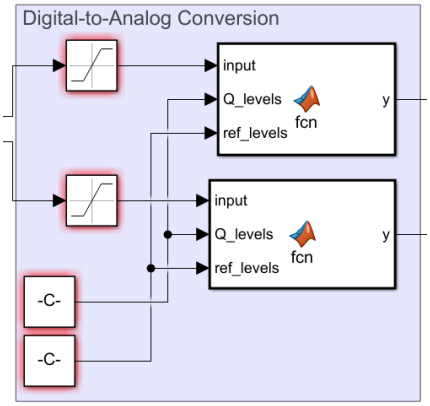


Figure A.4: A new quantization process has been designed to evaluate the spread in current levels in a typical current-steering DAC.

The quantization array used is slightly more complex than a simple quantization block. The following script is run prior to running the Simulink model to generate the DAC transfer curve.

```
A_c = 4; % normalized clipping level
upper_limit_clipping = A_c*sqrt(P_average_constellation);
lower_limit_clipping = -A_c*sqrt(P_average_constellation);

DR = upper_limit_clipping - lower_limit_clipping; % dynamic range of ...
% the signal to be quantized
N = 8; % Quantization resolution

unitcell_baseline = DR/2^(N); % LSB step without spread
unitcell_array = zeros(round(2^N),n_subcarriers) % prepare quantization array
quantization_array = zeros(round(2^N),n_subcarriers); % prepare quantization array
```

```
ideal_quantization = zeros(round(2^N),n_subcarriers); % prepare quantization array
a = zeros(round(2^(N)),1); % offset correction factor
INL_array = zeros(round(2^N),n_subcarriers); % Array containing INL values
DNL_array = zeros(round(2^N),n_subcarriers); % Array containing DNL values

for k = 1:n_subcarriers
length = size(unitcell_array(:,1));
for i = 1:length_
unitcell_array(i,k) = DR/2^(N)+normrnd(0,DR/2^(N)*2/sqrt(2^N)); % Take standard LSB...
% and add random sample from normal distribution to model spread
end

for i = 1:length
    quantization_array(i,k) = sum(unitcell_array(1:i,k)); % add ind. values to form curve
end

for i = 1:length
    ideal_quantization(i,k) = i*unitcell_baseline_QSI; % make ideal curve
end

a(k)  = (quantization_array(end,k)/ideal_quantization(end,k)); % compute offset corr.
INL_array(:,k) = (quantization_array(:,k)-ideal_quantization(:,k)*a(k))/unitcell_baseline;

for i = 1:length-1
    DNL_array(i,k) = (quantization_array(i+1,k)-quantization_array(i,k))/unitcell_baseline-1;
end

quantization_array(:,k) = quantization_array(:,k) - DR/(2^(N))*(length(:,1)/2+1);
ideal_quantization(:,k) = ideal_quantization(:,k) - DR/(2^(N))*(length(:,1)/2+1);
% move curves to mirror around zero
end
```

The arrays *quantization_array* and *ideal_quantization* are fed to the function blocks visible in Figure A.1a. There, they are processed as follows:

```
function y  = fcn(input,Q_levels,ref_levels)
    idx = ones(n_subcarriers,1); % prepare index array
    y = ones(n_subcarriers,1); % prepare output array
 for k = 1:n_subcarriers % for each subcarrier, check in ideal array which digital
    % code is required and output the corresponding sample from the array with spread.
    idx(k) = round((interp1(ref_levels(:,k),ref_levels(:,k),input(k),'nearest','extrap')...
    +DR/2)/unitcell_baseline_QSI);
    y(k) = Q_levels(idx(k),k);
 end
```

the effect of spread on the MER can be simulated after adding spread to the quantization levels. To analyze the impact of spread in the quantization levels on 'traditional' OFDM systems, the subcarrier mixing section in Model 1 is replaced by the iFFT section of Model 2, see Figure A.2. In contrary to Model 2, the DAC section of Model 1 is preserved.

For both models holds that the impact of quantization on the MER can be investigated by varying the parameters in the script used to generate the DAC transfer curve.

A.1.5 LTSpice model for verification of current leakage

To verify the limited output impedance of the cascode current sources, a simple model has been designed and implemented using LTSpice, refer to Figure A.5.

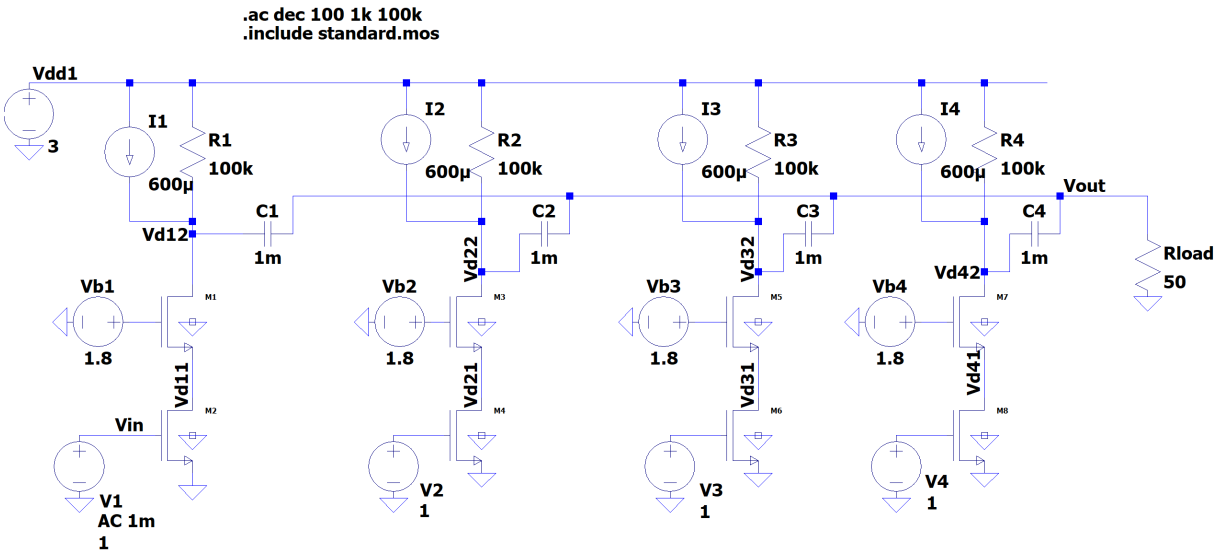


Figure A.5: LTSpice model used to verify current source leakage due to finite DAC output impedance.

The MOS-devices make use of a simple quadratic model identical to the one discussed in [38]. Targeted MOS-conductance is $4\ mS$. Using the quadratic MOS-equation, the desired bias current and gate voltages can then be calculated:

$$g_m = \sqrt{2 \cdot \mu_n C_{ox} \cdot \frac{W}{L} \cdot I_D} = \mu_n C_{ox} \cdot \frac{W}{L} \cdot (V_{GS} - V_{TH}) \tag{A.1}$$

Using $W/L = 100$, $V_{TH} = 700\ mV$ and $\mu_n C_{ox} = 1.31e-4\ [F \cdot m^2/(V \cdot s)]$, the bias circuits as used in Figure A.5 can be found. The capacitors are for bias decoupling to avoid the stages interfering with each other.

By increasing and decreasing the number of 'DACs' and plotting the transfer function, the total leakage can be estimated.

A.2 Simulink models used in Chapter 4

The following models will be covered in this Section:

1. Simulink baseband transmitter and receiver combination for signal splitting along the time-domain
2. Simulink model used for verification of PAPR reduction
3. Simulink model used for evaluation of transient effects

A.2.1 Baseband transmitter/receiver combination

As the same global system parameters hold for the time-domain signal splitting models as in the case of frequency-domain splitting, the parameters defined in Table A.1 also hold for these sections.

The baseband transmitter model has been depicted in Figure A.6. The corresponding receiver is identical to the one depicted in Figure A.1b.

The contents of the iFFT block is identical to what was presented in Figure A.2. Note that the only box checked in the iFFT block is 'Inherit FFT length from input dimensions'. The Real DAC and Imag DAC contain saturation blocks to implement clipping, but no quantization process. The 'track selection' blocks are if-action enabled subsystems that pass or hold their input based on the if-statements as annotated on the schematic in Figure A.6. The 'PA' blocks are empty and serve to designate the PA's location in a real system. Just after the B and C signals, a switch has been implemented to transmit the correct PA

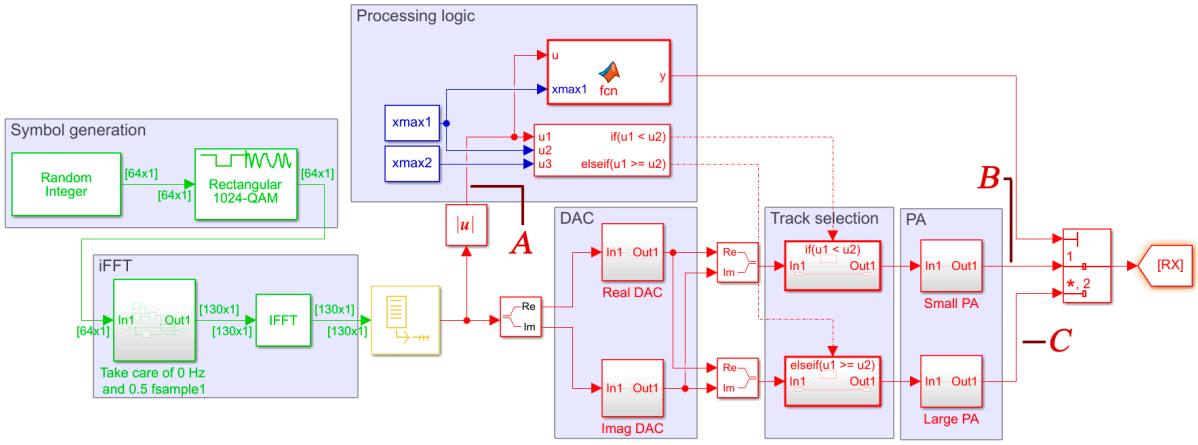


Figure A.6: Simulink model of the baseband transmitter for time-domain signal splitting.

output.

When it comes to the processing logic, the constants 'xmax1' and 'xmax2' are user-defined and provide the switching thresholds. The MATLAB function block drives the switch, refer to the following code:

```
function y = fcn(u,xmax1)
if u < xmax1
    y = 1;
else
    y = 2;
end
```

Table A.4 contains the parameters used for the transmitter model as included in Figure A.6.

Table A.4: Simulink TX model parameters for the baseband time-domain signal splitting model.

Block name	Block parameter	Simulink value
Symbol generation/Random integer generator	Set size	M_QAM
	Initial seed	0
	Sample time	1/f_spacing/n_subcarriers
	Samples per frame	n_subcarriers
Symbol generation/Rectangular QAM Modulator Baseband	M-ary number	M_QAM
	Constellation ordering	Gray
	Normalization method	Average power
	Average power	P_average_constellation
	Phase offset	0
DAC/Saturation (2x)	Upper limit	A_c*sqrt(P_average_constellation)
	Lower limit	-A_c*sqrt(P_average_constellation)

A.2.2 Model used for verification of PAPR reduction

To perform PAPR measurements on the PAPR of the signals at the input of both PAs, some additional blocks have been added to the baseband transmitter as depicted in Figure A.6, refer to Figure A.7.

For each of the PA signals, both the RMS value and the peak value are determined. To determine the RMS value, a running RMS block has been used. Note that it has been placed in an if statement enabled subsystem to have it take into account only those samples dedicated to that PA. The RMS value is subsequently squared to be able to use the result in the definition for the PAPR as defined in Equation 2.5. The maximum value is calculated using the 'Maximum' block in Simulink. This value too is squared. The ratio between the squared peak value and the squared RMS value gives the PAPR as visualized by the displays in Figure A.7.

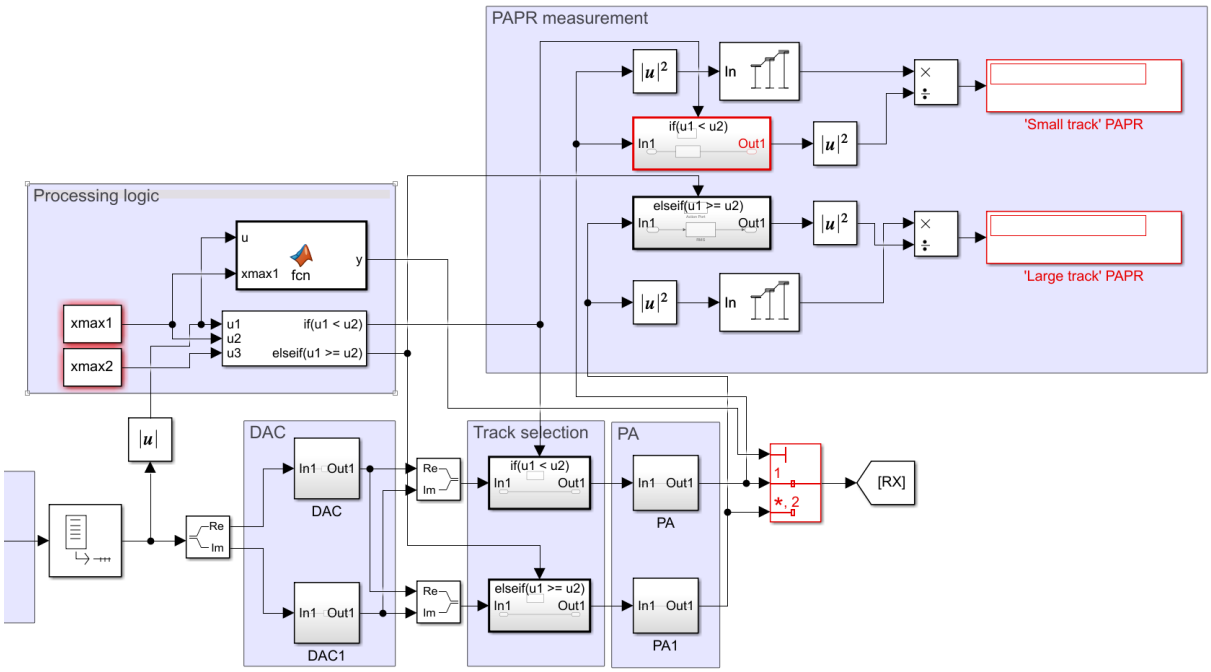


Figure A.7: The Simulink model in Figure A.6 extended with additional blocks to perform a measurement of the PAPR.

In case of three or four hardware tracks, as considered in Chapter 4 too, the other hardware tracks are also given their own RMS and peak value blocks. The if-statement should then be adjusted accordingly to have the samples correctly dedicated to the hardware tracks.

A.2.3 Model used for evaluation of transient effects

The transmitter model used to evaluate the impact of the transient effects on the EVM and ACPR contains four changes with respect to the baseband model as included in Figure A.6:

- Upsampling and FIR interpolation have been implemented.
- Upconversion to the carrier frequency has been implemented.
- The PA transfer functions have been incorporated.
- A measurement of the ACPR has been included.

In addition, the receiver now contains a downconversion block.

Upsampling and FIR interpolation

Upsampling and FIR interpolation have been integrated into the Simulink model. To that end, the area in the Simulink model of Figure A.6 designated with 'iFFT' has been replaced by a new area, refer to Figure A.8.

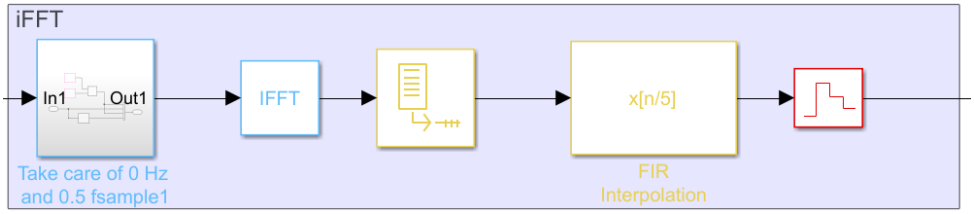


Figure A.8: Upsampling and FIR interpolation have been integrated into the Simulink model of Figure A.6 by including two additional blocks: 'FIR Interpolation' and a 'Zero-Order Hold'.

Inside the 'FIR Interpolation' block, the interpolation factor has been set to 5 and the input processing is based on 'Elements as channels (sample based)'. The Zero-Order Hold (ZOH) block resamples the signal to the same sample rate as used in the upconversion block to avoid sample rate conflicts, refer to the next Section.

Upconversion to the RF carrier frequency

Upconversion to the RF carrier frequency has been implemented through the use of a complex mixer, refer to Figure A.9. This mixer is placed directly behind the 'DAC' area as designated in Figure A.6 in place of the 'Real-Imag to Complex' blocks present there.

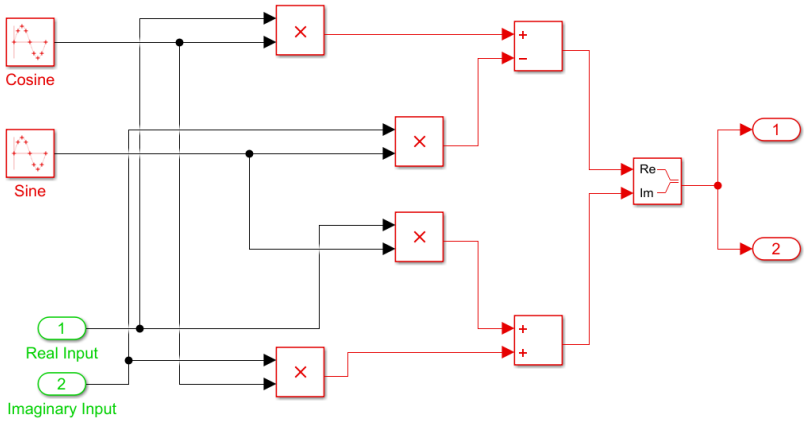


Figure A.9: Complex mier schematic used to upconvert the baseband signal to the desired RF carrier frequency of 26 GHz. The block is placed directly behind the 'DAC' area as designated in Figure A.6 in place of the 'Real-Imag to Complex' blocks present there.

Note how two identical outputs are provided to support the use of two hardware tracks. Table A.5 provides an overview of the parameters used for the cosine and sine blocks.

Table A.5: Complex upconversion mixer parameters used in the Simulink model of Figure A.9.

Block name	Block parameter	Simulink value
Cosine	Time (t)	Use simulation time
	Amplitude	1
	Bias	0
	Samples per period	sine_wave_accuracy
	Number of offset samples	sine_wave_accuracy/4
	Sample time	1/(f_RF*sine_wave_accuracy)
Sine	Time (t)	Use simulation time
	Amplitude	1
	Bias	0
	Samples per period	sine_wave_accuracy
	Number of offset samples	0
	Sample time	1/(f_RF*sine_wave_accuracy)

The sine_wave_accuracy parameter describes the amount of samples used to describe one sine-wave period. More samples results in more accuracy but also longer simulation time. In this model, this parameter was set to 20. The RF frequency is equal to 26 GHz.

Incorporating the PA transfer functions

To incorporate the TFs that were derived in Chapter 4, the empty PA blocks in the model in Figure A.6 have been replaced by the contents as depicted in Figure A.10.

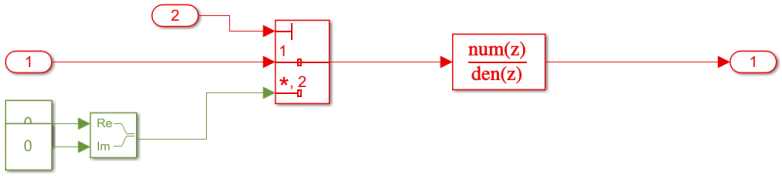


Figure A.10: The in Chapter 4 derived discrete TFs can be incorporated in the Simulink model through the Discrete Transfer Function block. To model the RZ behaviour of the DACs, an additional switch is used to present a signal with value zero when the hardware track is unused.

Note that to model the RZ behaviour of the DACs, an additional switch is used to present a signal with value zero when the hardware track is unused.

ACPR measurement

The ACPR has been measured using the built-in ACPR measurement in the Spectrum Analyzers in Simulink. A spectrum analyzer has been connected to the signal going to the receiver. Table A.6 summarizes the settings for both the spectrum analyzer itself as well as the ACPR measurement. Settings without a notation in Table A.6 have the default Simulink value.

Table A.6: Spectrum analyzer settings for performing an ACPR measurement.

Settings category	Parameter	Simulink value
Spectrum settings	Span (Hz)	$1e9$
	CF (Hz)	$26e9$
	RBW (Hz)	Auto
ACPR measurement	Span (Hz)	2*n_subcarriers*f_spacing
	CF (Hz)	$26e9$
	Number of pairs	1
	Bandwidth (Hz)	2*n_subcarriers*f_spacing
	Filter	None
	Channel Offset (Hz)	$8.3e6$

Downconversion in the receiver

Since the signal leaving the transmitter has been upconverted to the RF carrier frequency in this model, the receiver model as depicted in Figure A.1b has to be extended with downconversion functionality to be able to perform an EVM measurement. Figure A.11 presents the downconversion block used.

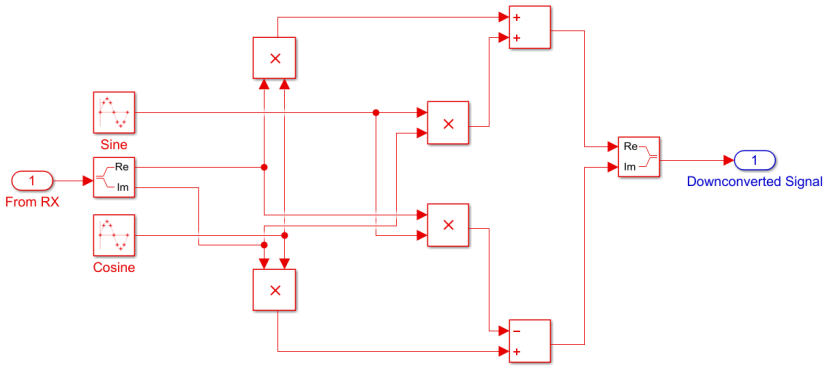


Figure A.11: A complex downconversion mixer has been used in the Simulink model to avoid image problems during downconversion that would complicate performing an EVM measurement. The sine and cosine properties have been included in Table A.5.

The sine and cosine properties used are identical to those in the upconversion mixer and have already been included in Table A.5.

B LINEARITY CONSIDERATIONS IN MIXERS

This Appendix provides an overview of linearity considerations related to the design of mixers to readers unfamiliar with the topic.

Linearity considerations in mixers

To be able to optimally balance the mixer's linearity to problems associated with high frequencies, physical mechanisms impacting mixer linearity have to be considered. In the end, the spectral pollution caused by the hardware has to be within spectral limits and the data transmitted should be of sufficient quality. Linearity considerations can roughly be categorized into three groups:

- In-band distortion: Nonlinear signal transfer in the transmitter will degrade data quality. In-band distortion is summarized in the EVM figure.
- Adjacent-band distortion: Intermodulation (IM) products will pollute nearby channels, the requirements for relatively close bands are strict and should be met. Out-of-band distortion is characterized using the ACPR.
- Spectral pollution in the rest of the spectrum: Harmonics generated by the transmitter are generally located further away from the used transmission band, but should be within spectral limits on pollution. Violation of harmonic content of the transmitter requirements is measured through the spectral mask.

An extensive discussion on the physical origin and impact of each of these non-linearities is beyond the scope of this thesis, the reader is referred to literature for more information [5]. Table B.1 provides a summarizing overview of issues to be encountered.

Table B.1: Overview of linearity considerations in mixers.

Linearity category	**Non-linearity phenomenon**	**Impact on mixers**
In-band	AM/AM conversion	Gain compression in mixers leads to compression of the outer QAM symbols.
	AM/PM conversion	As a result of non-linear (capacitive) effects, a change in Amplitude Modulation (AM) will also yield an undesired change in Phase Modulation (PM).
	Intermodulation	Nonlinear signal transfer in mixers leads to unwanted modulation of multiple closely spaced tones with each other. Especially a problem in OFDM, where many subcarriers are closely spaced and IM products fall on top of other frequency bins.
	Timing skew	In driving the mixer's LO port, if timing mismatch exists between parallel mixers, the offset in LO signals will lead to increased signal corruption.
	Phase noise	The LO signal itself will experience small variations in phase that corrupt the signal as well.
	Limited isolation	Leakage of signals between the various mixer ports leads to signal corruption, such as LO pulling.
	I/Q-imbalance	Imbalance in the in-phase and quadrature components that should ideally have 90 degree phase offset will lead to increased signal degradation.
Adjacent band	Smearing of the baseband signal	Sudden transitions in the baseband signal lead to spectral smearing.
	Intermodulation	Intermodulation products will also fall on top of bands directly adjacent to the desired band, corrupting its contents.
Harmonic content	Mixer spurious tones	Mixing might generate harmonic components in other (further away) bands that should be limited to satisfy the spectral mask.

Because the output of the DACs in Chapter 3 is a current and the magnitude of these currents is relatively large, the optimum mixer topology is a current-driven passive mixer, as it offers superior linearity while directly processing currents. Figure B.1 provides a schematic of a typical fully differential current-driven passive mixer.

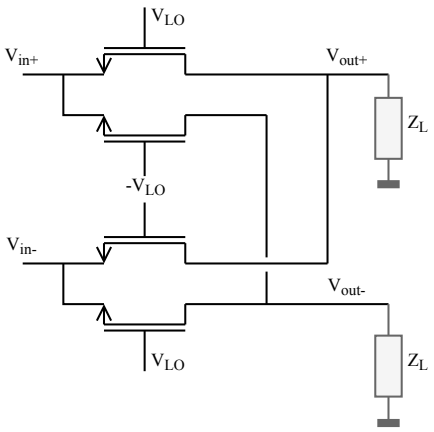


Figure B.1: Schematic of a fully differential current-driven passive mixer.

C DERIVATION OF TRANSFER FUNCTIONS

This Appendix contains the detailed derivation of two transfer functions as used in the main chapters of this thesis.

C.1 TF for evaluating impact of limited DAC output impedance

Figure C.1 provides a SSEC of the circuit under investigation.

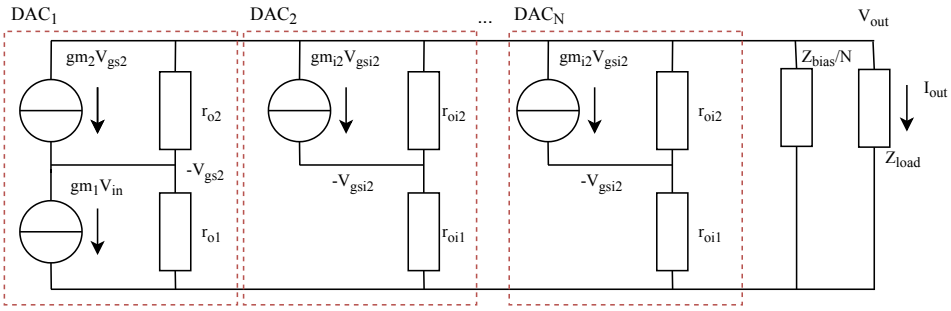


Figure C.1: Circuit schematics for the analysis of current leakage when driving a single load with multiple DACs.

The goal is to find $\frac{i_{out}}{v_{in}}$. As a first step, the current balance in the v_{out}-node will be considered:

$$i_{out} = v_{out} \cdot \frac{N}{Z_{bias}} + \sum_{i=2}^{N} i_{roi1} + i_{ro1} + gm_1 v_{in} \tag{C.1}$$

Through the load resistor, an expression for v_{out} can be substituted:

$$v_{out} = i_{out} \cdot Z_L \tag{C.2}$$

Next, consider an expression for i_{roi1}. Through current balance in the $-v_{gs_{i2}}$ node, an expression for this variable can be determined:

$$i_{roi1} = gm_{i2} v_{gs_{i2}} + \frac{(v_{out} + v_{gs_{i2}})}{r_{oi2}} \tag{C.3}$$

After substitution of $v_{gs_{i2}} = -i_{roi1} \cdot r_{oi1}$ and rewriting, one finds the following:

$$i_{roi1} = \frac{v_{out}}{r_{oi2}} \left(1 + \frac{gm_{i2}}{r_{oi1}} + \frac{1}{r_{oi1} r_{oi2}}\right)^{-1} \tag{C.4}$$

Next, an expression for i_{ro1} will be determined:

$$i_{ro1} = gm_2 v_{gs_2} + i_{ro2} - gm_1 v_{in} \tag{C.5}$$

Through substitution of $v_{gs_2} = -i_{ro1} r_{o1}$, $i_{ro2} = \frac{(v_{out} + v_{gs_2})}{r_{o2}}$ and rewriting one finds:

$$i_{ro1} = \left(i_{out} \frac{Z_L}{r_{o2}} - gm_1 v_{in}\right) \left(1 + r_{o1} gm_2 + \frac{r_{o1}}{r_{o2}}\right)^{-1} \tag{C.6}$$

$$i_{ro1} = \frac{i_{out} Z_L - gm_1 v_{in} r_{o2}}{r_{o2} + r_{o1} + r_{o1} r_{o2} gm_2} \tag{C.7}$$

$$i_{roi1} = \frac{i_{out} Z_L r_{oi1}}{r_{oi1} r_{oi2} + gm_{i2} r_{oi2} + 1} \tag{C.8}$$

Substitution of these results into the current balance of Equation C.1, isolating i_{out} and dividing by v_{in} yields the final result:

$$\frac{i_{out}}{v_{in}} = \frac{-gm_1}{1 + \sum_{i=2}^{N} \left(\frac{Z_{load}}{r_{oi2}} \left(\frac{1}{1 + r_{oi1} \left(\frac{1}{r_{oi2}} + gm_{i2} \right)} \right) \right) + \frac{Z_{load}}{r_{o2}} \left(\frac{1}{1 + gm_2 r_{o1} + \frac{r_{o1}}{r_{o2}}} \right) + N \cdot \frac{Z_{load}}{Z_{bias}}} \tag{C.9}$$

C.2 TF for evaluating transient effects

Figure C.2 provides a SSEC of the circuit under investigation.

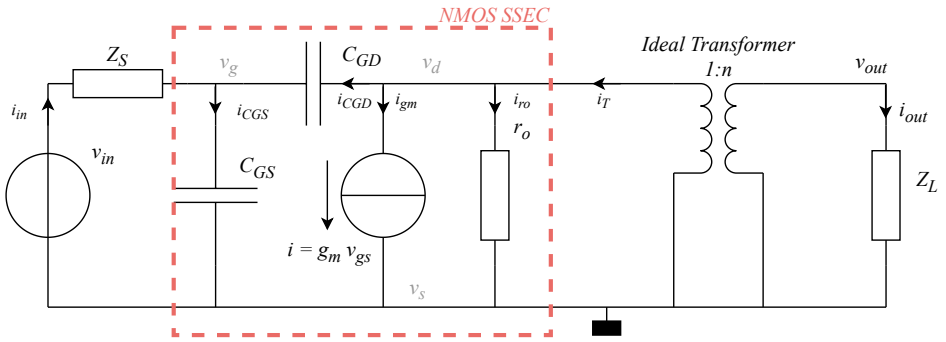


Figure C.2: Small signal equivalent circuit of one symmetric half of the class B amplifier in the configuration as included in Figure 4.15. Current direction definitions have been included.

The goal is to derive the TF from v_{in} to i_{out}. The first step will be to determine the current balance in the drain node:

$$i_T = j\omega C_{GD} (v_D - v_G) + g_m v_G + \frac{v_D}{r_o} \tag{C.10}$$

An expression for v_G can be derived by considering the relation for the capacitor:

$$v_G = \frac{i_{C_{GS}}}{j\omega C_{GS}} = \frac{j\omega C_{GD} (v_D - v_G) + i_{v_{in}}}{j\omega C_{GS}} \tag{C.11}$$

Using $i_{v_{in}} = \frac{v_{in} - v_G}{Z_S}$, one finds:

$$v_G = \frac{1}{j\omega C_{GS}} \left(j\omega C_{GD} (v_D - v_G) + \frac{v_{in} - v_G}{Z_S} \right) \tag{C.12}$$

Rewriting to isolate v_G yields:

$$v_G = \frac{j\omega C_{GD} v_D Z_S + v_{in}}{j\omega C_{GS} Z_S + j\omega C_{GD} Z_S + 1} \tag{C.13}$$

Now, the expression for i_{out} can be found through the transformer:

$$i_{out} = -\frac{i_T}{n} = -\frac{1}{n} \left(j\omega C_{GD} (v_D - v_G) + g_m v_G + \frac{v_D}{r_o} \right) \tag{C.14}$$

An expression for v_D can also be found through the transformer relation:

$$v_D = \frac{i_{out} Z_L}{n} \tag{C.15}$$

Substitution of Equations C.13 and C.15 into Equation C.14 and rewriting yields:

$$\frac{i_{out}}{v_{in}} = \frac{-g_m + j\omega C_{GD}}{n\left(j\omega C_{GD} Z_S + j\omega C_{GS} Z_S + 1\right)\left(\frac{Z_L}{n^2 r_o} + \frac{j\omega C_{GD} Z_L}{n^2} + \frac{j\omega C_{GD} Z_L Z_S (g_m - j\omega C_{GD})}{n^2 (j\omega C_{GD} Z_S + j\omega C_{GS} Z_S + 1)} + 1\right)} \tag{C.16}$$

D ADDITIONAL TRANSFER FUNCTION ANALYSIS RESULTS FOR TIME-DOMAIN SIGNAL SPLITTING

In support of the main analysis into PA handover effects, additional research has been conducted into both the step response of the TFs as analyzed in Chapter 4 and the impact of the parameters on the TFs.

Analyzing the step response
In addition to the frequency-dependent transfer as plotted in Figure 4.20, evaluating the transfer function's step response will provide more insight into the behaviour of the TF as described in Equation 4.5. Plotting the step response will allow determining its settling time after a step in its input value, e.g. when pulling the gate to ground to disable the NMOS devices or restoring intended input voltage to enable the PA. The differences in step response will lead to distortion in the combined output as was already plotted in Figure 4.18. Using the parameters in Table 4.4, the step responses of the TF in Equation 4.5 have been plotted for the large and for the small PA, refer to Figure D.1.

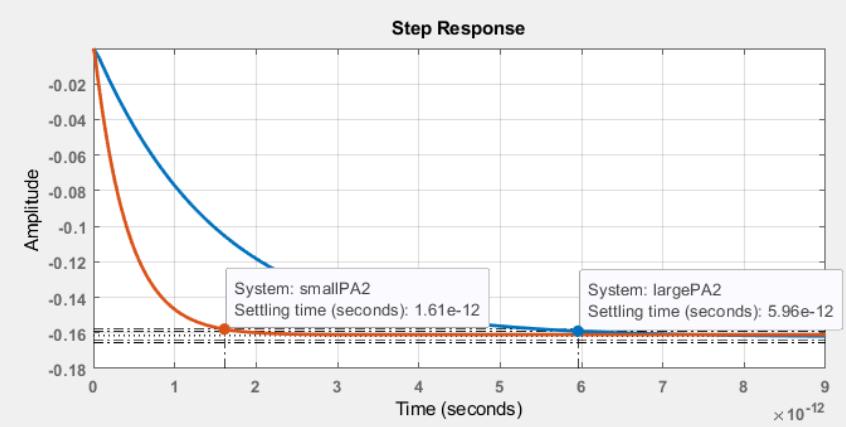


Figure D.1: Step response of the TF in Equation 4.5 using the parameters as included in Table 4.4 for the large and small PA. The input has been stepped from 0 to 1 V. Settling times for both PAs have been indicated.

Figure D.1 shows that the small PA settles a rough factor 4 faster than the large PA. As a result, if the small PA would be enabled while the large PA has to be disabled, the smaller PA would settle faster, leading to an increase in the combined output current, because the large PA is slower in turning off. Similar transient effects occur when disabling the small PA and enabling the large one.

These transient effects will introduce distortion and deteriorate both the EVM and ACPR. Before a Simulink model will be designed to provide an estimate of the impact of the PA handover on the EVM and ACPR, first the impact of the TF parameters as included in Table 4.4 will be examined more closely.

Impact of parameters on the transfer function
Since selecting fully representative values for the parasitic capacitances and output resistances of the NMOS devices would require implementation details, an educated guess for both values was used. The impact of these values on the transfer function will now be evaluated.

First, the impact of the size of the parasitic capacitances will be evaluated. Both the C_{GD} and C_{GS} ca-

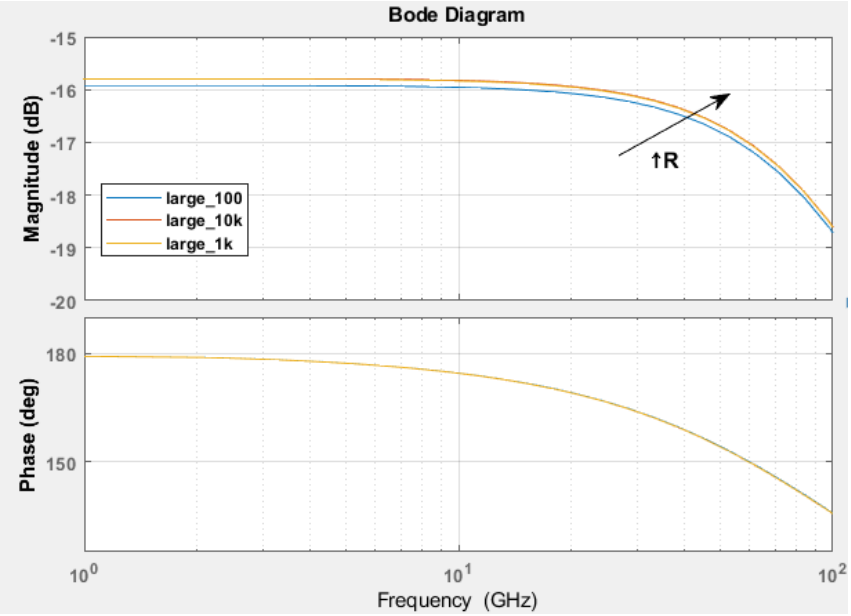


Figure D.3: Bode plots of the TF in Equation 4.5 for various values of the parasitic NMOS output resistance. Values for each curve have been indicated in the legend. The other TF parameters are equal to those for the large PA as included in Table 4.4.

pacitances will be given two additional values: $1.00\ fF$ and $100\ fF$, Figure D.2 provides the resulting bode plots in addition to the one already plotted in Figure 4.20. The other TF parameters are equal to those for the large PA as included in Table 4.4.

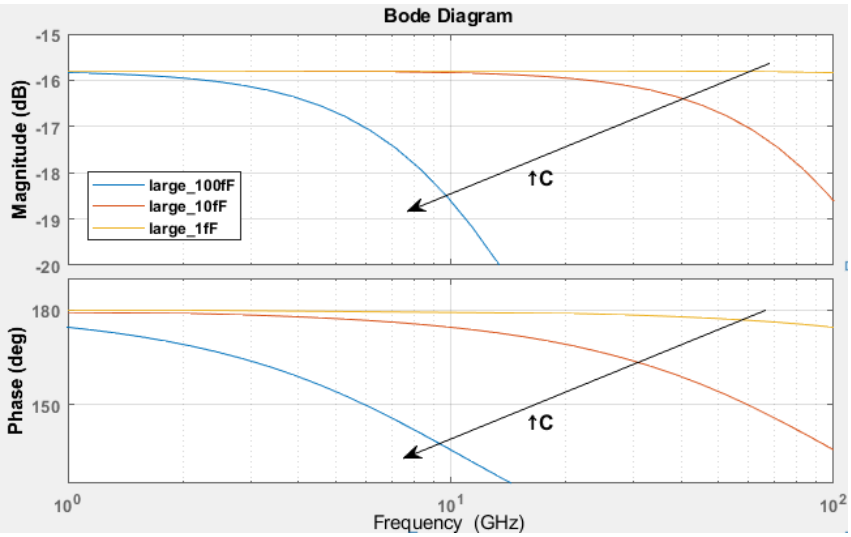


Figure D.2: Bode plots of the TF in Equation 4.5 for various values of the parasitic capacitances. Values for each curve have been indicated in the legend. The other TF parameters are equal to those for the large PA as included in Table 4.4.

A similar evaluation of the impact of a parameter on the TF in Equation 4.5 will now be conducted for the output resistance of the NMOS devices. It will be scaled in a similar manner as the parasitic Cs; r_o will be set to $100\ \Omega$, $1\ k\Omega$ and $10\ k\Omega$. Figure D.3 shows the resulting bode plots. Again, the other TF parameters are equal to those for the large PA as included in Table 4.4.

Combining the results in Figures D.2 and D.3 allows determining what is limiting the TF's magnitude at higher frequencies. Considering that the magnitude fall-off for a larger NMOS output resistance is much

smaller than that for a smaller parasitic capacitance and that there is no noticeable change in phase transfer for varying output resistance indicates that the band-limiting behaviour originates from the combination of the input source resistance with the gate-source parasitic capacitance.

The insensitivity of the phase transfer to the variation in NMOS output resistance is a strong indicator that the limited loss in the transfer's magnitude in Figure D.3 is due to the fact that a small part of the current flows into the output resistance instead of to the load. However, since the matching network lowers the load impedance and the output resistance of the NMOS device is still relatively high, this leakage is small (e.g. only - 0.2 dB at 26 GHz using $r_o = 100\ \Omega$ with respect to $r_o = 10\ k\Omega$).

To verify the impact of the source resistance of the input voltage source on the transfer function, it has also been swept, refer to Figure D.4.

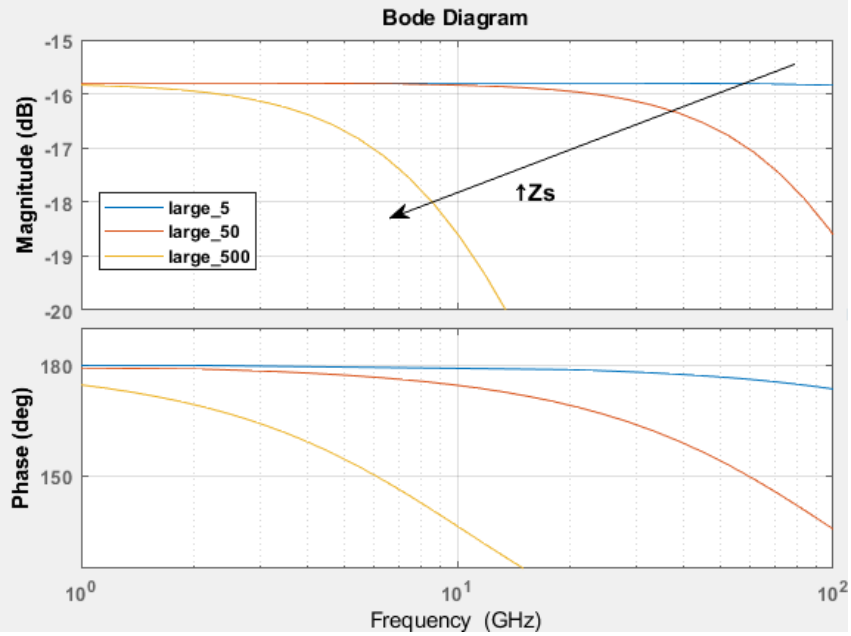


Figure D.4: Bode plots of the TF in Equation 4.5 for various values of the source resistance Z_S. Values for each curve have been indicated in the legend. The other TF parameters are equal to those for the large PA as included in Table 4.4.

Both the magnitude and phase transfer as plotted in Figure D.4 show strong resemblance to those of Figure D.2, confirming that the source resistance combines with the parasitic NMOS capacitances to form a low-pass network at the NMOS gate. Ideally, the source resistance would be adapted for the small PA to counteract the change in parasitic capacitance due to scaling of the NMOS device, complicating the design of the predriver network.