TECHNISCHE UNIVERSITÄT WIEN
Vienna | Austria

PhD Thesis

On Memory Induced, Nonlinear Behavior of All-Digital Radio Frequency Transmitters

unter der Leitung von
em. O. Univ.-Prof. Dipl.-Ing. Dr. Gottfried Magerl
Ass. Prof. Dipl.-Ing. Dr. Holger Arthaber
Institute of Electrodymamics, Microwave and Circuit Engineering

eingereicht an der Technischen Universität Wien,
Fakultät für Elektrotechnik und Informationstechnik

von

DI Norbert LEDER
Matr.Nr. 0525421
Bahnstraße 45
2285 Leopoldsdorf i.M.
geb. 26.3.1985

Wien, im August 2017

Abstract

This work discusses the design and modeling of all-digital radio frequency transmitters (DRFTx). Transmitters based on nonlinear radio frequency (RF) modulators can be energy efficient and universal with respect to center frequencies and modulation formats. This ability is highly desirable for the steadily increasing pace and multitude at which RF-modulation standards are developed and deployed.

However, despite its potential, the DRFTx hasn't been adopted for today's mainstream applications, mainly for three reasons. Firstly, the RF-circuitry requires low loss filters with high quality factors which are inherently bulky and expensive. Secondly, the modulator introduces undesired signal components that need to be sustained throughout the signal chain in order to preserve certain properties of that signal. However, this renders many design paradigms for high efficiency power amplifiers (PA) useless. Dealing with the many issues of a completely novel design methodology is a hurdle in itself. Thirdly, due to its entanglement between the software based modulator and the attached RF circuitry the modulator and PA cannot be developed separately. Both of these entities are highly nonlinear and show memory affected behavior. In order to get the best overall system performance a co-design approach is required, which again is not established yet.

This work demonstrates some of these dependencies and shows that a system approach is required in order to gain an optimized DRFTx design. An important stepping stone is the availability of a circuit-centric model. This allows linking circuit measures like energy efficiency, effective memory depth or nonlinear distortion to properties of the excitation signal generated by the modulator. By using a model that captures these dependencies, a joint optimization between modulator settings and circuit variants can be conducted. Ultimately, this allows to balance nonlinear distortion, circuit complexity, and power efficiency in order to gain an overall viable system.

There are different model types that can serve for this purpose which are presented as an overview in this work. However, the auto regressive exogenous (ARX) and a type of look up table model (LUT) have been investigated and compared in detail. The latter one has been found to be the more suitable option and it was possible to improve it by introducing a hierarchical data structuring technique.

All the results presented in this work are based upon or verified by measurements of a purpose built DRFTx setup. It is based on a laboratory test bench which was built to produce a 20 MHz LTE signal at a center frequency of 2.5 GHz utilizing a 10 W GaN HEMT.

Kurzfassung

Diese Arbeit beschäftigt sich mit dem Entwurf und der Modellierung von voll-digitalen Hochfrequenz-Sendern (DRFTx). Der Einsatz nichtlinearer Hochfrequenz-Modulatoren ermöglicht es Sender zu realisieren, die gleichsam hoch effizient und software-konfigurierbar sind. Damit würde man den lang ersehnten Brückenschlag zwischen der rasanten Entwicklung und der im Vergleich dazu schleppenden Verbreitung neuer Funkstandards ermöglichen.

Obwohl der DRFTx theoretisch großartiges Potential verspricht, wird er bis dato nicht großflächig eingesetzt. Das lässt sich vor allem durch gewisse Eigenheiten des Konzepts und mehrere inhärente Eigenschaften erklären. Erstens müssen verlustarme Filter mit hohen Gütefaktoren eingesetzt werden, um die gewünschte Funktionalität zu erreichen. Diese sind aber schwer zu integrieren, groß und teuer. Zweitens erzeugt der Modulator aufgrund seiner Funktionsweise eine Vielzahl an Signalkomponenten, die nicht dem eigentlichen Ausgangssignal zuzurechnen sind. Diese müssen, um die Funktionsweise des Gesamtsystems zu ermöglichen, bis zum Leistungsverstärker (PA) durch die Signalkette aufrechterhalten werden. Das steht allerdings im Widerspruch zu den meisten Entwurfsparadigmen für hocheffiziente Leistungsverstärker. Um solche Sender zu realisieren, wäre daher ein komplett neues Entwicklungskonzept notwendig, was eine große Hürde darstellt. Drittens, die Abhängigkeiten zwischen dem Modulator, der ausschließlich aus Software besteht, und jenen der Schaltungstechnik verhindert eine separate Entwicklung dieser Einheiten. Um das Gesamtsystem zu optimieren, ist hier ein übergreifender Entwurf notwendig, der momentan sehr unüblich ist.

Diese Arbeit zeigt, dass für die Optimierung des Gesamtsystems die Abhängigkeiten zwischen Modulator und Schaltung berücksichtigt werden müssen. Ein Schaltungsmodell welches das leistet, ist ein erster und wichtiger Schritt dazu. Das Modell muss dazu auf Basis der digitalen Ausgangssignale des Modulators das Verhalten der Schaltung prädizieren können. Das erlaubt schaltungsspezifische Kennwerte, wie Leistungseffizient oder Speichertiefe, mit den Eigenschaften und Einstellungen des Modulators zu verknüpfen. Nur wenn das Modell diese Abhängigkeiten korrekt abbildet, ist es möglich, das Gesamtsystem zu optimieren. Als übergeordnetes Ziel kann dadurch eine Realisierung gefunden werden, die eine optimale Balance zwischen Schaltungskomplexität, Leistungseffizienz und nichtlinearer Verzerrung aufweist.

Es gibt eine Fülle von Modelltypen, die für diese Anwendung in Frage kommen. Diese Arbeit handelt die wichtigsten überblicksmäßig ab und vergleicht zwei davon detailliert. Dabei handelt es sich um die Klasse der äuto regressive exogenous"(ARX)

Modelle und eine Tabellen basierterte Lösung (LUT). Der Tabellen Ansatz hat sich im Rahmen der Analyse als die handhabbarere herausgestellt. Außerdem konnte diese durch den Einsatz einer hierarchischen Daten-Strukturierung für diese Anwendung optimiert werden.

Alle Ergebnisse, die in dieser Arbeit präsentiert werden, basieren auf Messungen an einem DRFTx Test-Aufbau. Dieser ist Teil eines Laboraufbaus zur Untersuchung dieser Sendertopologie, welcher ein 20 MHz LTE Signal bei einer Mittenfrequenz von 2,5 GHz erzeugt. Die Leistungsendstufe verwendet einen weit verbreiteten 10 W GaN HEMT von *CREE*.

Acknowledgements

During the five years while I was working on the overarching project this work is based on, it consumed a lot of my time and attention. I want to thank my family and friends for their continuous support during this time, without it, this work would have not been possible.

As most of this project was carried in the workplace I also want to thank all my colleges for their help, hints and pats on the back. However, some of them stand out as supporters, companions and/or drinking buddies.

Firstly, the microwave group, a team of self-proclaimed nerds with an impressive skill-set that runs on freshly prepared meals and humor, helped me out countless times and I want to thank them for all the time, effort and wit. It was a pleasure to work side by side with Bernhard Pichler on many of the experiments leading up to this thesis and the corresponding papers. He is a true optimist and an excellent scientist. Holger Arthaber managed this project and provided dangling carrots, as well as, the creative deadline management, which is sometimes required for an efficient paper writing process. Thomas Faseth spent many hours, often on short notice, proof-reading paper drafts and giving highly valued feedback. Gottfried Magerl provided many brilliant hints, funny anecdotes, and know-how for almost every situation. Last but not least, I want to thank Margit Maritzen, as one of the good souls behind the scenes. She eased many administrative obstacles and paved the way for a better work environment.

Contents

1 Motivation and Thesis Outline **2**
- 1.1 The DRFTx from a Technological Perspective . . . 3
- 1.2 Circuit Design Considerations . . . 5
- 1.3 Modeling Considerations for DRFTx Systems . . . 7

2 Test Bench for Digital RF Transmitters **10**
- 2.1 Signal Source . . . 12
- 2.2 Booster Amplifier . . . 14
- 2.3 Power Amplification Stage . . . 17
- 2.4 Output Filter . . . 21
- 2.5 Output Coupler . . . 23
- 2.6 Measurement Setup . . . 24

3 Design of DRFTx Circuit Modules **27**
- 3.1 Bias Module for Experimental RF-Amplifiers . . . 27
- 3.2 Load Pull Measurements and Setup . . . 29
- 3.3 Matching Network Design . . . 37
- 3.4 Circuit Synthesis Based on SRFT . . . 41
- 3.5 Measured Matching Networks . . . 48
- 3.6 Design of the Output Coupler Section . . . 50

4 Nonlinear System Characterization and Modeling **55**
- 4.1 Characterizing Nonlinear Memory Affected Distortion . . . 57
- 4.2 Modeling Requirements for DRFTx . . . 61
- 4.3 Model Estimation and Verification . . . 64

5 The ARX Model **68**
- 5.1 Formal Definition of an ARX Model . . . 68
- 5.2 Determining the Required ARX Model Complexity . . . 70
- 5.3 ARX Modeling Results . . . 72

6 LUT Models **76**
6.1 Formal Definition of LUT Model Parameters 77
6.2 Generation of LUT Model Entries 79
6.3 Gauging LUT Models 82
6.4 Hierarchical LUT Models 84
6.5 Verification of LUT Models 88

7 Model Comparison **92**
7.1 Model Estimation and Model Complexity 93
7.2 Simulation and Prediction-Error 96
7.3 Conclusion and Outlook 105

Appendices **114**

CHAPTER 1

MOTIVATION AND THESIS OUTLINE

Transmitters based on nonlinear radio frequency (RF) modulators are subject of ongoing research. Pursuing this technology is largely driven by the need for transmitters that are energy efficient and universal with respect to center frequencies and modulation formats. At least in theory, these requirements can be addressed by employing RF pulse width modulation (PWM) in RF transmitter circuits. However, this class of digital RF transmitters (DRFTx) hasn't yet arrived in today's circuit mainstream and that for good reasons. Despite, or maybe just because of that DRFTx are an interesting system which is influenced by the findings of many research fields, ranging from analog RF-circuit design over nonlinear signal processing to memory affected circuit modeling.

This work is based on a laboratory test bench for a DRFTx which was built to produce a 20 MHz LTE signal at a center frequency of 2.5 GHz utilizing a 10 W GaN HEMT. Since the DRFTx in this setup was designed entirely within the time the work on this thesis was conducted, it allowed for a very experimental approach. In contrast to the following chapters which explain the setup and methods involved in detail, this chapter introduces the basic principles behind the DRFTx and the theoretical aspects that need to be considered in order to build and constructively analyze a DRFTx setup.

1.1 The DRFTx from a Technological Perspective

Designing and building highly efficient, flexible yet broadband RF signal generators is the subject of research spanning decades. Although, all approaches that overcome the stringent efficiency limits dictated by linear power amplifiers[1] do that by employing nonlinear phenomena in their power amplification stage.

Basically there are two ways to do so: Firstly, using amplifiers that behave dominantly linear are driven into saturation. This boosts their efficiency, although, causes some unintentional nonlinear distortion, like compression for example. These effects are combated by countermeasures restoring the intended signal. Typical examples are amplifiers in conjunction with predistortion, Doherthy amplifiers, and many more.

Secondly, it is also possible to intentionally operate the amplifier permanently in saturation resulting in the family of switched mode power amplifiers (SMPA). Of course, these amplifiers behave inherently nonlinear. However, harmonic components that originate from these nonlinear effects, in conjunction with harmonic controlled matching networks, can be utilized to reach even higher efficiencies. For narrow-band excitation, SMPA concepts are reasonably complex to implement and well understood [2]. In general, the nonlinear behavior of the amplifier also generates signal components within its baseband. These components are usually undesirable and their relative power grows with higher relative signal bandwidth of the amplifier [3]. It is important to note, that building a SMPA always requires to include some kind of filter. The terminology varies for different approaches. In some cases a harmonic controlled matching network, which also performs a filter functionality, is sufficient to remove the unwanted signal components. Other designs may require a dedicated reconstruction filter to reach sufficient suppression.

While this work is inclined to the SMPA category, there is one key difference to it. Harmonic controlled amplifiers are driving an SPMA with an analog RF-signal, very similar to the desired output signal with respect to waveform and spectral content. This is equivalent to the assumption that the signal present at the active element of the amplifier is similar to the desired output signal. All additional signal components at the active component do not account for the dominant share of the overall power. For the system presented in this work, things are very different.

The signal exciting the power amplification stage of the system is generated us-

ing a pulse width modulator which generates the baseband signal using pulse width modulation (PWM) and performs the up conversion to the desired center frequency [4]. However, it has a rather low coding efficiency which expresses the ratio if the energy of desired signal components and the overall energy of the signal. For the presented system the coding efficiency is on the order of 1-10 % [5], depending on specific modulator settings, which results in wide-ranging implications regarding circuit design and modeling of the circuit. Ultimately, this is equivalent to envelope elimination and restoration approach with only a single bit encoding of the envelope and a binary drive signal.

The concept of employing a PWM modulator with noise shaping techniques together with switched mode power amplifiers, is well established [6], and found nowadays primarily in base-band applications of moderate bandwidth. The combinations of switched mode amplifier and noise shaping modulator have substantial benefits, namely superb linearity, broad band operation, and high efficiency [7]. However, that comes at the expense of operating the whole circuit, including the power amplifier, at frequencies orders of magnitude higher than the bandwidth of the system and the need for low-loss yet very selective filters to eliminate unwanted modulator noise. While this is not a big issue, when dealing for example with audio signals, expanding this concept to radio frequency (RF) systems in a straight forward manner is not possible. One key idea to get around this problem is to further incorporate the up-converter into the system. Works like [8] [9] [10] [11] have demonstrated concepts that are based on this idea and are more suitable for typical RF modulation schemes. Furthermore, they are realizable with readily available components. However, designing a setup capable of delivering significant output power at center-frequencies in the GHz domain is a challenging task.

Additionally, there is an ongoing trend of incorporating as much as possible into the digital signal processing chain of the RF system which led to the concept of digital-SMPAs [9]. Here, an SMPA is directly excited by the binary output of a PWM modulator operating the power stage permanently in a saturated regime. Determining the overall energy efficiency of such a system is not a trivial task and is discussed in works like [12]. However, the resulting excitation signals are very broadband what makes dealing with nonlinear effects a challenging task. Figure 1.1 provides an overview of a DRFTx system including the spectrum of the signals in question at relevant sections

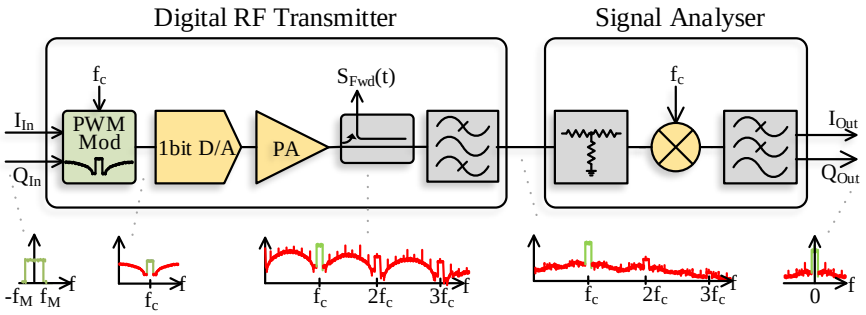


Figure 1.1: Operation principle of a DRFTx.

in the circuit.

The inner workings of the PWM modulator discussed within this thesis were published in [5] [13], and, [14]. Figure 1.1 illustrates the most important interrelations and signals related to the PWM modulator. Basically, it generates its binary output signal reassembling the desired output signal from a digital input signal and a predetermined set of baseband symbols. In a first step, a baseband version of the output signal is formed by the PWM algorithm employed in the PWM modulator. The algorithm is operating at a higher sampling rate than the input sequence which allows to utilize noise shaping. Although the sequence already reassembles the desired output signal, it is still located in the baseband of the modulator. The digital up converter shifts this frequency band to the RF domain by remapping the baseband symbols to RF symbols. These RF symbols need be chosen appropriately according to the different sampling rate of input sequence, modulator sampling frequency, as well as the center frequency. A single bit digital to analog converter (DAC) is required in order to gain a time continuous version of the signal used to drive the following stages of the DRFTx circuit.

1.2 Circuit Design Considerations

Figure 1.2 provides an overview on the circuit modules of a DRFTx and typical signals at important interfaces. A DRFTx system uses a digital PWM modulator to generate the desired RF-signal within its modulation bandwidth. Therefore, the input signal of the analog section is binary and requires a high bit rate in the range of 10 Gbit/s to

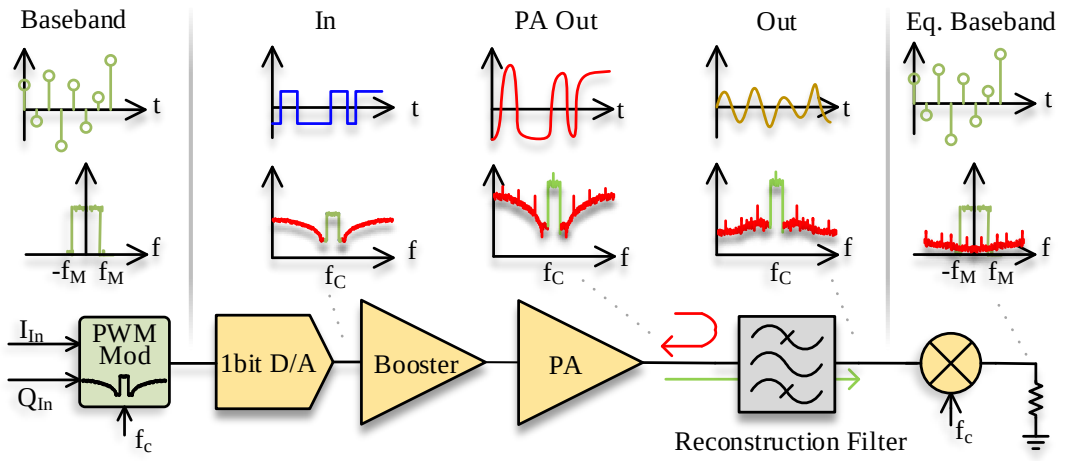


Figure 1.2: Circuit modules and typical signals within of a DRFTx.

be feasible for the task at hand. Practically all digital sources capable of delivering such a data rate use low voltage differential signaling (LVDS). From a circuit design perspective, such a signal is very weak compared to the drive levels required in order to reach saturation within the PA stage. This calls for a booster amplifier in between the LVDS output of the modulator and the PA stage. However, broadband high gain limiting-amplifiers are notoriously nonlinear. Although not all nonlinear effects degrade the overall linearity of the system due to the special nature of the signal, any nonlinear contribution that produces signal components within the modulation band reduces the overall linearity. This is discussed throughout chapter 4, where different methods are discussed that allow to test a given circuit for memory affected nonlinear behavior.

Since the majority of the energy of the signal is located outside the modulation band, the PA's mode of operation is dominantly defined by the side products of the PWM modulator, instead of the intended output signal. Thus, the design of the PA stage in this setup is strongly dependent on the output signal of the modulator. This is at odds with classical PA design approaches and the utilized approach is documented in section 2.3. Furthermore, the PA design strategy is also very closely related to the design of the reconstruction filter that also makes up a significant part of the PA's output matching network. Designing this filter is a balancing act. On one hand, high suppression of modulator noise at the output and the requirement for highly reflective

terminations in harmonic controlled amplifiers is in favor of a very narrow band reconstruction filter that requires high quality factors. On the other hand, by having highly reflective terminations, load-modulation effects at the output of the active device are poorly dampened. If these effects exhibit time constants, that are not negligibly short compared to the sequence length of the PWM-modulator, quasi static design considerations will fail. So to reduce distortion introduced by nonlinear memory effects an absorptive, low quality factor filter would be beneficial. This kind of memory induced, nonlinear behavior was in fact observed in the lab setup, so the corresponding interpretation and analysis of these effects is one of the major parts in this work. Thirdly, the need for a higher order filter that is required to handle multiple times the output power of the system, features a tuneable center frequency, and behaves highly linear is disastrous at best. In fact, herein lies the most problematic design aspect of this concept. While piezo electric tuneable cavity filters [15] have the potential to make this problem approachable, no solution that utilizes current technology will be neither economically feasible, nor will it be a physically compact one. Nevertheless, this problem is very interesting from an academic point of view.

1.3 Modeling Considerations for DRFTx Systems

Finding a balance between the complexity of the modulator, the circuit design, and nonlinear distortion is essential to build an overall efficient DRFTx setup. Doing so in an experimental manner is possible, however, time consuming and requires new hardware iterations for any new circuit to be tested. It was demonstrated [16] that the observed memory induced nonlinear effects can be measured and modeled effectively in time domain, which makes it possible to efficiently perform circuit-modulator co-design.

Modeling complex systems is nearly as much about describing specific dependencies, as it is about splitting a complicated system into more manageable parts that can be analyzed separately. As indicated in Figure 1.2, a DRFTx consists of numerous parts that need to be considered in order to capture the behavior of such a system. The first step towards partitioning this system is to have a separate model for the PWM-modulator. This is straight forward, since the modulator is implemented in software and has clear defined interfaces. It is excited by the desired output signal provided in

the equivalent base-band and requires some parameters that define its resulting center frequency, noise shaping, and its resulting bandwidth. While these parameters may affect the circuit with regards to in-band linearity or overall energy efficiency, the functionality of the modulator is not influenced by any effects in the circuit. Therefore, the modulator is considered a black-box in this work and its inner workings are not investigated, however, it is extensively described in [14].

In contrast to the modulator, the circuit cannot be segmented as extensively as it is done in many more traditional design approaches. A partitioning of the circuit is hindered due to two main differences in the way this circuit operates:

Firstly, traditional design approaches usually consider the amplifier of a TX-system as quasi-stationary nonlinear. In the presented case, however, the load modulation of the active device causes an entanglement between the nonlinear circuitry and the statistics of the input signal. Measurements [16] show that the memory depth of the circuit is in the same range as typical burst lengths of the the modulator, which clearly violates the assumed quasi-stationary behavior. Therefore, methods like passive tuner-based load-pull cannot be applied directly for such circuits since they require quasi-static load conditions for the DUT. Chapter 2 elaborates this issue in greater detail. For the same reasons, assigning a class of operation to a DRFTx is not meaningful. Dependent on the excitation sequence, the active device's mode of operation changes drastically as demonstrated in chapter 4.

Secondly, a DRFTx produces the desired RF waveform by means of a binary modulator. The spectrum of the binary signal resembles the desired spectrum around a chosen center frequency [17], however, the modulator also introduces modulator noise [5]. The output signal of the modulator is composed of many components. Its power spectral density is illustrated in Figure 1.2, which also indicates the frequency-band occupied by the desired RF-signal. The cumulative power of the modulator noise usually supersedes the overall power of the intentional signal. The corresponding power ratio of intended RF signal and the overall signal power is denoted as coding efficiency [12], which is 2-5 % for the described system. Therefore, in order to achieve a certain behavior of the active device, like power efficient operation, the entirety of signals present at the active device need to be known. Even though, around 95 % of these signals do not contribute to the desired output signal.

Out of this set of possibilities a lookup table (LUT) model has been found to be the

most practical choice, because it provides a compact nonlinear model which allows to adaptively capture the memory depth of the circuit. It is based on a LUT, consisting of the output waveforms of the system for all possible input symbol combinations up to a certain memory depth. In order to predict the output of the system, it combines LUT entries according to the input sequence in question. In a general case such an approach is impractical since it would require a huge LUT to capture the multitude of possible input signals. However, this model exploits the fact that any excitation signal, as provided by the PWM-modulator, is composed of a limited set of binary symbols. This greatly simplifies the structure of the LUT and allows to cover all possible input signals with a discrete set of input symbol combinations. Since this model takes advantage from the limited signal space used by the modulator, it is a less general approach than a filter model. However, while this limits the model to a specific set of PWM parameters it does not anticipate a specific baseband encoder/modulation algorithm. The PWM parameters are strongly linked to the target application and technology used. Hence, this limitation doesn't limit the universality of the model in practice.

Chapter 2 elaborates what capabilities are required from a test bench in order to characterize a DRFTx. The following chapter 3 explains the design of the DRFTx, that served as the DUT in the test bench. The modeling issues specific to the DRFTx, as well as its ramifications, are explained in chapter 4. Since there are many possibilities how to describe memory induced, nonlinear behavior chapters 5 and 6 show how to obtain practical relevant models based on auto-regressive- and LUT-model approaches. The results and noteworthy aspects of each of these are discussed in chapter 7 which also concludes this work and provides an outlook on potential expansions of that concept.

CHAPTER 2

TEST BENCH FOR DIGITAL RF TRANSMITTERS

In contrast to Software Defined Radios (SDR), which are the functional siblings of DRFTxs, there are no standardized hardware platforms to conduct research, mostly due to the fact that DRFTx setups need to be built purposely for a given frequency and power range. As explained in chapter 1, building a DRFTx is about building a transmitter that operates power efficiently across different modulation formats and power levels and not about finding a universal transmitter architecture like a SDR. This ambitious goal requires to co-develop RF-modulator algorithms and DRFTx circuitry. A DRFTx is a complex piece of circuitry and its modules are strongly depending on each other. Therefore, it requires some dedicated design methods which utilize data gained from different functional units to optimize the overall functionality which is impossible to illustrate for the individual unit. The focus of this chapter is to illustrate these dependencies and how they can be addressed. This is complemented by chapter 3 which explains the more cohesive design methods and measurements that have been used.

This chapter describes the DRFTx that served as DUT as well as the setup of the test bench itself. The test bench must allow testing different DRFTx algorithms and circuitry against each other. Therefore, modularity and flexibility have been given top priority over available yet rigid solutions. It also explains why nonlinear modeling is a necessary step to develop nonlinear memory affected circuits that operate efficiently

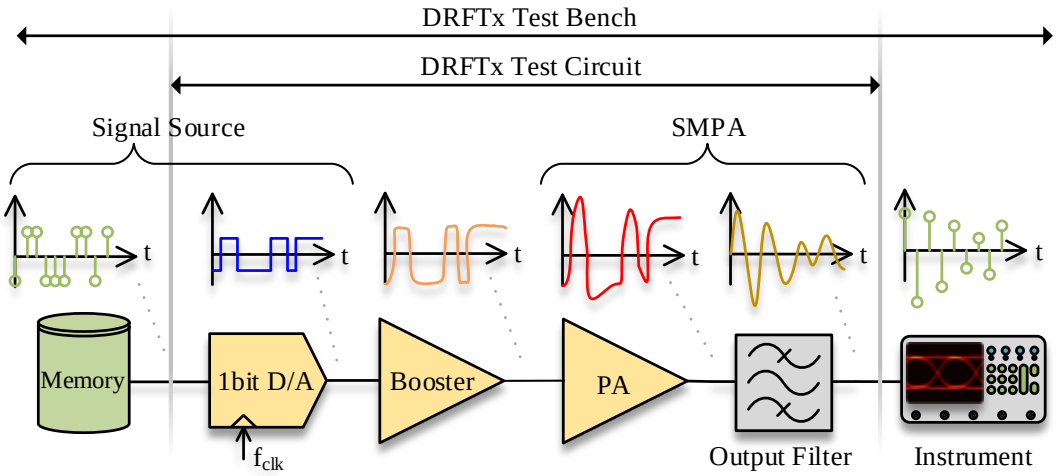


Figure 2.1: Test bench overview

under modulated excitation. While this chapter has a focus on the interaction between the different functional modules, chapter 3 complements this by elaborating the inner workings of these models and the design choices that went into them. Chapter 4 explains how the nonlinear memory affected behavior can be gauged, even though modulators are not available or their performance is insufficient. Modeling the circuit is in many ways superior over just characterizing it. Different modeling approaches suitable for this type of circuits are described and compared in chapters 5 to 7.

The experimental nature of this work is reflected in not having stringent circuit parameters to be met. However, there are some overarching design goals and parameters that are derived from the potential target application, namely an LTE transmitter: The DRFTx was designed to work as an LTE transmitter for a channel bandwidth of $f_{Mod} = 20\,\text{MHz}$ around a center frequency about $f_c = 2.5\,\text{GHz}$. The exact center frequency can be adjusted by means of the employed modulator [14]. The utilized modulator symbol length is $n_s = 4\,\text{bit}$, resulting in a bit rate of $f_{bit} = 10\,\text{Gbit/s}$ for the excitation signal. This in return defines the modulator's fundamental frequency to be $f_1 = {}^{f_{bit}}/_{n_s} = 2.5\,\text{GHz}$. A measured spectrum of such an excitation signal is depicted in Figure 2.2.

In order to provide comparability to other works, the very popular *CGH40010F* HEMT from *Cree* was chosen as the active device in the PA stage rated for about 10 W of saturated output power. Other stages like the booster amplifier or the output

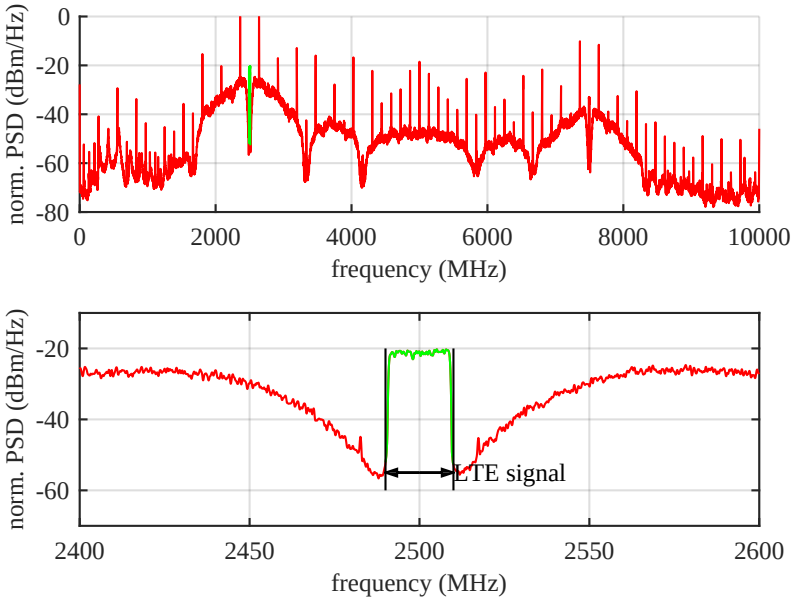


Figure 2.2: Power spectral density of the PWM modulated excitation signal

filter where designed to fit in with the requirements defined by the application as LTE transmitter.

2.1 Signal Source

A binary signal source, capable of operating at 10 Gbit/s with sufficient memory to store a modulated LTE frame is required to excite the attached circuitry. A *BPG 40G* from *Sympuls Aachen* was used throughout this work. However, since the unit does not provide an internal bit clock signal an RF source (*Agilent E8244A*) acts as a configurable bit clock. It is used to drive the bit pattern generator and to synchronize the attached measurement equipment with the bit pattern generator. To generate a specific excitation sequence, the bit-pattern-generator can be loaded with an arbitrary bit sequence which is played back repetitively.

As the bit-pattern generator is designed to test digital communication equipment, it operates on a double data rate principle. This, however, is problematic with respect to the task at hand since it causes asymmetric bit durations. Any offset error in the

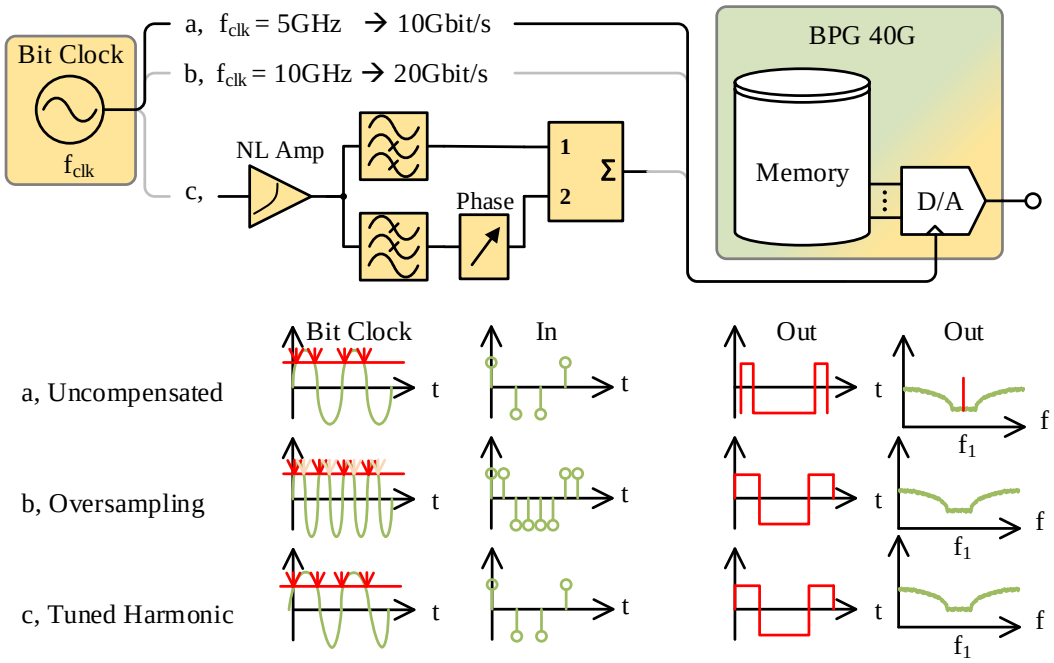


Figure 2.3: The digital signal source and used bit length balancing

bit-pattern generator's clock input or any asymmetry of the circuitry that goes with it affects the bit duration. Therefore, the duration of a bit triggered by the falling edge will be slightly different to one triggered by a rising edge. The distortion due to this effect is most noticeable in the frequency domain since it causes additional spurs. In contrast to a random clock jitter, which causes an evenly raised noise floor in the DAC's output signal, a deterministic variation in bit duration gives rise to undesired spectral lines [18].

Since the bit clock is an integer multiple n_s of the modulator's base frequency f_1, it causes significant distortion around the system's center frequency. Strictly speaking this effect is one example for memory induced, nonlinear distortion since two adjacent bits affect their individual bit length depending on their values. Although, as this effect only shows a memory depth of two RF-bits, it is rather short in comparison to others discussed in this work. Section 4.1 discusses how this effect can be quantified. In any case this effect reduces the dynamic range of the overall system. Therefore, it needs to be compensated. Two methods have been investigated to tackle this issue:

The first approach simply oversamples the input sequence by a factor of two and

increases the clock rate to twice the bit rate. While this method eliminates the effect perfectly without the need for tuning and for any possible clock rate, it effectively halves the the bit-pattern generator's memory. Since some of the excitation sequences require memory on the order of 10^6 bits, this method is not optimal in every sense.

Secondly, a nonlinear amplifier is used to generate a distorted version of the clock signal. A diplexer then separates the fundamental frequency from the predominant first harmonic while a phase shifter allows varying the phase relation of these two components. Doing so, allows the waveform to be modified in order to compensate the observed uneven bit's length. However, it requires the phase shifter to be tuned according to measurements. On top of that, this approach also suffers from a significant drift and temperature dependence. Therefore, the oversampling method was used when ever possible. Figure 2.3 depicts this issue and the two methods that have been used to mitigate this behavior.

2.2 Booster Amplifier

The booster amplifier stage is necessary to amplify the small voltage swing provided by the signal source to levels suitable for driving an RF transistor into saturation. It consists of an integrated limiting amplifier namely the *TGA4954-SL* from *TriQuint*. This device is intended for driving Mach-Zehnder modulators up to about 12.5 GBit/s with +30 dBm output power. Its properties make it suitable as gate driver as well. Furthermore, the TGA4954-SL's output waveform can be continuously tuned by its bias voltages. This basically affects both, the output voltage swing, and the pulse symmetry, which are both critical parameters for controlling the overall nonlinear behavior. Setting the bias for this unit in a consistent manner is a nontrivial task that is discussed in section 3.1.

From a design perspective it is important to find an optimal bias setting with respect to nonlinear distortion. As explained in section 4.1, static nonlinear distortion does not affect the linearity of a binary system. However, memory afflicted nonlinear behavior limits the maximum dynamic range of the system. Therefore, the bias settings of the booster amplifier were optimized to reduce these effects in a setup as depicted in Figure 2.4. Works like [19] and [20] have demonstrated that this is vital for the quality of the overall setup.

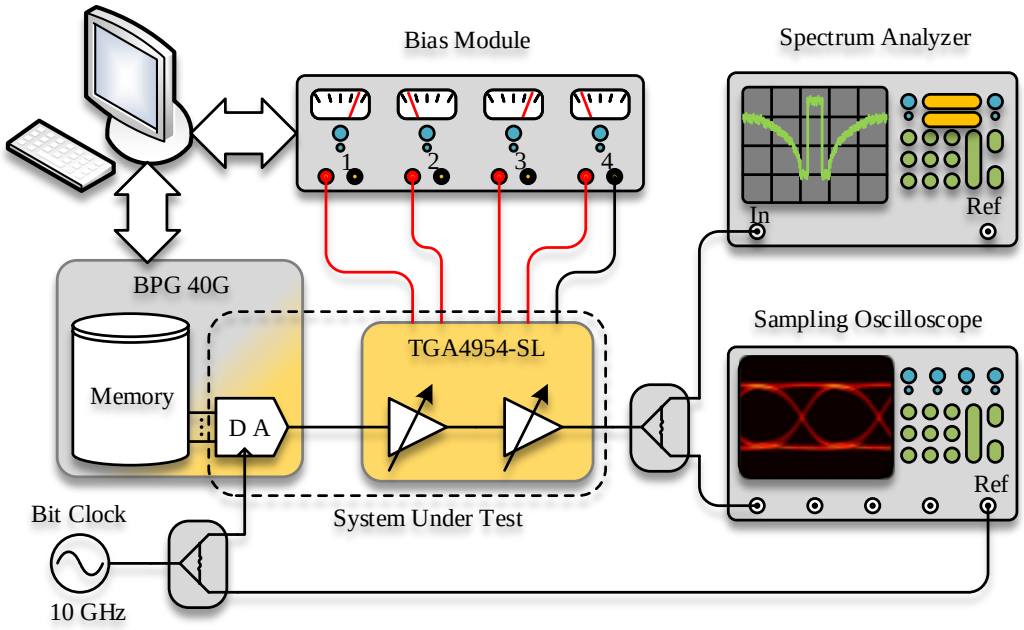


Figure 2.4: Setup used to investigate the booster's nonlinear distortion.

Figure 2.5 illustrates this effect by visualizing the achieved dynamic range depending on the bias setting. The measurement method is described in section 4.1. In the following, only bias points in the favorable region that form the central ridge in this graph have been used to bias this stage.

In the ideal case the resulting dynamic range is solely limited by the employed modulator. Therefore, deriving measurements based on the modulator's output signals can be sub-optimal. This can be avoided by utilizing test methods that use very generic excitation signals that don't require a modulator. For this work a characterization method based on eye-diagrams was investigated (section 4.1) as well as an auto regressive behavioral model (chapter 5).

The eye diagram is a convenient way to visualize the amplifier's variable performance depending on its bias. Figure 2.6, provides an overview. Some settings produce very symmetric eye diagrams, while in other cases they result in less regular patterns. Nonetheless, how this affects the resulting RF-signals' quality is not easy to predict. Although, the effect can be quantified by using the method described in section 4.1.

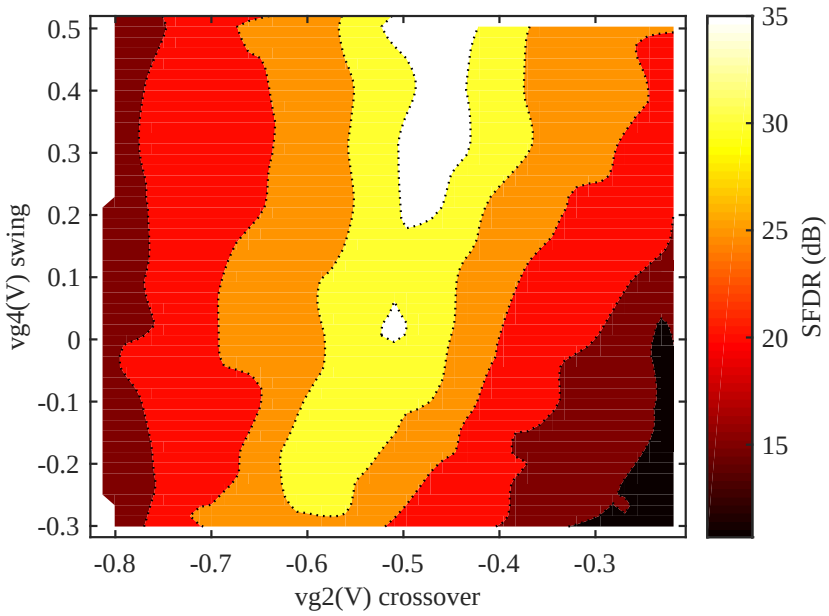


Figure 2.5: Memory induced, nonlinear distortion as function of the booster stage's bias settings measured in frequency domain.

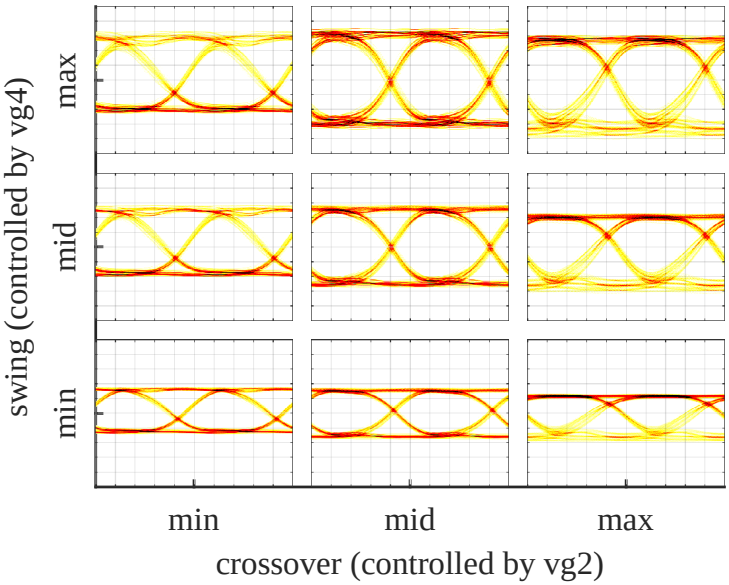


Figure 2.6: Eye diagram of the booster amplifier's output depending in the applied bias settings.

The results are depicted in Figure 2.7 and verify that the beneficial bias settings, as identified using frequency domain measurements, are correct.

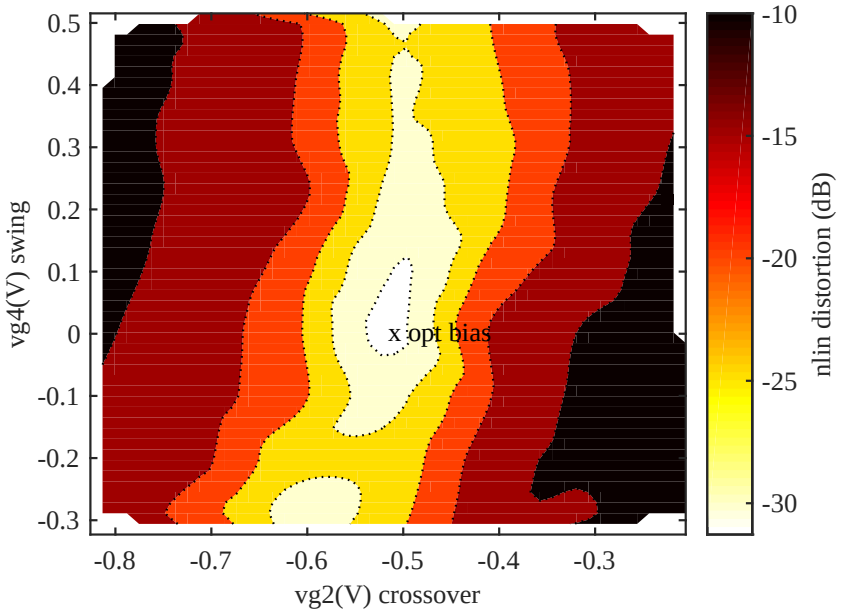


Figure 2.7: Memory induced, nonlinear distortion as function of the booster stage's bias estimated based on its eye diagram.

One drawback of these methods is that they allow characterizing only the memory induced, nonlinear distortion, however, they provide no information about the memory depth of this effect. By finding a model for the circuit it is possible to gain more insight by evaluating the resulting modeling error over the model's complexity. This is explained in greater details in chapter 5.

To provide a more tangible illustration of the hardware making up this section, the booster amplifier and its peripherals can be found in Figure 2.14. The booster amplifier in this pictured is attached to the bias controller using a ribbon cable and are marked with the same blue color frame.

2.3 Power Amplification Stage

Switch-mode power amplifiers (SMPAs) are promising candidates for the power amplification stage in DRFTx systems. Their potential high power efficiency, which is

gained by exploiting the nonlinear nature of the transistor, makes them attractive for this kind of application. By using a matching network and output filter combination that is highly reflective at the signal's harmonics, the current and voltage waveform at the transistor's output can be adjusted. Depending on the configuration, a specific class of operation can be achieved [2]. In its basic variation this only holds true for the design frequency. However, continuous modes have been proposed [21] that allow to gain additional design space and realize highly efficient operation over a wider frequency range. SMPAs in continuous mode have demonstrated operational frequency ranges up to one octave [22]. Although, it is important to note that the efficiency achieved in these works is measured using continuous wave (CW) excitation of varying frequency, not for modulated excitation. If modulated signals are used to drive the circuit, the power efficiency significantly reduces [23] [12]. Therefore, having a SMPA that is delivering good performance over a wide frequency range is not equivalent with achieving high efficiency under broadband modulated excitation.

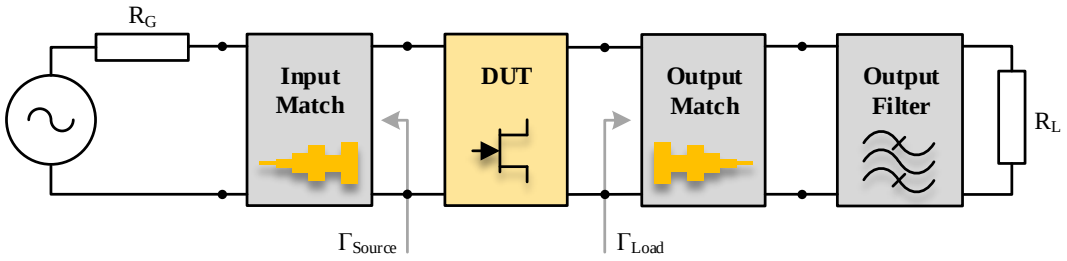


Figure 2.8: Circuit elements of the power amplifier stage

The signal's components provided by the PWM modulator can be separated into two categories. Firstly, there is the intended RF signal located around the desired center frequency, secondly, there is the modulator noise. This is depicted in Figure 2.2 and 2.9 which depict the spectrum of the excitation-signal and at the PA's transistor respectively. The red signal components indicate the modulator noise and the green components correspond to the desired signal. The ratio of the powers of these two components is usually referred to as coding efficiency [13] and is determined by the PWM modulator's parameters. For the utilized modulator and a 20 MHz LTE signal the coding efficiency is typically ranging from 2-5 %.

In both cases, about 90 % of the signal's power is distributed over a bandwidth of

310 MHz around the center frequency, which underlines the wide bandwidth of the system. The desired LTE signal is located in the center and occupies only 20 MHz of bandwidth. The frequency bands denoted in Figure 2.9 indicate which parts of the spectrum are relevant for individual functional units. While the modulator-band contains the desired RF-signal, the PWM-band is the part of the spectrum that is considered within the equivalent baseband of the modulator. The optimizer-band is the sum of all fundamental frequencies that are used to design the power amplification stage. A more formal definition of those frequency bands can be found in subsection 3.3.

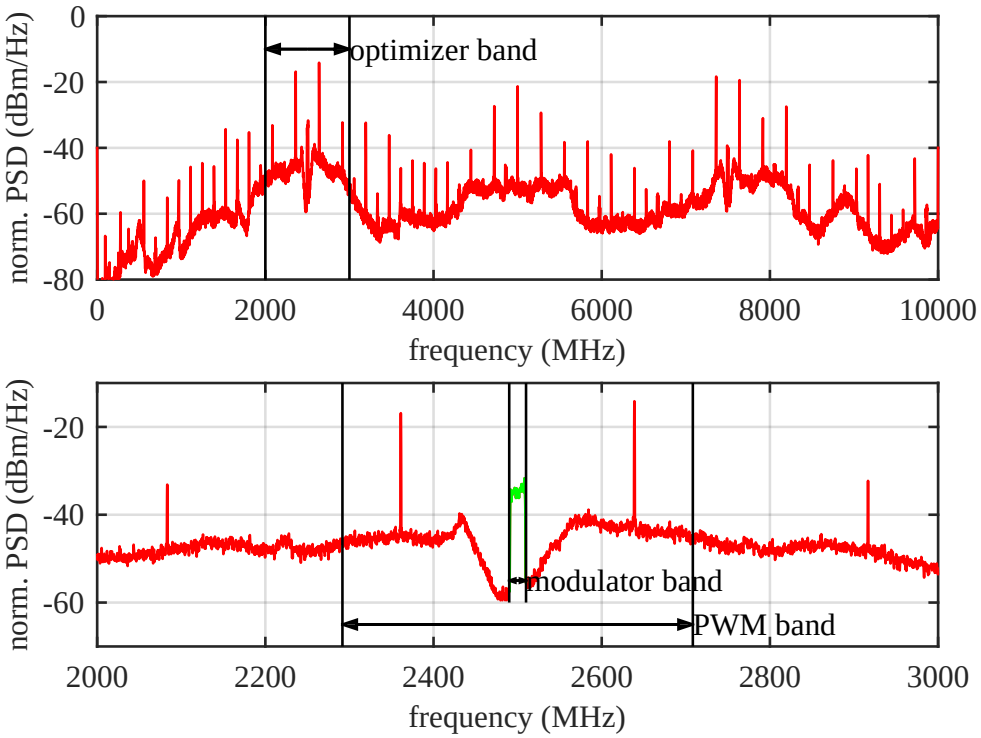


Figure 2.9: Power spectral density of DRFTx's output signal measured at its coupler for LTE excitation.

This calls for a specialized design approach since straightforward amplifying all signal components in the excitation, even if accomplished in energy efficient manner, is not suitable. Since the modulator noise is accounting for 95-98 % of the overall power, using an absorptive output filter would limit the overall power efficiency to

2-5 % and render the system to be inefficient per definition [11].

Filtering the modulator noise signal in between the booster and the PA stage would eliminate the need for a specialized design approach of the PA. However, this would also destroy the binary nature of the signal at this point. Binary signals represent a special case when dealing with nonlinear systems because their potential dynamic range does not degrade when applied to a static nonlinear system [18]. This makes binary excitation signals very attractive for nonlinear power amplifier designs and interesting from a scientific perspective. Therefore, it was decided to investigate solely the option of driving the PA with a binary signal.

Broadly speaking, typical continuous mode SMPAs utilize matching networks that optimize a single figure of merit over a given design frequency range implicitly assuming narrow-band excitation. In the case of continuous mode Class-F amplifiers for example, the power added efficiency (PAE) is optimized over their design frequency range [21] including their harmonics, however, without addressing any other signal components outside this frequency range.

The starting point for designing the PA was to adapt an existing continuous mode PA design [24] to the specification of this project. This was described in [25], although, this work expands the method by employing a banded approach. The idea behind it is to distinguish between modulator noise and intended RF signal already in the design phase and adapt the matching network and filter components accordingly, as described in chapter 3, section 3.3. By distinguishing these signal components, the target is to minimize the power consumption attributed to sustain the modulator noise while achieving efficient amplification for the desired output signal.

This is a paradigm shift with respect to many other RF-design approaches as the design evolves as much around the modulator noise as with respect to desired output signal. The very frequency selective transition between these bands is dictated by the design of the output filter. However, in order to achieve certain operational properties of the PA specific source- and load-trajectories have to be realized at the PA's transistor. The filter alone does not offer enough design space to accomplish this. However, by introducing a matching network between the active device and the output filter it is possible to overcome these issues.

From a design perspective this leaves two issues to be solved: Firstly, favorable source- and load-conditions need to be identified in order to optimize a design for such

a banded approach. Section 3.2 explains the relevant issues regarding this method and explain how this method was used in this project.

Secondly, a realizable matching network needs to be found that maximizes a figure of merit based on the load-pull contours. The theory behind this step is discussed in section 3.4 alongside the realization of the matching network for this work.

Practical designs are hindered mostly due to the huge bandwidth required for the load pull measurements. However, it has been demonstrated that designs based on load-pull measurements that cannot utilize modulated excitation [26] improve the DRFTx efficiency over standalone SMPA designs.

2.4 Output Filter

In a DRFTx the output filter serves a double purpose. On one hand, it is intended to suppress the modulator noise at the overall system's output. As depicted in Figure 2.9, the majority of the PA's output signal energy is located outside the modulation band. Therefore, the need for high suppression and the requirement for highly reflective terminations in harmonic controlled amplifiers is in favor of a very narrow band, high quality factor (Hi-Q) reconstruction filter.

On the other hand designing a matching network incorporating a high-Q factor filter is much more challenging and less stable with respect to manufacturing tolerances. It also introduces a potentially nonlinear feedback between active device and filter which gives rise to the generation of additional unwanted signal components [27] and causes prolonged load-modulation driven by the energy stored in the filter. If these effects exhibit time constants in the range of the PWM-modulator's sequence length or longer, this effect will significantly contribute to overall memory of the system. To ease the matching network design and to reduce nonlinear memory effects, theoretically, an absorptive or low quality factor filter would be beneficial.

Finally, the filter needs to be able to handle the system's output power, which is in the ballpark of the transistor's saturated output power of about 10 W.

This ultimately results in a trade off between potential energy efficient operation, memory depth and design complexity. Since this work aims at energy efficient operation, a highly reflective termination of the active device was chosen [11]. To ensure low insertion loss at a relatively low bandwidth of 1.6 % a fourth order interdigital filter

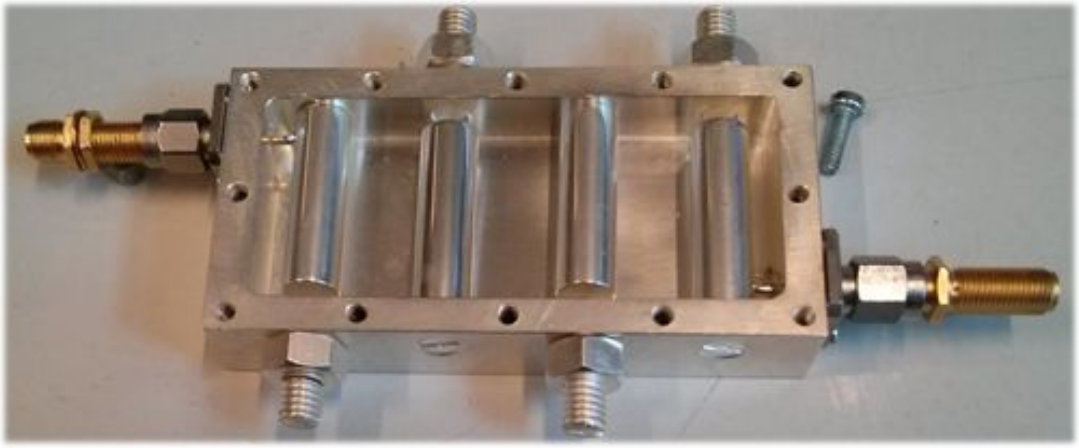


Figure 2.10: The output filter used in the DRFTx test bench.

with a bandwidth of 40 MHz was chosen. The filter was based on a computer aided design, carried out in *HFSS*, and custom built to provide best control over its properties. A photograph of the filter can be found in Figure 2.10.

Measured results of the filter's S-parameters are depicted in figure 2.11. Around the center-frequency of 2.5 GHz this unit provides the frequency selectivity needed to separate the desired RF-signal from other signal components. However, at the harmonics of this frequency the filter also couples to its output as well as at some other higher modes occurring above 5 GHz. If only the filter would be used to set the load condition to the transistor this would be an issue due to leaking harmonic content. Although, the matching network has a low pass characteristic which rejects this components before reaching the output filter. To get a more holistic result of this circuit section, see section 3.5.

Settling on a specific filter, PA and modulator combination is the most important design decision for this class of circuits. It is important to note that the nonlinear contributions of all of these parts are well researched as isolated phenomenons. However, only their interaction gives rise to memory induced, nonlinear behavior. Since this type of distortion is very interesting from an academic standpoint and not very well researched, it was decided to use a filter that will contribute to a deeper nonlinear memory instead of building just another nonlinear amplifier with short memory. Inherently this causes potentially bad distortion performance that goes with it, how-

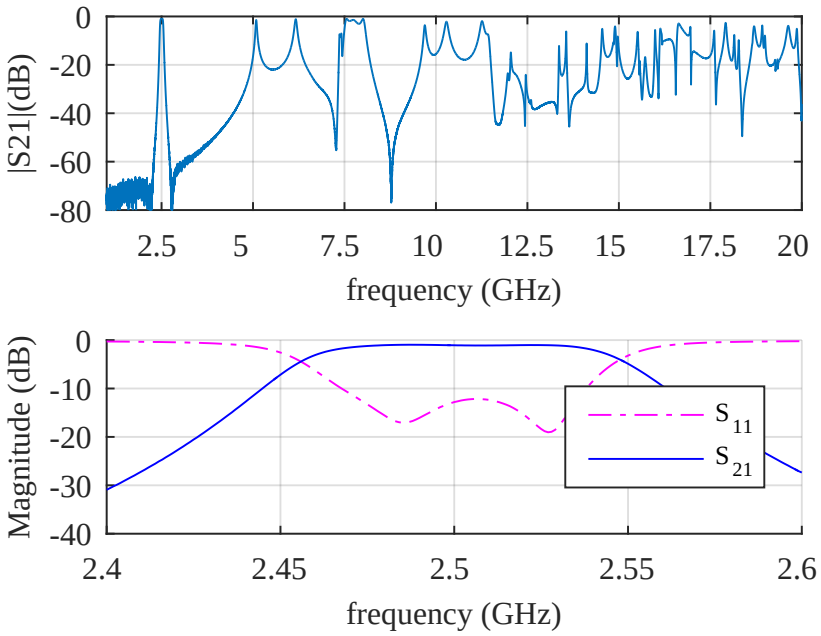


Figure 2.11: Measured characteristics of the filter utilized as the output filter in the DRFTx test bench.

ever, this is a more radical approach towards efficiency and also allows to study more pronounced memory induced, nonlinear behavior.

2.5 Output Coupler

The output filter is a necessary component within a DRFTx, however, when it comes to characterizing the circuitry it causes some issues. The output filter's narrow relative bandwidth makes many frequency components unobservable through the system's output. Especially when using very broad band instruments which have inherently lower dynamic range. Conducting circuit characterization solely based on the signal components present within the modulator band can be sufficient for distortion analysis, however, lacks most of the modulator noise which is mostly defining the PA's efficiency.

To overcome this issue, a coupler is integrated into the matching network between the transistor and the filter. This allows, in combination with a calibration step, to

gain access to almost the entirety of signals present at this stage. On top of this common deembedding techniques can be applied to directly measure current- and voltage-waveforms at the intrinsic transistor. The design and calibration of a coupler designed for this application was already published in [28], however, the most important aspects are summarized in this section.

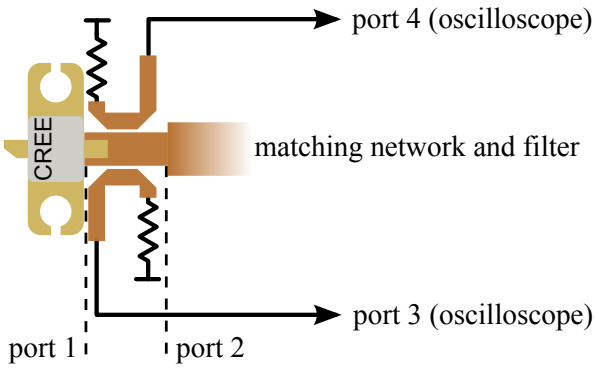


Figure 2.12: Dual directional coupler

For the design of the directional coupler several design issues have to be considered. Firstly, a microstrip line based topology is required to allow an integration into the transistor's matching network as described in section 3.3. Secondly, a high bandwidth is a major design goal to be able to cover a bandwidth of about 500 MHz up to 20 GHz.

To achieve a reasonable high bandwidth, the coupled length has been chosen such that the quarter wave resonance is out of the desired frequency band. Hence, the coupled length becomes small which lowers both coupling factor and directivity. While the low coupling factor is beneficial for the intended application, its other characteristics are suboptimum. However, due to the calibration presented in [28] the error of the directional coupler can be mathematically corrected to allow for exact measurement results.

2.6 Measurement Setup

The sum of all these components including some measurement instruments makes up the resulting DRFTx test bench. A conceptual diagram of it is depicted in Figure 2.13.

It allows to gain a multitude of different time- and frequency-domain measurements at important interfaces of the DRFTx. The individual interfaces are also named within Figure 2.13 to clarify from which signal specific measurements are derived.

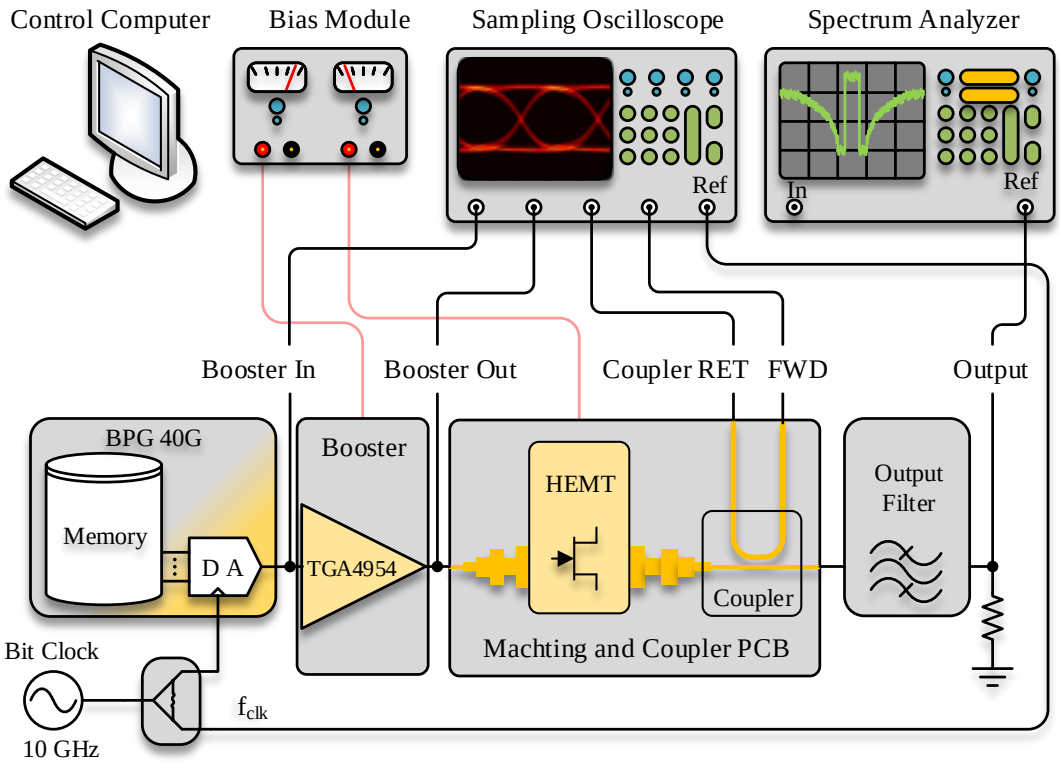


Figure 2.13: Test bench setup used for modeling

To put that into perspective, Figure 2.14 provides a photograph of the test bench.The picture also indicates some important signals, indicated by red lines with arrows. However, one has to keep in mind that several different measurements were conducted using this setup and only a typical configuration is depicted here. When discussing individual measurement results and the corresponding modeling in chapter 4, any deviation from this setup is explicitly stated.

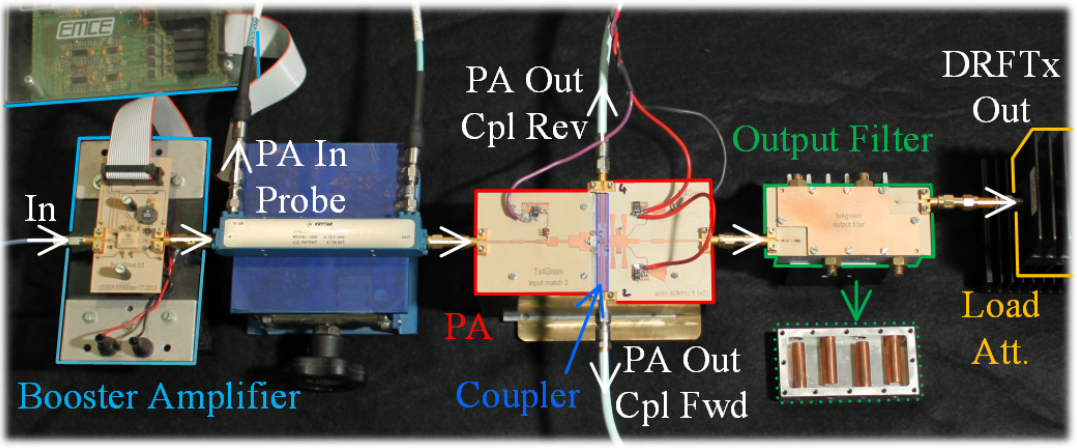


Figure 2.14: Photograph of the DRFTx used in the test bench.

CHAPTER 3

DESIGN OF DRFTX CIRCUIT MODULES

This chapter discusses the RF-circuit design specific issues evolving around the DRFTx setup that is explained in chapter 2. Some circuit components are vital for the practical use and functionality of the testbench, however, they are not directly related to the functionality as a signal source. As a result they are discussed in this chapter which is built around implementation. This is especially true for sections 3.1, 3.3 and 3.6.

As described in chapter 2, a DRFTx is a complex piece of circuitry and on top of that its modules are strongly interdependent. Chapter 2 illustrates how some design methods, which utilizes data gained from different functional units, optimize the overall functionality which is impossible to demonstrate for the individual unit. This chapter explains the more cohesive methods and the manufacturing challenges used in this work. Most prominent, the simplified real frequency technique (SRFT) is explained in section 3.4.

3.1 Bias Module for Experimental RF-Amplifiers

RF-amplifiers that operate close to their physical limits are inherently problematic in a test setup. They can fail due to many different reasons, however, thermal or electrical breakdown being the most common. In most cases it destroys the active device and, even more problematic, potentially also destroys the attached stages. This not only

results in financial inconvenience because simply replacing the active device is not enough to repair the setup. As active devices show significant component variations a redesign and re-manufacturing of the damaged circuits within the system as well may be required.

Sadly when designing highly efficient power amplifiers there is little to no space for safety margins in circuit design. Firstly, amplifiers need to operate in compression to function as SMPAs. This can be achieved by operating a device close to its electrical breakdown or utilizing a transistor designed for higher powers and restricting it artificially by lowering its bias. But, larger devices always suffer from lower gain and poorer frequency response which makes them ill suited for highly power efficient designs. Secondly, the waveform at the intrinsic amplifier cannot be measured in most cases. Therefore, under varying load conditions it is hard to predict if an active device is operated outside its specifications. If so, the device may spontaneously malfunction. Finally, the overall gain of cascaded amplifiers can be high enough that RF-leakage, for example through the bias circuits, can cause oscillations. Since this is a parasitic effect, that can occur not only at design frequencies, it is hard to predict. If it occurs it will likely lock up and destroy something by overheating it.

Since these problems cannot be addressed by overbuilding the circuit, they can be encountered by having failure protocols in place. If something fails, it is important to detect it and immediately power down the setup to a safe state. This is especially important since the test-setup is used to conduct many measurement runs in an automated but unattended fashion. In this work a purpose built bias unit was used to perform this functionality. It is a dedicated PCB with its own micro controller unit to decouple this functionality from the overall measurement setup. It allows biasing a circuit that requires up to four control voltages and two bias rails using a state machine. During measurements it monitors these voltages as well as currents and temperatures to ensure stable conditions.

This unit has been used extensively during this project to bias and monitor the booster amplifier, as well as the PA's transistor in some cases. Having a dedicated module to control the booster amplifier's bias in an automated manner is necessary since the integrated limiting amplifier, namely the *TGA4954-SL*, requires a specific biasing sequence. Its four control voltages and two supply rails that need to be turned on and off in a specific sequence and some of the bias parameters depend on each other.

For example, two control voltages are found by measuring the idle current consumption of the overall circuit and adjusting the control voltages iteratively until a certain quiescent current ratio is achieved. Whenever it is required to alter the bias parameters for different measurements, it is necessary to power down the signal source and the booster module in a controlled manner and to start over with the adaptive bias routine.

3.2 Load Pull Measurements and Setup

Biasing a transistor consistently is the first step to ensure RF-measurements are reproducible. The measurements gained from monitoring the bias can help to determine if the device is still functioning and well, however, these results carry little information about the RF-properties of the device. In recent years sophisticated computer-models and nonlinear measurement techniques have made an effort to replace lab experiments as the design method of choice to build RF-PAs. Although, especially when operating in a strongly nonlinear regime some measurement based characterization techniques have stood the test of time up to this day. Load pull setups for example are common place in today's RF-PA design [29]. While being around for decades, the need to deal with non-unilateral and nonlinear devices still poses design challenges that are hard to tackle using solely simulations.

Load pull setups allow to measure the performance of transistors directly by exciting the active device using an RF signal source while manipulating the source- and load-impedance using so called tuners. It allows to record different performance measures as a function of the impedance imposed by the tuner. Typically, signal power levels and DC power consumption would be used to identify favorable input- and output-reflection coefficients for the given device. What is ultimately chosen to be used as the figure of merit and how to turn this information into a circuit defines the design of the PA.

A very basic load pull setup is depicted in Figure 3.1. There are many alternatives to this kind of setup that expand the capabilities with respect to bandwidth, harmonic tuning, modulated excitation, and so on. Choosing a specific setup strongly depends on the requirements imposed by the overarching project. In fact there are so many that discussing all possibilities is out of the scope of this work, however, suitable literature

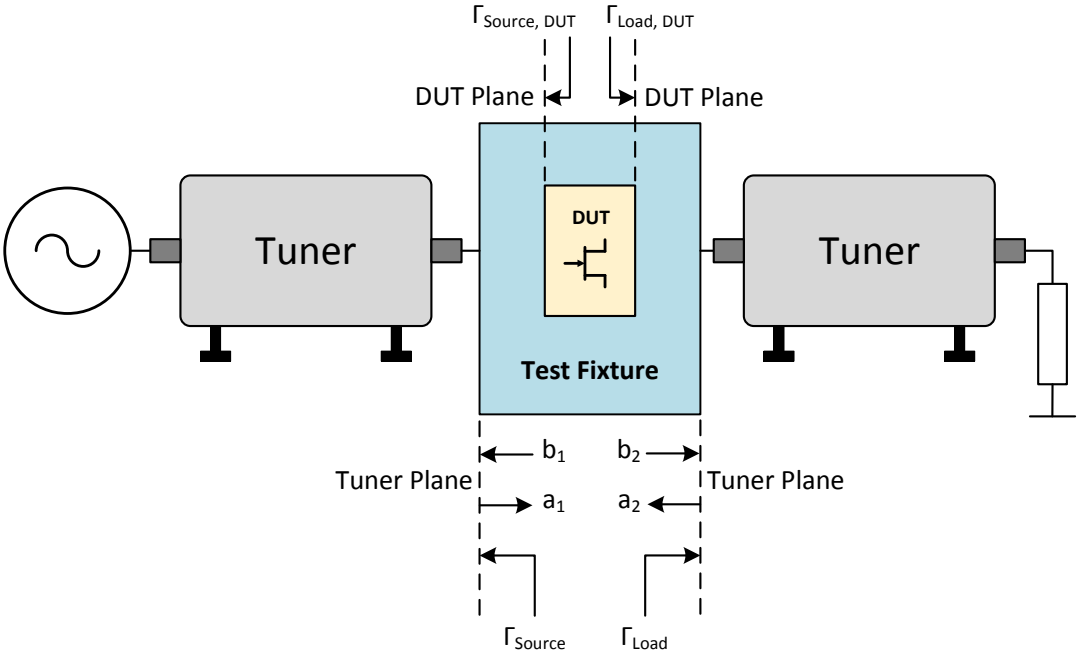


Figure 3.1: Basic Load Pull Setup

is available [29]. Since a passive tuning setup was available, it was decided to utilize it for this project. Although, this still leaves room for many implementation options. While the basic idea is always the same, each approach has some intrinsic limitations or benefits. The relevant alternatives are briefly discussed in the following:

Passive tuner setups suffer from a tuning range that is bounded by having smaller than unity reflection coefficients. Additionally, passive tuners, which are usually based on a double slug tuner approach, are intrinsically narrow band and, therefore, can only be utilized with a CW or narrow band excitation.

Harmonic tuning concepts allow to control impendances at the fundamental and the harmonics of the excitation signal. If realized with a passive system they require multiple tuners, and therefore, have higher insertion losses. Depending on the implementation, they require sophisticated signal combiners to separate and combine the individual harmonic components, that intrinsically limit the available frequency range of the setup. Alternatively, multiple tuners can be cascaded to gain the desired additional design space. In the latter case, the dependencies between the individual tuners are captured by a more sophisticated calibration, although, achieving less precise re-

sults.

Active concepts can deal with modulated excitation and have no intrinsic limitation with respect to the synthesized reflection coefficients. However, since the wave traveling towards the active device (a_x) is artificially generated, these setups require some sophisticated active circuitry. The active tuner needs to handle the power required without distorting the signal, which is in the ballpark of the active device's saturated power. They need to synthesize given impedance trajectories for the design bandwidth, which directly translates to bandwidth demands of all the circuits involved. When harmonic concepts are employed, the individual signal bands around the harmonics need to be handled separately requiring additional active hardware. The overall bandwidth is further limited due to the fact that the signal's bandwidth at the harmonics grows with its order.

Finally, there are many kinds of hybrid approaches that combine some of these concepts, like using a passive tuner for the fundamental frequency and active branches at the harmonics for example.

So choosing the most suitable load pull setup out of this variety is not a trivial task. The design of the PA for the DRFTx in this work would require a baseband bandwidth in the range of hundreds of MHz due to the employed modulator. That would require an extraordinarily expensive active load pull setup. Out of the passive options, cascaded tuning has been found to be the best compromise between setup complexity, measurement control ability, and, most important, bandwidth. Such a setup is depicted in Figure 3.2 and was used extensively throughout this project.

The setup uses only one tuner at the input, allowing to control solely the fundamental impedance at the active device's input, however, three tuners at the output to jointly control the impedances at the fundamental, first, and second harmonic at this interface. Circuit simulations have shown that this tuner combination yield sufficient control over the circuit's behavior to optimize gain and power efficiency. Figure 3.3 depicts the composition of three different reflection coefficients, realized at the fundamental- f_1, the second- f_2 and third- harmonic f_3 frequency. The contribution of the individual output tuners (F1, F2, and F3) are color coded correspondingly.

Since the resulting impedance at the individual frequencies is dependent on the mutual state of all load tuners, the mutual state of the tuners that result in the desired reflection coefficients needs to be found by pre-characterization measurements. This

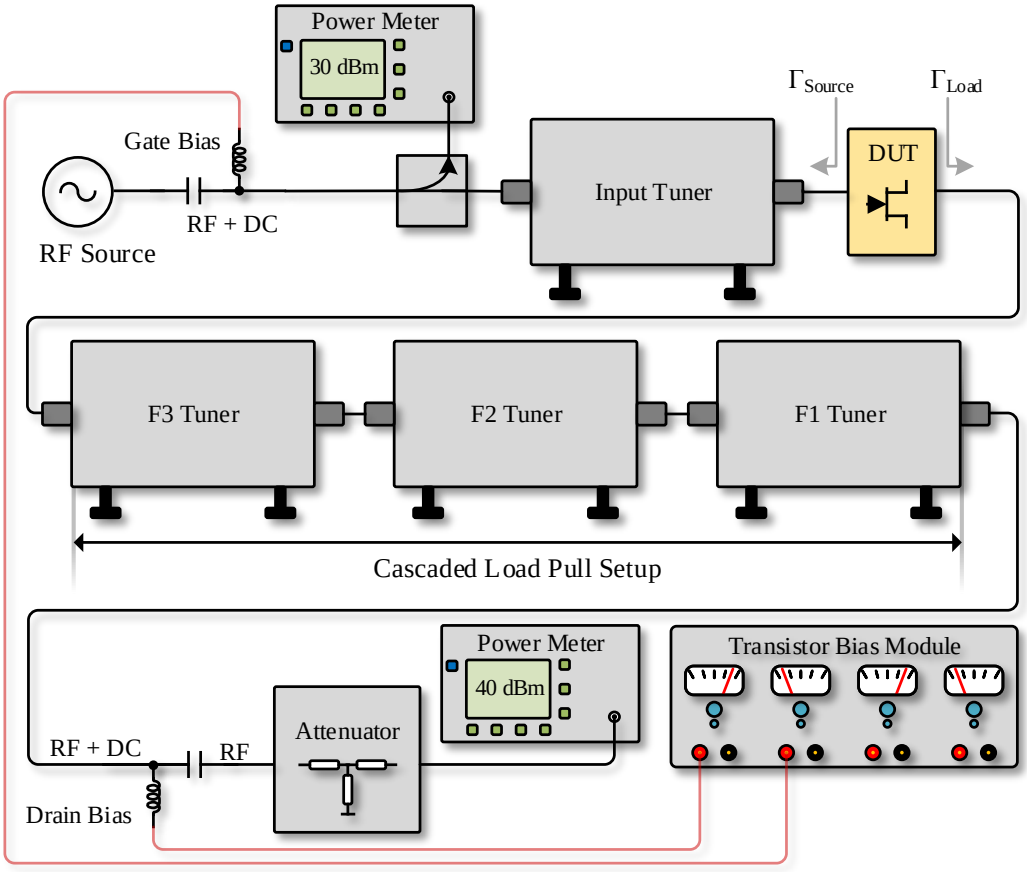


Figure 3.2: Basic harmonic load pull setup

is based on a software package by *Maury Microwave* called *ATS5*. While the cascaded setup provides the opportunity to expand the capabilities to harmonic tuning, the mutual dependency also causes higher uncertainties when synthesizing specific impedance values. This design is based on [30] which explains these issues in greater detail.

The setup covers a broad frequency range f_{design} around the chosen fundamental f_1 and is, therefore, often considered broadband. This is a problematic nomenclature since the load pull setup uses a swept CW source while the DRFTx target system (see Figure 2.2) uses a constant envelope, yet heavily modulated source to drive the amplifier. So while the setup is capable of covering the whole bandwidth, that is occupied by the PWM-modulator's excitation signal, it is important to note that the RF-excitation used in both setups is significantly different. However, works like [12] allow to pro-

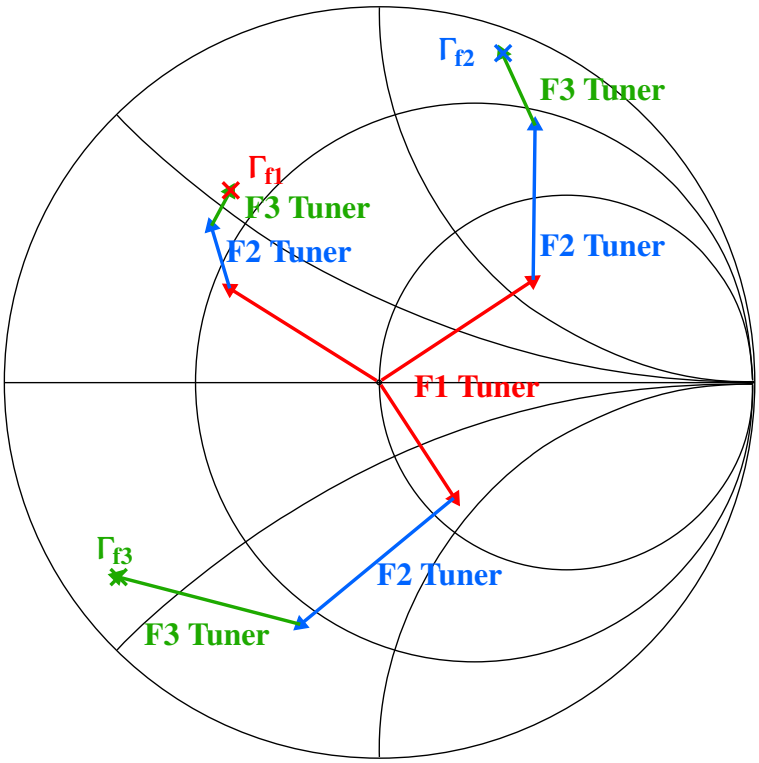


Figure 3.3: Tuner dependencies in cascaded tuning setups

vide an estimated power efficiency based on load pull results. Therefore, it is possible to derive constructive system design measures for broad band modulated systems even though the measurements are derived from a narrow band load pull setup.

Measurements and simulations indicate that the basic setup depicted in Figure 3.2 lacks tuning range with respect to the maximum reflection coefficients imposed by the system's losses. Especially for the harmonic signal components at the transistor's output, highly reflective terminations are required to achieve beneficial load trajectories, which are identified in section 3.3. So cascading tuners alone is very useful to cover a very large proportion of the design space governed by the passive reflection coefficients. This way allows to narrow down the design space to dedicated regions of interest (RoI). However, after a specific range of reflection coefficients have been identified as RoI, additional circuitry is required to improve the maximum reflectivity

and resolution in this region. So called pre-matching circuits are introduced between tuner and DUT to accomplish this.

A setup including this extension is depicted in Figure 3.4. It is basically a passive structure that provides a controlled static reflection close to the active device. Since this is another lossy device, it also reduces the overall tuning range, however, it shifts it towards a RoI within the Smith-chart as illustrated in Figure 3.5.

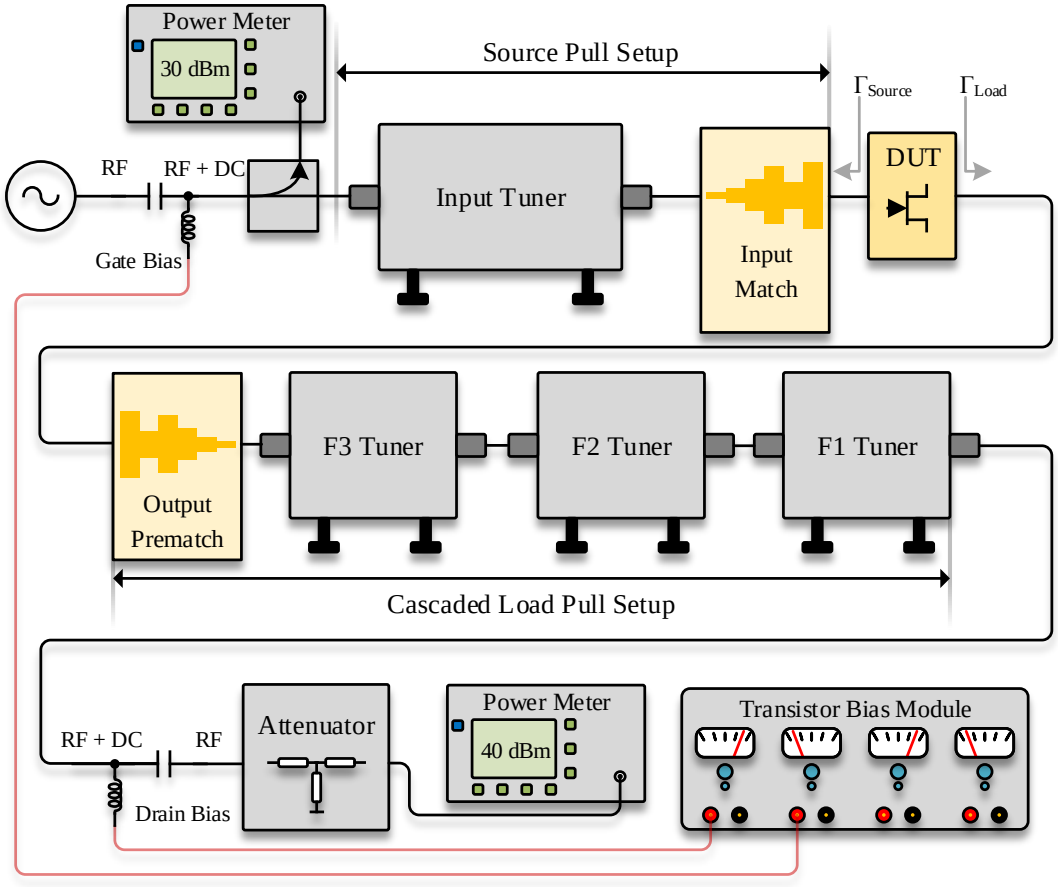


Figure 3.4: Prematched harmonic load pull setup

The calculation of the resulting reflection coefficient Γ_{Total} is shown in eq. (3.1), where S_{12}, S_{21}, and S_{22} are the S-parameters of the pre-match. As a result, the maximum achievable reflection coefficient Γ_{max} is greater than the maximum reflection coefficient of the tuner alone.

$$\Gamma_{Total} = \Gamma_{Prematch} + \frac{S_{12}S_{21}\Gamma_{Tuner}}{1 - S_{22}\Gamma_{Tuner}} \tag{3.1}$$

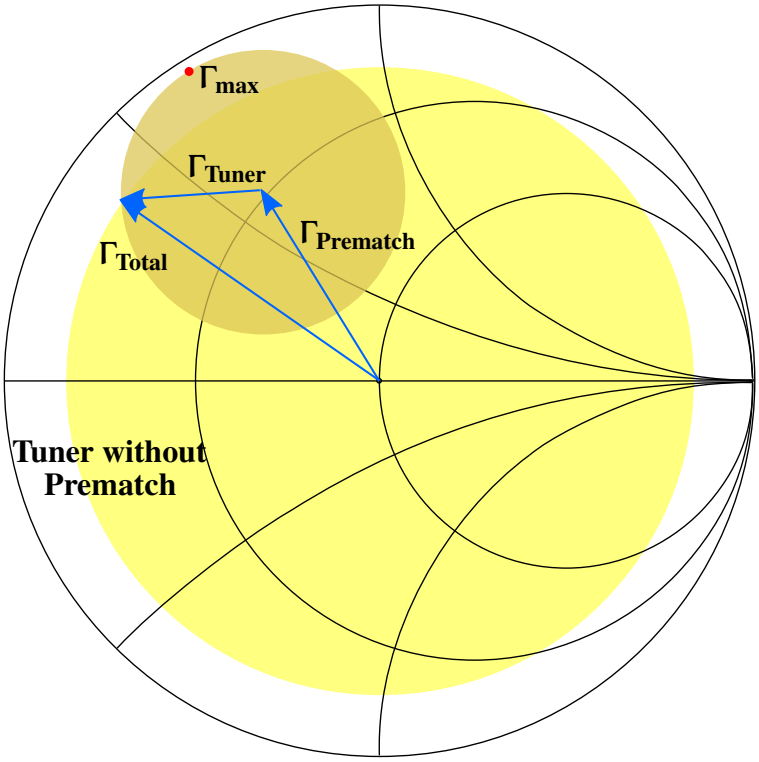


Figure 3.5: Tuning range with and without pre-match

Pre-matching is very useful to improve the coverage of a load pull setup for a limited RoI. However, in order to complement a cascaded tuning setup the pre-match circuit needs to take on a specific value for all harmonics that are considered in the given setup. In this work the number of harmonics considered is denoted by $N_{PA} = 2$. As aforementioned, by using cascaded harmonic tuning it is possible to cover a wide measurement bandwidth. However, in order to make use of this flexibility the RoI dictated by the pre-match circuit needs to encompass the optimum within the Γdesign space for all considered design frequencies which are denoted as DBW_i at the individual harmonics $i = 1..N_{PA}$. Therefore, prior knowledge of the DUT's behaviour is required in order to tailor the pre-match circuits according to these requirements. By using

the cascaded tuning measurement setup without pre-match, RoIs for the individual frequencies can be estimated through interpolation. The measured results were in good agreement with circuit simulations. The synthesis of circuits based on this data is described in section 3.4.

Since this design procedure is based on having multiple frequency bands and some also corresponding harmonics, the individual frequency bands are also summed up in eq. (3.2).

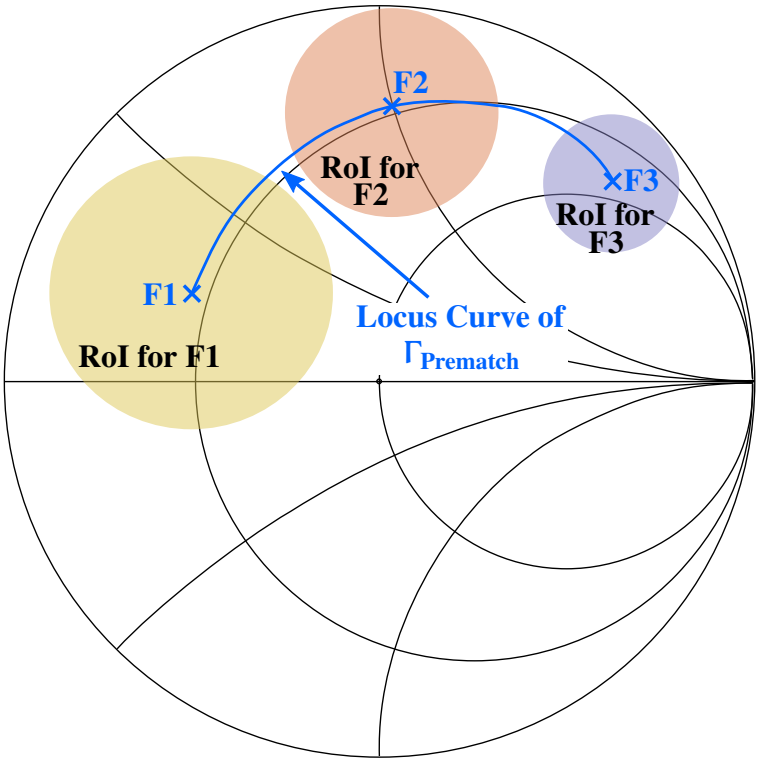


Figure 3.6: Tuning range with prematching for three harmonics

The optimum reflection coefficient $\Gamma_{ideal}(f)$ for the given circuit can be located using the pre-match design based on the framework as denoted in eq. (3.1). Figures 3.5 and 3.6 illustrate the transformation of the tuning ranges to a certain area of interest for the design bandwidth DBW_i at each considered harmonic.

3.3 Matching Network Design

The spectrum of a typical PWM modulated signal for a 20 MHz LTE signal, is illustrated in Figure 3.7. It shows that the spectrum is segmented into the intended output signal $f_{Mod} = 20\,\text{MHz}$ centered around $f_1 = 2.5GHz$, resulting in the occupied bandwidth OB and the modulator noise covering almost the rest of the spectrum.

However, the majority of the power is concentrated around the harmonics f_i. This is the reason why the design band DB was chosen to cover most of the overall signal power, resulting in a design bandwidth of $f_{design} = 1000\,\text{MHz}$. If the signal source's DAC would have a perfectly flat frequency response the excitation spectrum would by cyclic in $2 \cdot f_0$. Therefore, also the higher order design bands DBW_i need to be considered. It was found that considering the first two harmonics is sufficient for the design since the gain of the active device is very low beyond this band. Therefore, the the design bands DBW_i have been specified according to n_{PA}.

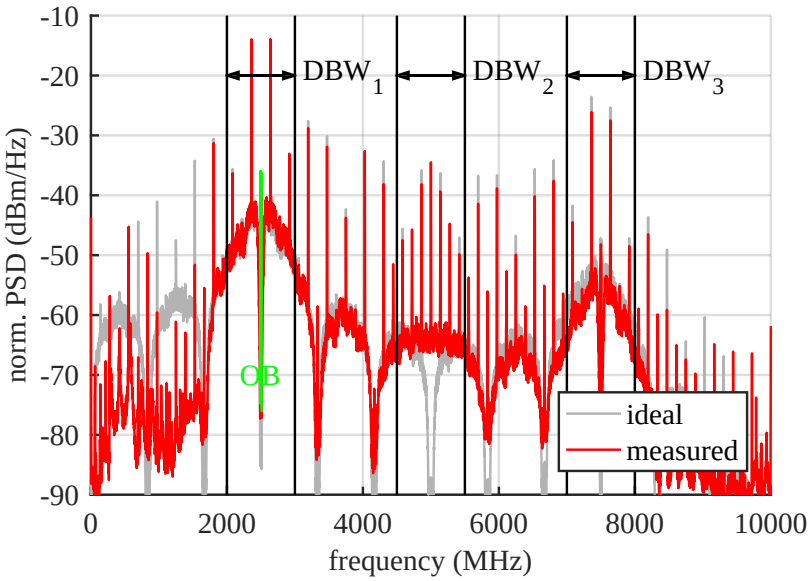


Figure 3.7: Normalized power spectral density of the LTE excitation signal.

$$
\begin{aligned}
f_i &= 2.5GHz \cdot i \\
f_{Mod} &= 20\,\text{MHz} \\
n_{PA} &= \{1..N_{PA}\} = \{1,2,3\} \\
OB &= f_0 + \left[\frac{-f_{Mod}}{2}, \frac{f_{Mod}}{2}\right] \\
f_{design} &= 1000\,\text{MHz} \\
DBW_i &= f_i + \left[\frac{-f_{design}}{2}, \frac{f_{design}}{2}\right] \\
NLBW_i &= f_i + [\frac{-f_{BW}}{2}, \frac{f_{BW}}{2}] \cdot i \\
DB &= \bigcup_{i=0}^{N_{PA}} DBW_i \\
NB &= \bigcup_{i=0}^{N_{PA}} DBW_i \cap OB
\end{aligned}
\tag{3.2}
$$

The signal components in these bands stem from the amplified excitation as well as from nonlinear signal components produced in the transistor. This distinction marks another important difference to more mainstream SMPA designs, in which case the bands occupied by the nonlinear mixing products require a bandwidth that is proportional to its order as in $NLBW_i$. This is depicted in Figure 3.8 where the output signal of the active device as observed through the coupler, is compared to the measured input signal. The design band DB consists of all frequencies considered for designing the matching network, while the noise band NB is a subsection of it. All these frequency ranges are defined in equation eq. (3.2) and are used to distinguish them for the banded matching network design approach described in this subsection.

By using a load pull setup it is possible to gain many load-pull contours that measure a load pull parameter as function of the realized input- and output-reflection-coefficients. For a broad band problem also the frequency is a variable, since the contour will change significantly with frequency f. So every load-pull parameter results in a function of form $LP(\Gamma_{in}, \Gamma_{out}, f)$.

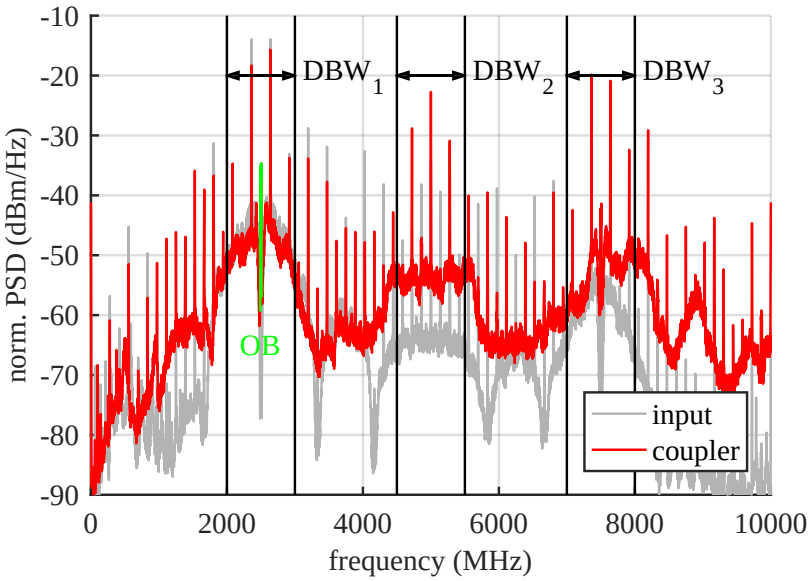
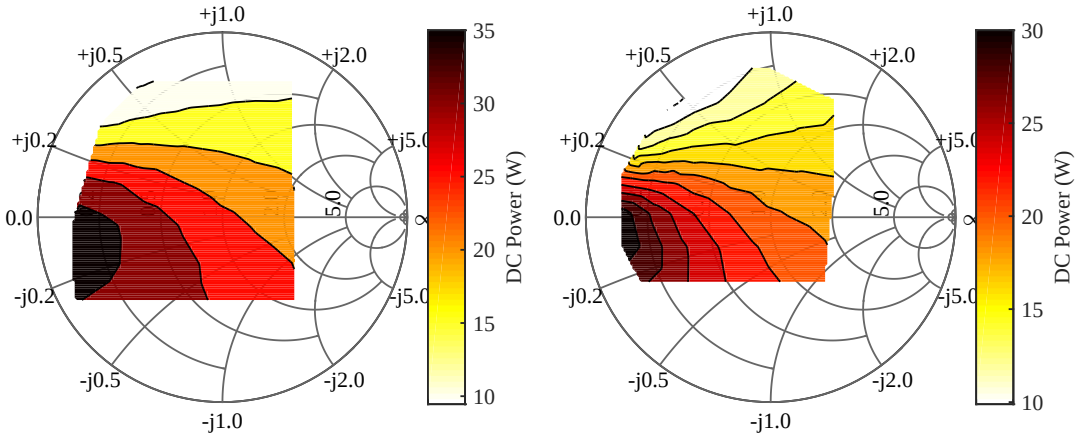


Figure 3.8: Normalized power spectral density of the PWM-modulated LTE signal at the in- and output of the PA.

$$\begin{aligned} FG(\Gamma) &= |Gain(\Gamma)| \\ FPAE(\Gamma) &= |PAE(\Gamma)| \\ FPD(\Gamma) &= -|P_{Bias}(\Gamma)| \\ FOUT(\Gamma) &= \begin{cases} FOM_{PAE} \cdot cPAE, f \in OB \\ FOM_{PD} \cdot cPD, f \in NB \end{cases} \end{aligned} \tag{3.3}$$

While this is the most general description, it somewhat problematic since it requires many points to be measured in order to analyze a single parameter. Since the data gathering is based on mechanical tuners this translates to long measurement times. However, this can be avoided by first starting with an optimization step. For every considered frequency point, the load-pull setup conducts a joint optimization of the input- and output- load condition for every analyzed system parameter. When the ideal load combination is found, two extensive measurement runs are conducted at the load- and source- side while the respective other is fixed to a specific value. Early in the design this is the before found optimum value, as the design progresses, the most likely value will be used. This allows to analyze the impact of a reflection coefficient deviating

from its optimum, however, it requires an optimum to be defined for each parameter. The resulting contours are called load pull contours and Figure 3.9 provides an example. As a result a load pull contour is a simpler function that only depends on one reflection coefficient as in $LPC(\Gamma_{test}, f)$. These contours allow to indicate favorable source and load conditions depending on a given application. However, in order to gain a single parameter that can be optimized figures of merit need to be defined. They use a source or load trajectory $\Gamma(f)$ in order to gauge how well a circuit parameter is optimized for this trajectory. Some examples are denoted in eq. (3.3). Usually harmonic controlled amplifiers rely on two static FOMs for in- and output- to gauge the optimality of a matching network.



(a) 2.0 GHz (b) 3.0 GHz

Figure 3.9: Measured DC power consumption load pull contour.

For the presented solution banded FOMs are employed at the output to differentiate between the parts of the spectrum that are occupied by the intended output signal present in *OB* and the modulator noise occupying *NB*. Within the signal band the same FOM is used as in many popular class F amplifier designs. It is a FOM that tries to optimize the power added efficiency for all the components in this bandwidth ($FPAE(\Gamma)$). Outside this band and within regions where a significant amount of modulator noise is present (*NB*) a minimum power consumption ($FPD(\Gamma)$) is important for this design. Therefore, this FOM is used in these sections of the spectrum. Figure 3.9 illustrates the results of these load pull measurements related to the power consump-

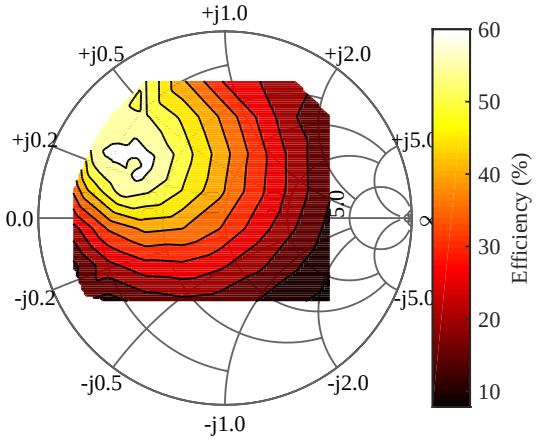


Figure 3.10: Measured PAE load pull contour at 2.5 GHz.

tion. The location of the optimum with respect to the output power maximum is depicted in Figure 3.10. They are contradictory, however, by using a banded approach as described by (FOM_{output}) a compromising load trajectory can be found based on this FOM. The coefficients *cPAE* and *cPD* are used to adjust the weight of the FOMs with respect to each other. Deriving a circuit from these results asks for an optimizer based circuit design which is described in section 3.4.

For the PA's input a single FOM is employed that is used throughout the overall design bandwidth *DB*. In this the gain of the transistor is maximized via $FG(\Gamma)$ in order to have plenty of output signal available.

3.4 Circuit Synthesis Based on SRFT

As described in section 3.3, the source- and load-contours for optimum operation of the transistor were determined with the help of load pull measurements. Especially the load-contours are affected by the banded approach since the figure of merit changes frequency dependent from power consumption to PAE.

In a next step, a realizable matching network for the transistor's input and output needs to be found based on these load pull measurements. The so-called Simplified Real Frequency Technique (SRFT) method was used for this task. Because of the problem's broadband nature analytical approaches, based on the classical gain-bandwidth

theory, cannot be applied in order to gain a realizable circuit.

The SRFT is a semi-analytical approach which is superior in many ways. It allows to find a lossless circuit that maximizes an arbitrary figure of merit instead of only optimizing power transfer. This method was proposed in [31] and has been refined continuously [32]. It was used in many other works [33] [34], as well as in the works [25] [30] related to the overarching project. A modified version was used in [26] which already employs a banded approach and forms the basis of this section.

As illustrated in Figure 2.8, matching networks connect a DUT to other circuit sections which realize a certain impedance trajectory. The method's goal is to design two lossless two-ports that match the generator- and the load-side of the network where both can take on complex impedances. Matching is here used in a broader sense since it not necessarily means to optimize solely the power transfer, but to optimize some FOM for this section. Finding a circuit that fulfills this condition is often referred to as a double matching problem, since both sides of the matching networks can take on complex impedances.

The SRFT allows to do so by, firstly, finding an analytic description for all realizable two ports using polynomials. Secondly, it allows to compute S-parameters describing the matching networks from these polynomials. In conjunction with measured S-parameters of source- and load-impedances the reflection coefficients at the transistor's interface can be computed. The figure of merit 3.3 is now used to compute an error function. Thirdly this error function is minimized by a nonlinear optimizer. Since the polynomial description ensures a valid result, the final optimized polynomial also represents a realizable circuit.

In the next step the optimized polynomial is used to compute circuit parameters. In this case a transmission line implementation was chosen. Characteristic impedances and electrical lengths are used to describe the circuit to be manufactured.

While the SRFT ensures that the results are meaningful in a physical sense they can very well be impossible to manufacture, due to the for example ridiculously small characteristic impedances or so on. Finally, computer aided circuit design and EM-simulations are utilized to design a manufacturable circuit with parameters as close to the theoretical ones as possible. The individual steps of this process are described in the upcoming subsections.

The SRFT algorithm is independent of the chosen type of circuit elements used

for the matching network and derives a circuit of given predefined order. It consists of cascaded circuit sections, so called unit elements (UE), the values of which are found by minimizing a cost function or respectively maximizing a figure of merit. Each of these UEs is defined by its characteristic impedance Z_0 and length l and the propagation constant β. When a lossless transmission line is attached to a generator with impedance Z_G and terminated with an arbitrary load impedance Z_L at the distance $z = -l$, the input impedance Z_{IN} can be written as:

$$Z_{\mathrm{IN}} = Z_0 \frac{1 + \Gamma e^{-2j\beta l}}{1 - \Gamma e^{-2j\beta l}} \tag{3.4}$$

By replacing the reflection coefficient Γ by its definition, the input impedance computes as:

$$Z_{\mathrm{IN}} = Z_0 \frac{(Z_L + Z_0) + (Z_L - Z_0)e^{-2j\beta l}}{(Z_L + Z_0) - (Z_L - Z_0)e^{-2j\beta l}} = Z_0 \frac{Z_L + Z_0 j \tan \beta l}{Z_0 + Z_L j \tan \beta l} \tag{3.5}$$

Applying the Richard's transform [35], as defined in eq. (3.6), allows to simplify the input impedance to:

$$\tau = j\Omega = j \tan \theta = j \tan \beta l \tag{3.6}$$

$$Z_{\mathrm{IN}} = Z_0 \frac{Z_L + Z_0 \tau}{Z_0 + Z_L \tau} \tag{3.7}$$

In these equations τ is called Richard's variable. By specializing the circuit to using only line sections of a fixed length, the input impedance of a terminated line can be denoted with the help of a rational function [36, section 4.14][37]. It is important that the electrical length βl is shorter than $90\,^\circ$, because of the singularity of the transform. Figure 3.11 illustrates the cascaded connection of two UEs.

As in eq. (3.7), the impedances $Z_{\mathrm{IN},1}$ and $Z_{\mathrm{IN},2}$ in the Richard's domain are given

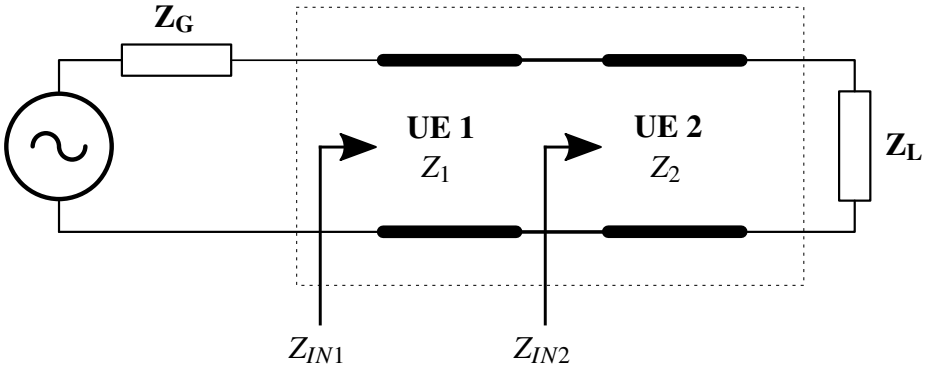


Figure 3.11: Cascaded Connection of two UE's

by:

$$Z_{\mathrm{IN},1} = Z_1 \frac{Z_{\mathrm{IN},2} + Z_1\tau}{Z_1 + Z_{\mathrm{IN},2}\tau} \tag{3.8a}$$

$$Z_{\mathrm{IN},2} = Z_2 \frac{Z_L + Z_2\tau}{Z_2 + Z_L\tau} \tag{3.8b}$$

By inserting $Z_{\mathrm{IN},2}$ in $Z_{\mathrm{IN},1}$, eq. (3.8a) changes to:

$$Z_{\mathrm{IN},1} = \frac{Z_1^2\tau^2 + Z_1Z_2(Z_1+Z_2)\tau + Z_1Z_2}{Z_2^2\tau^2 + (Z_1+Z_2)\tau + Z_1Z_2} \tag{3.9}$$

This is again a rational function in the Richard's variable τ. Hence, the input impedance of commensurate transmission lines can be generalized and formulated as rational function in nominator coefficients a_n and denominator coefficients b_n (eq. (3.10)). In eq. (3.10) and 3.11, n is the numbers of UEs in the input and output matching network.

$$Z_{\mathrm{IN}} = \frac{N(\tau)}{D(\tau)} = \frac{a_n\tau^n + a_{n-1}\tau^{n-1} + \ldots + a_1\tau + a_0}{b_n\tau^n + b_{n-1}\tau^{n-1} + \ldots + b_1\tau + b_0} \tag{3.10}$$

This is a very convenient formulation, consisting of a nominator polynomial $N(\tau)$ and a denominator polynomial $D(\tau)$. By normalizing the UEs' wave impedances, the input reflection coefficients can be calculated as:

$$\begin{aligned} S_{\mathrm{IN}}(\tau) = S_{11}(\tau) = \frac{Z_{\mathrm{IN}}-1}{Z_{\mathrm{IN}}+1} = \frac{N(\tau)-D(\tau)}{N(\tau)+D(\tau)} = \\ = \frac{h(\tau)}{g(\tau)} = \frac{h_1\tau^n + h_2\tau^{n-1} + \ldots + h_n\tau + h_{n+1}}{g_1\tau^n + g_2\tau^{n-1} + \ldots + g_n\tau + g_{n+1}} \end{aligned} \tag{3.11}$$

Since the reflection coefficient's magnitude of passive systems is bounded by one, $S_{11}(\tau)$ must be strictly free of open right half plane poles. Thus, the polynomial $g(\tau)$ must be strictly Hurwitz and furthermore $D(\tau)$ and $N(\tau)$ have to be Hurwitz polynomials as well [36, chapter 4]. Because losslessness is assumed, the squared magnitude of S_{21} of a passive two-port fulfills the following conditions:

$$|S_{21}(\tau)|^2 \leq 1 \tag{3.12}$$

$$|S_{21}(\tau)|^2 = 1 - |S_{11}(\tau)|^2 \tag{3.13}$$

As a consequence, the losslessness condition can be reformulated as follows:

$$\begin{aligned} |S_{11}|^2 + |S_{21}|^2 = 1 \\ S_{11}(\tau)S_{11}(-\tau) + S_{21}(\tau)S_{21}(-\tau) = 1 \end{aligned} \tag{3.14}$$

With this knowledge, S_{21} can be calculated as:

$$S_{21}(\tau) = \frac{f(\tau)}{g(\tau)} = \frac{\tau^q(1-\tau^2)^{\frac{k}{2}}}{g(\tau)}, \tag{3.15}$$

where q is the number of transmission zeroes at DC and k the number of UEs. If it is defined that $q = 0$ the result is a low pass filter. For $q > 0$, a bandpass filter is designed. By applying the reciprocity theorem, the full S-parameter matrix can be denoted as:

$$S(\tau) = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} = \begin{bmatrix} \frac{h(\tau)}{g(\tau)} & \frac{f(\tau)}{g(\tau)} \\ \frac{f(\tau)}{g(\tau)} & -\frac{f(\tau)h(-\tau)}{f(-\tau)g(\tau)} \end{bmatrix}. \tag{3.16}$$

Therefore, filters consisting of commensurate transmission lines can be described with the help of rational functions, which is called the Belevitch form. By computing the resulting load trajectory at the active device, using the S-parameters of the matching network and the load itself, the overall FOM can be calculated. This completes one

cycle in an optimizer run in order to solve the broadband double matching problem using SRFT.

In the SRFT's standard form the total power transfer is optimized by changing the transfer function of the equalizer filter based on an optimizer. Because the transfer function, which is described using the Richard's variable, is highly nonlinear, a nonlinear optimizer is required. However, any other FOM can be used as well to provide feedback to the optimizer.

To get started with the SRFT algorithm, the coefficients of the nominator polynomial $h_{\mathrm{IN}}(\tau)$ have to be initialized. The polynomial $f_{\mathrm{IN}}(\tau)$ is generated by including the transmission zeros at DC and the denominator polynomial $g_{\mathrm{IN}}(\tau)$ is generated as a strictly Hurwitz polynomial by spectral factorization of the losslessness condition:

$$g_{\mathrm{IN}}(\tau)g_{\mathrm{IN}}(-\tau) = h_{\mathrm{IN}}(\tau)h_{\mathrm{IN}}(-\tau) + f_{\mathrm{IN}}(\tau)f_{\mathrm{IN}}(-\tau) \tag{3.17}$$

The coefficients $h(\tau)$ are optimized iteratively by minimizing the cost function, or in this case maximizing the FOM, over the frequency band of interest.

The scattering matrix has to be transformed into impedance values to construct a circuit using either lumped elements or commensurate transmission lines. This procedure is called Richard extraction and the normalized characteristic impedance of each UE is given by:

$$Z_i = Z_{i-1}\frac{1+\Gamma_i(1,Z_{i-1})}{1-\Gamma_i(1,Z_{i-1})}, i = 1,2,\ldots,n \tag{3.18}$$

The initial reflection coefficient Γ_1 can be calculated as follows.

$$\Gamma_1(1,Z_0) = \frac{h_1+h_2+\ldots+h_n+h_{n+1}}{g_1+g_2+\ldots+g_n+g_{n+1}} \tag{3.19}$$

With the normalized impedance Z_0 equal to one, the first elements impedance can be calculated. The following reflection coefficients can be written as a function of the actual reflection coefficient.

$$\Gamma_{i+1}(\tau,Z_i) = \frac{1+\tau}{1-\tau}\Gamma_i(\tau,Z_i), i = 1,2,\ldots,n \tag{3.20}$$

Therefore, the actual reflection coefficient of the UEs is calculated by:

$$\Gamma_i(\tau, Z_i) = \frac{K(i-1,i) + \Gamma_i(\tau, Z_{i-1})}{1 + K(i-1,i)\Gamma_i(\tau, Z_{i-1})}, i = 1,2,\ldots,n \tag{3.21}$$

$$K(i-1,i) = \frac{Z_{i-1} - Z_i}{Z_{i-1} + Z_i}, i = 1,2,\ldots,n \tag{3.22}$$

Consequently, the impedance values of each UE are calculated using the Richard extraction. By building a circuit consisting of commensurate transmission lines, the width of each line is determined by the impedance. However, to finally determine the geometry also the line length needs to be specified. Because of the Richards transform (see equation 3.6) all elements need to be of equal electrical length.

All circuits in this project, that have been synthesized using SRFT, were designed as a low pass filter. Therefore, a cutoff frequency f_{max} needs to be introduced which defines the length of the commensurate line elements. From an implementation point of view, the cutoff frequency should be as low as possible, since this results in having cyclic repetitions of the elementary transfer function at rather high frequencies. Otherwise the repetitions of the pass-band may interfere with harmonics. From the realization point of view, the line length should be as large as possible to limit the influence of manufacturing tolerances and the influence of fringing fields. By introducing a design parameter κ, the stop- and pass-band roll-off can be designed to achieve a trade-off between the limitation of manufacturing tolerances and the influence of the periodicity of the equalizer. Therefore, the line length can be calculated as

$$\begin{aligned} \beta l &= \frac{\pi}{2\kappa} \\ l &= \frac{\pi\lambda}{2\pi 2\kappa} = \frac{1}{4 f_{max}\kappa}. \end{aligned} \tag{3.23}$$

This step concludes the circuit synthesis, however, the resulting characteristic impedances and line lengths must be realized in a way that can be manufactured. The following section 3.5 explains how this can be addressed using computer aided circuit design.

3.5 Measured Matching Networks

The SRFT delivers impedance values for each UE. These UEs can be realized using a multitude of transmission line technologies to build a commensurate transmission line filter. It is even possible to incorporate lumped elements into a transmission line design if necessary, as long as the result reproduces the desired circuit parameters. This process is iterative in itself as it requires a circuit design based on values gained from the theory which needs to be tested in order to see if it meets these expectations. By simulating the resulting matching network with the help of an EM-circuit simulator, this process can be sped up significantly. Figure 3.12 depicts the circuit parameters of the individual stages of the input and output matching network. For the realization of the matching network, a Rogers RO4003 substrate material with a board thickness of 0.508 mm, a copper thickness of 17 µm, and an $\varepsilon_r = 3.38$ was selected.

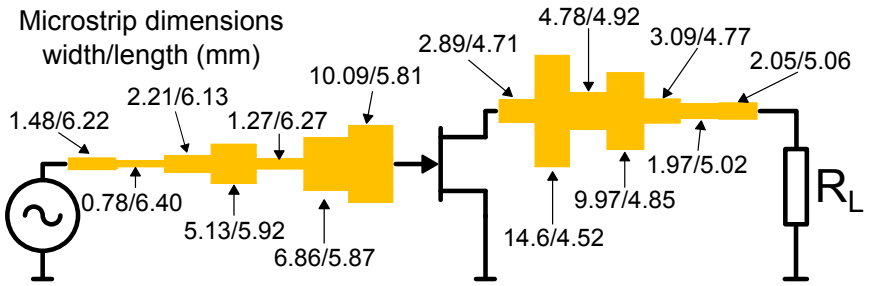


Figure 3.12: Amplifier layout without bias

The simulation environment used to conduct this step was *NI AWR Microwave Office* using a planar electromagnetic solver named *AXIEM*. Using a simulator to predict the behavior of the circuit before manufacturing is essential when designing filters like this, where a high dynamic range in characteristic impedances is required. When designing transmission line filters which vary much in thickness, the effective dielectric constant also changes significantly which prohibits the use of simple line models to predict their parameters. In extreme cases, they may even be overmoded, which means higher order modes can be excited within segments. Even, though this can be predicted using EM-simulators it is not advisable to utilize this effect in designs. In such cases it is better to switch to another type of circuit element to realize this section.

As depicted in Figure 3.13, the input matching network was realized using solely microstrip lines. However, for the output matching network, this was not possible.

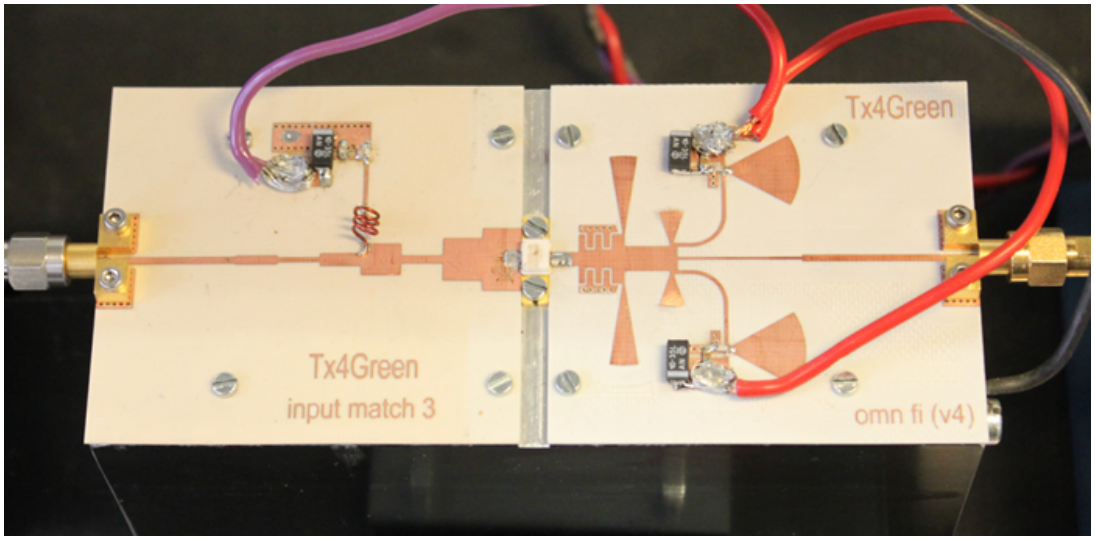


Figure 3.13: Constructed PA

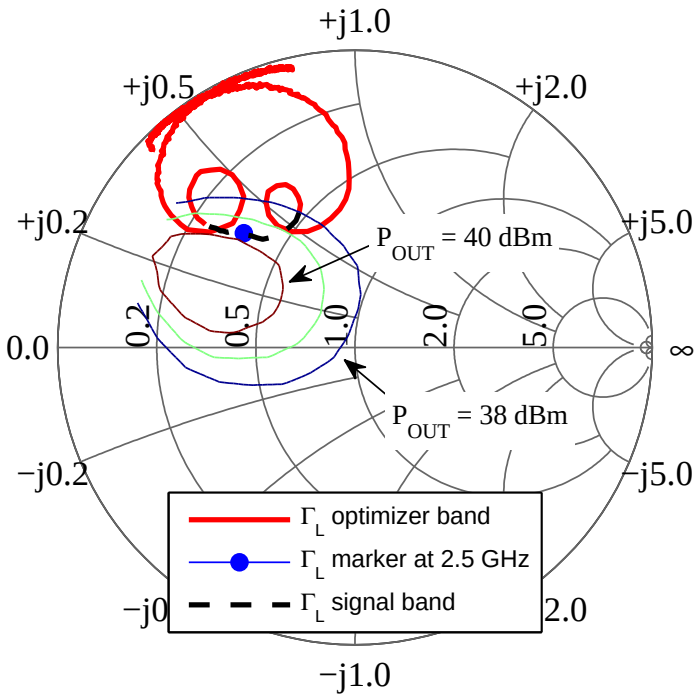


Figure 3.14: Measured locus of the output matching network including the output filter with the corresponding load pull contour for output power at 2.5 GHz

Due to the required impedance values, a hybrid design with inter-digital structures and radial stubs was used to realize the circuit. Figure 3.15 depicts a simulation of the PA's output matching network with the overlayed current density. The color scheme illustrates the current density on an uncalibrated logarithmic scale. It is important to note that the directivity of the coupler as well as the RF-rejection of the bias circuitry work well in these simulations.

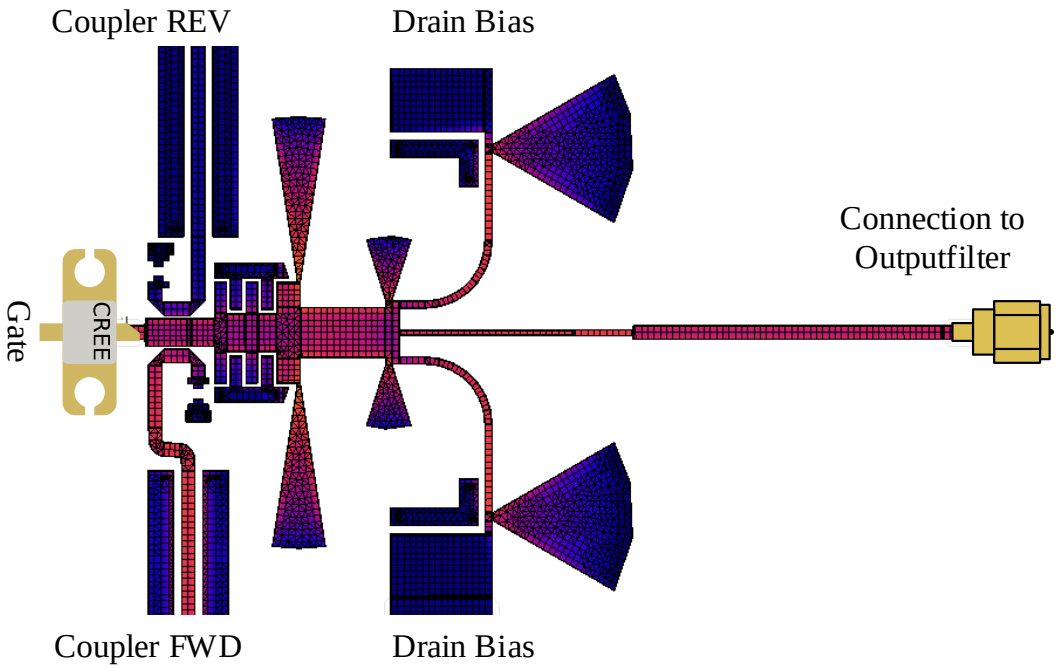


Figure 3.15: Logarithmic current density (blue low, red high) at 2.5 GHz in the PA's output matching network.

Measured points of Γ_L in the optimizer and signal bands (see eq. (3.2)) with respect to the measured load pull contours at 2.5 GHz are depicted in Figure 3.14.

3.6 Design of the Output Coupler Section

A DRFTx requires a highly selective output filter and a highly reflective matching network in order to function. Per design, this contraption rejects the majority of the signal components present at the active device, namely all frequency components that are exclusively occupied by the modulator noise (see eq. (3.2)). However, this also

prohibits any meaningful time domain measurements of the active device's waveform since the narrow relative bandwidth makes many frequency components unobservable through the system's output. Especially when using very broad band instruments which have inherently lower dynamic range. To overcome this issue, a coupler is integrated into the matching network between the transistor and the filter to allow direct access to the desired signals.

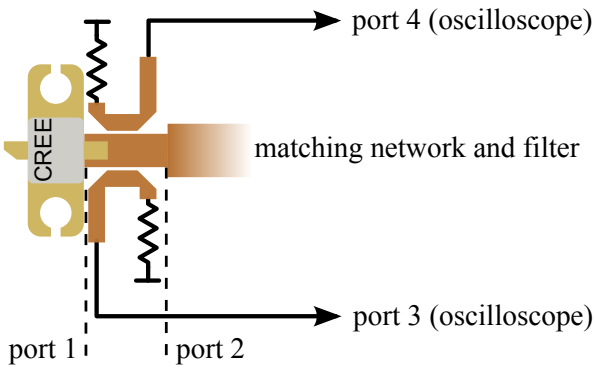


Figure 3.16: Test PCB for the directional coupler used in the PA

A directional coupler was chosen to tap into the circuit. In order to be compatible with the matching network a microstrip line based topology was used. The bandwidth of the coupler also needs to cover all relevant design frequencies ranging from about 500 MHz up to 20 GHz. To achieve a reasonable high bandwidth, the coupled length was chosen significantly below quarter wave resonance. Hence, the coupled length becomes small which lowers both coupling factor and directivity. While the low coupling factor is beneficial for the intended application, the limited directivity is the major design trade-off.

Figure 3.16 depicts a test PCB containing only the utilized coupler. To increase the isolation of the coupled ports, a dual directional coupler design was chosen. This improves the potential quality of a coupler calibration, and, therefore, the resulting measurements. The realized directional coupler is depicted in Figure 3.17 and the measured results can be found in Figure 3.18.

These measurements indicate that insertion loss and bandwidth are satisfactory for the task at hand. However, in order to apply common deembedding techniques to directly measure current- and voltage- waveforms at the intrinsic transistor an additional

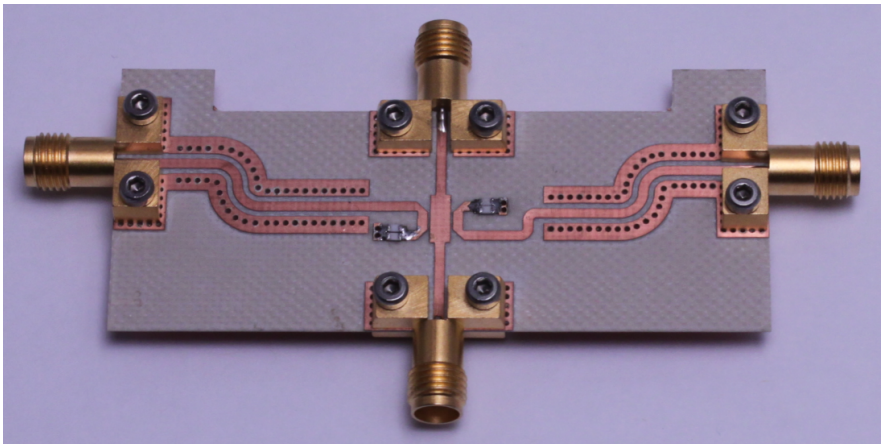


Figure 3.17: Realized directional coupler

calibration step is required which was already published in [28]. However, the most important aspects are summarized in this section.

While the imperfection of the coupler alone can be easily identified, and potentially corrected, based on the measured S-parameters. It was found that having a joint correction, that includes the imperfections of the attached sampling oscilloscope, is quicker and yields better results. Since an oscilloscope is a passive instrument a SOL (short, open and load) calibration or similar methods are not feasible. Additionally all calibration methods that are based on harmonic excitation require many measurement runs in order to cover the coupler's bandwidth. Furthermore, the directional coupler's coupling factor is very small at lower frequencies. Hence, the dynamic range of the sampling oscilloscope is not sufficient to measure the signal accurately. Therefore, a two-stage calibration procedure was chosen. Firstly, a calibration which allows to eliminate the imperfections of the directional coupler, cables, mismatch of the instrument, and so on has to be conducted. These values can be computed from the available S-parameters of the coupler and the oscilloscope's inputs.

Secondly, a receiver calibration has to be carried out on top of that to eliminate the influences of the sampling oscilloscope's detectors. The contributing factors are mainly gain imbalance and timing skew between the channels. To determine these systematic errors, induced by the sampling oscilloscope, the waveform of a pre-calibrated harmonic phase reference (HPR) was recorded. A block diagram of the measurement setup is depicted in Figure 3.19.

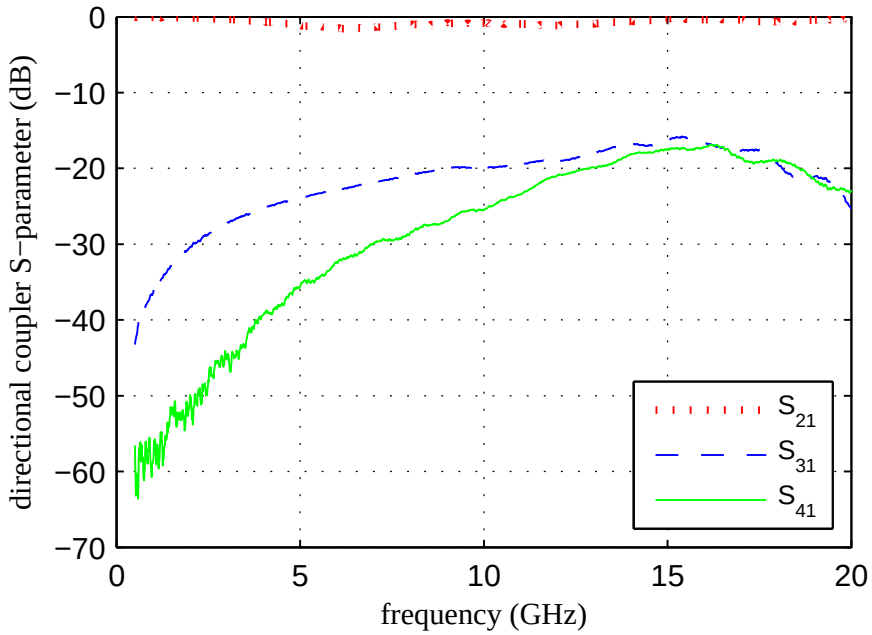


Figure 3.18: Measured directional coupler

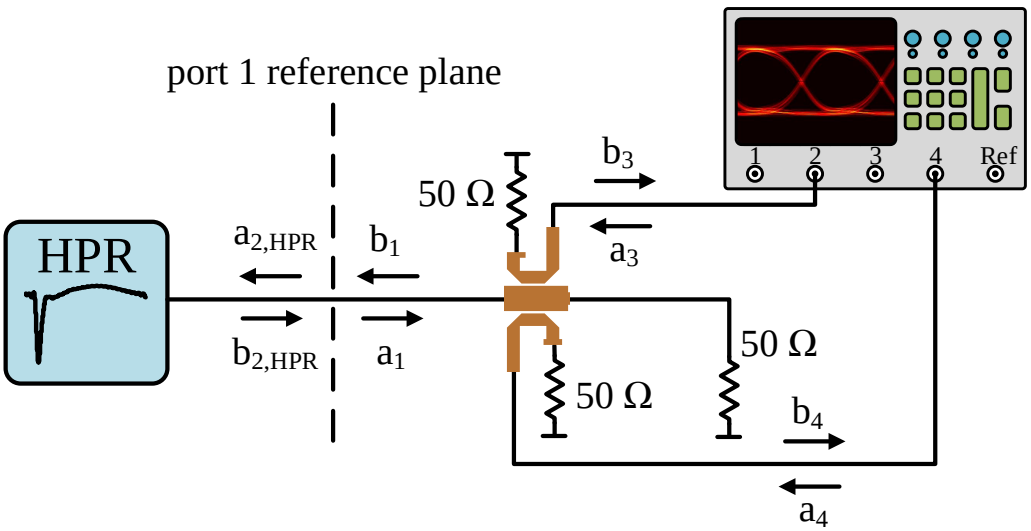


Figure 3.19: Sampling oscilloscope calibration block diagram

In Figure 3.20 the output voltage wave of the HPR ($b_{2,\mathrm{HPR}}$) is compared to the calculated input voltage wave at port 1 (a_1) resulting from the sampling oscilloscope measurement. Both waveforms show excellent alignment indicating a valid correction or calibration technique. Since this measurement system is targeted towards time domain measurements, dispersion and channel-skew are much more problematic than high dynamic range. Therefore, a time domain verification was preferred over a lengthy frequency domain method based on harmonic excitation.

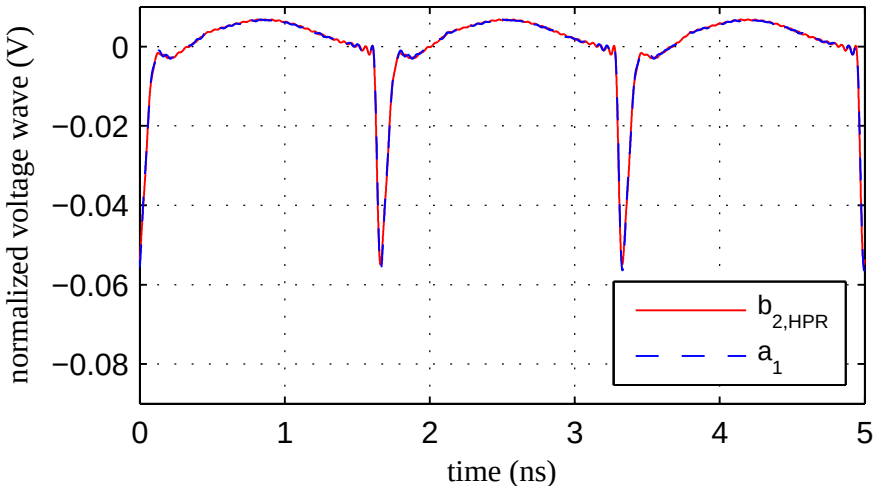


Figure 3.20: Output of HPR compared to measured results with applied calibration

CHAPTER 4

NONLINEAR SYSTEM CHARACTERIZATION AND MODELING

Recent developments in the field of wireless communications emphasize the urgent need for flexible radio frequency (RF) front ends. Not only a multitude of frequency bands needs to be supported, also modulation formats come in a greater variety and tend to develop at a faster pace. While receiver topologies have been adapted by employing software defined radio (SDR) concepts, translating these developments to transmitters was hindered, mostly due to the lack of power amplifiers that are energy efficient, linear, and broadband at the same time. All digital radios are a promising, however, not a very mature option to overcome this issue. While offering the desired flexibility and potential energy efficient operation, these setups exhibit relatively long memory effects and significant nonlinear distortion.

This chapter explains why a DRFTx design doesn't lend itself well to a modularization of the design process and how this can be addressed. Many design choices require modulator- and circuit-co-design to provide good overall performance. There are two main reasons for the dependencies between the different functional units: Firstly, DRFTx systems utilize lengthy binary excitation sequences to drive the attached power amplifier. These sequences are generated by the DRFTx system's modulator in order to generate the desired RF signal. Typically, those modulators compose their output

signal from a predefined set of modulator symbols, although the algorithms selecting these may be different. In any case these modulators have a rather long memory depth which results in a great variety of output sub-sequences that follow no simple pattern. The composition of the modulator's output signal for a given baseband signal is described by the umbrella term modulator statistic. As a consequence, the length of individual patterns in the modulator's output depends significantly on the settings of the modulator. This can be also described in terms of typical pattern durations.

Secondly, only a portion of the modulator's output signal resembles the desired RF-signal while most of the signal's overall energy constitutes of modulator-noise. Both components are occupying distinct frequency bands and for the utilized modulator [4] the power ratio between the desired signal components and the noise components is typically 2-5%. To suppress the modulator noise a highly selective reconstruction filter is required, which also needs to be reflective in order to maintain energy efficient operation [11]. Therefore, the majority of the PA's output signal is reflected by the reconstruction filter and causes persistent load modulation of the active device. The duration of these transient effects is mainly determined by the PA's output section.

If the typical durations of these effects are significantly different they can be easily separated as one can be considered pseudo-stationary with respect to the other. However, measurements have indicated [16] that this is not possible for the presented setup. In contrast to other more common PA designs, the PA circuitry in DRFTx systems rarely reach an equilibrium state of operation which entangles both memory mechanisms.

Having a nonlinear circuit model at hand, that allows to freely parameterize every aspect of the system, would make "brute force" optimization possible. However, even with modern computer aided circuit design software this would be rather inefficient. One way to overcome this issue is to use different characterization methods that allow comparing functional units based on their impact on performance of the overall system.

Especially in an early design phase, where many aspects of the DRFTx circuit are still in question, characterizing specific aspects is more practical than a modeling approach. In contrast to models characterization methods allow to compare how the performance of the overall system is affected by a specific piece of circuit design or operational parameters. Section 4.1 explains how memory affected nonlinear distortion, introduced by specific circuit modules, can be gauged in frequency- and time-domain.

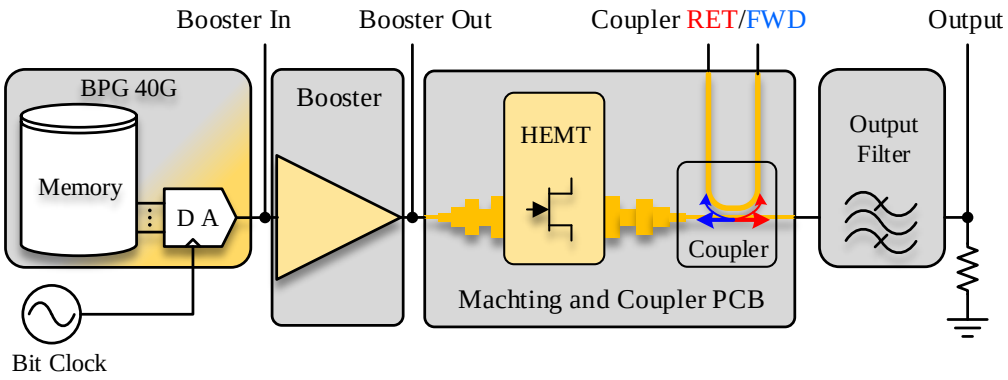


Figure 4.1: Typical signals in a digital RF transmitter (DRFTx) and the scope of different model types

Ultimately the DRFTx system needs to be modeled, however, in order to optimize the design. Key to a modulator and circuit co-design is a model representing the nonlinear memory affected behavior of a DRFTx and predicting design relevant parameters, like power efficiency. By using such a model, it is furthermore possible to find an optimal trade off between modulator complexity, nonlinear distortion, circuit complexity, and power consumption of a DRFTx already during the design phase. Consequently, the model needs to be computationally efficient to cope with the long excitation sequences required to capture the statistics of real world signals. Section 4.2 provides an overview about different model types and explains which ones are suitable for the task at hand. Finally, chapters 5 and 6 provide a detailed analysis for the models that have been investigated in this work.

4.1 Characterizing Nonlinear Memory Affected Distortion

All binary modulators that generate an analog signal do that by oversampling and subsequent filtering. The filtering separates the desired signal from undesired components which are often referred to as modulator noise. Gauging the linearity of such a system is problematic, since the device is inherently strongly nonlinear, but, due to the nature of the signal, it is not affected by static nonlinearity. However, as described in [18], [38], and [19] memory effects in conjunction with nonlinear behavior can give

rise to different undesired signal components. Depending on the context, there are different names for this distortion effect. In communications nonlinear inter symbol interference (nISI) is common place while circuit modeling describes this as memory induced, nonlinear behavior, which is the more general term and is also used in this work.

The presence of distortion due to this effect is generally undesirable because it reduces the spurious free dynamic range (SFDR) of the overall system. As described in section 2.1 and 2.2, these effects can often be compensated on a circuit level. However, in order to find a setting that results in a maximum SFDR, it is important to quantify the effect.

Frequency Domain Characterization

The most straight forward method to quantify these effects is to measure these undesired components directly in the frequency domain. Although in order to detect undesired signal components, they need to be separated from the desired signal. The simplest way to do that is to use the PWM-modulator to generate a signal that contains no signal components within its modulation bandwidth [18]. An example spectrum is illustrated in Figure 4.2. By using a spectrum analyzer, the output signal source is analyzed for the presents of unwanted spurs.

However, this method has some shortcomings in practical applications. When cascading multiple circuit sections, an absolute distortion power density is of little use comparing the contribution of different sections. Hence, measuring SFDR is more universal since these values can be compared from stage to stage. To compute a SFDR, the noise level within the band of interest needs to be compared to another well controlled signal component.

By deliberately adding static tones within the signal band of the modulator, this can be achieved. Since the frequency response of the modulator ought to be flat around these frequencies, it detaches the result from the noise shaping properties of the modulator. It also allows to gain measurements of sections which require a filter to operate, like the output filter in the PA stage (see section 2.3).

Figure 4.3 illustrates the test result gained from the presented test setup. However, the two test tones could mask some other higher order components why this test has

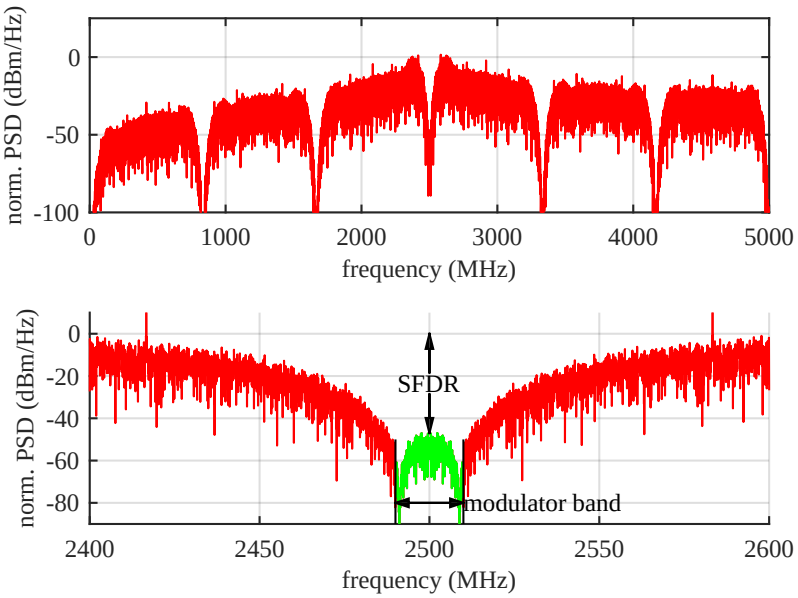


Figure 4.2: Nonlinear characterization test spectrum

to be conducted at least twice with different tone spacings to rule out the presence of such components.

This method has been used throughout this work, however, it requires a PWM-modulator that can provide a higher SFDR than the DUT and a source that allows to generate an excitation for which the same is true. As in 2.1, where the signal source itself is the DUT, conducting this experiment is not possible and other methods have to be used.

Time Domain Methods

A simple yet effective method to gauge whether memory induced, nonlinear behavior is present in a binary system is based on the DUT's eye diagram. The DUT is excited with a pseudo random binary signal and its output is recorded in time domain. In a next step, the signal is segmented into sections covering three adjacent bits. All of these sections are overlayed in a plot to form an eye diagram, as exemplified in Figure 4.4. The pseudo random bit sequence needs to be long enough to capture the memory depth in question and must asymptotically contain the negated version of any subsequence present within the sequence. M-sequences for example satisfy this

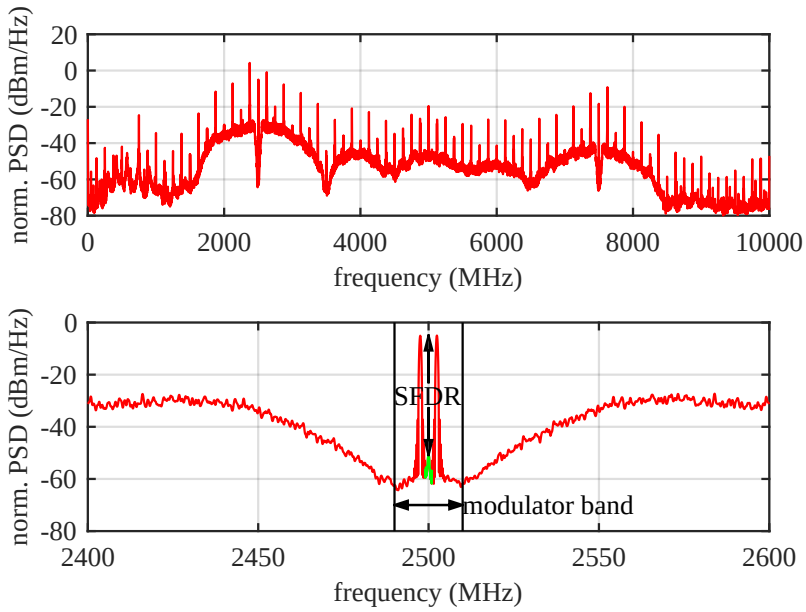
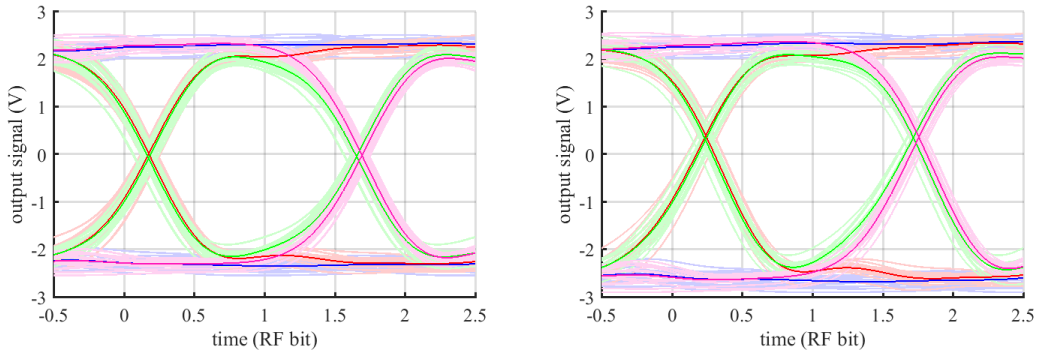


Figure 4.3: Signal source test spectrum

condition [39]. In a next step the averages of all 2^3 trace sets are computed to gain unified representations for each set. An average of the overall set is computed and subtracted from all the unified representations to remove any bias. Now the unified representation is subtracted from its negated version and the difference is the resulting error signal. The commutative error energy is normalized to gain a universal measure. This process is depicted in Figure 4.5 for the set $[101]$ for a suboptimal setting (a) and an optimized setting (b).

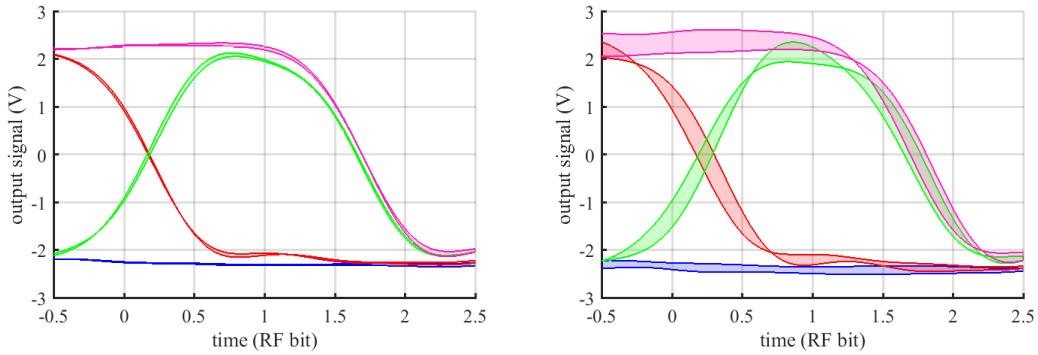
This relation was published in [19] and [20] which presents also other options to compute nonlinearity measures from a measured eye-diagram. This measure allows to test almost any DUT in a straightforward manner and is in fact a special case for the waveform extraction of the LUT modeling method presented in chapter 6. For practical applications this method is sub optimal for several reasons: Firstly, circuits that have a very dynamic frequency response, the synchronization required to derive the eye-diagram gets problematic. Secondly, it offers only a scalar overall error measure without gaining information about memory depth. Finally, it is not constructive in order to co-design modulator algorithms and the circuits that go with them.

It could be improved by utilizing more sophisticated subsequences than the ones defined inherently by the eye diagram. However, this would result in very many sub-



(a) suboptimal setting (b) optimized setting

Figure 4.4: Eye diagram based test for nonlinear memory effects



(a) suboptimal setting (b) optimized setting

Figure 4.5: Pulse-symmetry test for nonlinear memory effects

sequences that need to be tested in order to gain a complete assessment of a DUT. The LUT model reduces the complexity by utilizing the limited signal space of the modulator and provides also a model based on its alphabet. Therefore, the LUT approach is superior if a modulator is available. Very early in the design, this is a quick and valuable tool for benchmarking circuits.

4.2 Modeling Requirements for DRFTx

At this point it is important to note the entanglement between modulator algorithm design, distortion management, and circuit driven efficiency engineering. While these topics are usually addressed separately in literature, a joint approach is required in this

case. To optimize the circuit's overall efficiency for example, requires to predict the waveforms at the active device. These waveforms are partly caused by the excitation from the PWM modulator's algorithm and partly from the memory effects present in the circuit. Due to the highly nonlinear nature of that problem, these effects cannot be separated. Therefore, a nonlinear memory afflicted model of the circuit, covering its entire bandwidth, is required to perform joint optimization.

This approach is different from many other nonlinear modeling approaches discussed in literature since it is circuit centric and does not limit itself to predicting nonlinear distortion within the communication band of interest. While there are plenty of model types to choose from, when it comes to modeling nonlinear systems [40], many of them have significant complexity issues if the system shows long memory and is very broadband at the same time [41]. Since this project requires a model for the lab setup presented in chapter 2, a measurement based black-box modeling approach was pursued, although a circuit simulator based option was investigated too.

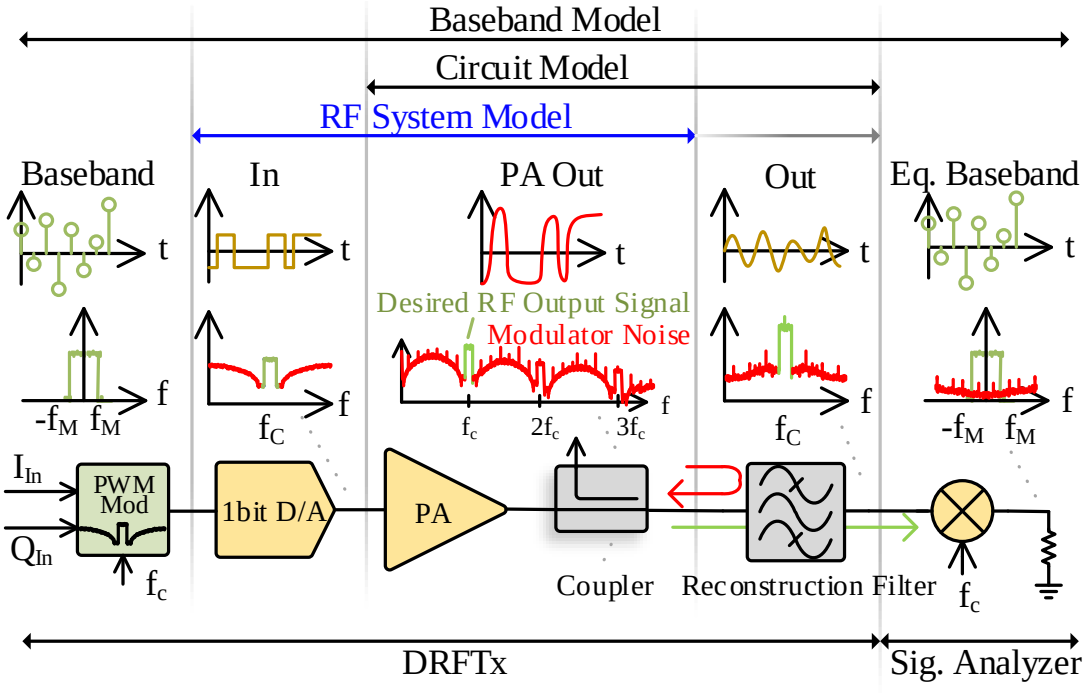


Figure 4.6: Typical signals in a digital RF transmitter (DRFTx) and the scope of different model types.

To emphasize that different modeling approaches describe different sections within

a RF-system, this relation is depicted in Figure 4.6. While baseband models provide an end-to-end description of a system, they operate at a moderate bandwidth dictated by the system's bandwidth and the observed degree of nonlinearity. Equivalent base-band models are not suitable for circuit-centric optimization of a DRFTx modulator and circuit designs since they neglect a majority of the overall signal. For the presented case 95%-98% of the PA output signal's energy is outside the equivalent baseband. This needs to be considered in order to analyze whether a circuit shows the desired behavior, like energy efficient operation. Nevertheless, [42] describes an interesting case since it shares the same hardware, although, this model was designed solely to predict the in-band distortion within the bandwidth of the modulator. Therefore, it features a very limited bandwidth in which it achieves a lower model error of about 10 dB better with respect to the mean NMSE. The modulator and different base band models are extensively discussed in [14].

Nevertheless, modeling methods that are associated with this class of models, like Wiener-Hammererstein or Volterra-Models, are also applicable for circuits. These circuit models are very universal since they can be excited by any time-domain waveform. However, as described in [41], they lend themselves not very well in applications where long memory is required. Mostly because the number of parameters that need to be identified grows significantly with memory depth [41] [40]. On top of the growing complexity, the stability of these models is another issue that get more problematic with deeper memory. Due to their nonlinear nature, the model's stability can be analyzed often only empirically. Chapter 5 discusses the performance that can be achieved using a nonlinear auto regressive model (ARX) to describe a DRFTx circuit. It was found that nonlinear models scale dramatically with the nonlinear properties of a circuit. On top of that, the literature does not provide comparable results. Therefore, describing the presented DRFTx using such a model was done primarily to have a performance- and complexity-benchmark using a well established model. These results can be directly compared with other approaches.

For this work, however, a measurement based RF system model is introduced that reduces the model's complexity by exploiting the modulator's limited signal space [16]. The excitation is completely defined by base-band symbols provided by the modulator. This is similar to base-band models in that it drastically limits the variety of possible excitation signals. However, this model captures the nonlinear behavior

of the circuit by tabulating typical versions of output waveforms for whatever bandwidth required to solve the design problem at hand. This can be done with moderate memory demand for all possible combinations of input symbols within the expected memory depth of the circuit. Such an approach avoids the need to drive the model from time continuous waveforms. It can be directly derived from measurements of the desired output waveform and the input symbol sequence. While related methods are used in control theory [43], they are not directly applicable for this kind of DRFTx circuitry. Chapter 6 introduces this modeling approach based on measurements gained from the DRFTx lab setup. It describes how such a model can be generated from measurements and compares the results with verification sequences that resemble a 20 MHz LTE frame.

There are also some hybrid approaches that reduce the complexity by employing different modeling approaches side by side and combining the individual outputs [44]. They are most useful if circuits show effects that can be explained by different mechanisms operating independently of each other. In this case each mechanism can be modeled by a different sub-model. A typical example would be a slow self heating model with deep memory alongside a fast EM-model with little memory depth. This work aims at an instance where a separation is not possible, therefore, hybrid approaches have not been investigated in this work. Although, the models in this work could easily be expanded to make use of such a hybrid approach.

4.3 Model Estimation and Verification

In order to generate a model of a circuit, there are some steps that need to be taken:

Firstly, all models require some modeling parameters to set. Usually the complexity of the model is defined by these settings and in many cases this is a step that requires some guesswork and tuning. A typical example would be the memory depth of the circuit or the type of nonlinearity used to model the observed behavior. The model's complexity needs to be chosen suitably for the problem at hand, however, the model's complexity needs to stay manageable. Therefore, it is essential to find the shortest but still feasible parameter set. Since this work assumes a black box scenario, this is usually accomplished in a rather straightforward manner by estimating models of steadily increasing complexity and testing their modeling error.

In order to gauge the prediction error of a model there are some commonly used figures of merit. The ones used in this work are listed in eq. (4.1). To ensure that these values allow a fair comparison between different systems, models, model orders, or amplifier settings, the absolute error's energy is normalized by the predicted waveform's energy.

$$\begin{aligned} model\,fit &= (1 - \frac{\|Y - \hat{Y}\|_2}{\|Y\|_2}) \cdot 100\% \\ RMSE &= \frac{\|Y - \hat{Y}\|_2}{\|Y\|_2} \\ MSE &= \frac{(Y - \hat{Y})^2}{Y^2} \end{aligned} \tag{4.1}$$

Secondly, after the model's parameters are set the model's coefficients need to be estimated. This step is often called model training and in this step an algorithm finds the set of coefficients that result in the smallest prediction error. The model is trained using measured results from the test setup. However, training a model based on the system's target signal may be inefficient or even impossible. Therefore, often specialized sequences which fulfill some properties specific to the model type used are required. In any case the training step will return a prediction error for the training sequence. Although, this number is strictly speaking a prediction error, it is called training error since the model is per definition optimized for this signal.

Finally, the model's performance needs to be verified for a broader range of potential target signals in order to ensure stable performance. This step is called model verification. Since this step is usually computationally less expensive than the model training, longer and more complicated signals can be used. Since all models in this work used dedicated training sequences, the system's target signal namely a 20 MHz LTE signal was used to perform this step. The resulting error is called model error and was used to compare the performance of different model types to each other.

Analyzing the model's prediction error as function of different model parameters allows to estimate important circuit parameters. For example, the effective memory length of nonlinear circuits is not easily determined since methods like impulse response are not applicable. Nonlinear models on the other hand allow to test for spe-

cific dependencies. For example, how much the prediction improves by tuning the models memory depth, which would point to residual unexplained memory effects in the measurement.

Another approach to study the suitability of a model type for a given application is to build a model for different test scenarios and compare the results. This setup lends itself very well to this approach, since there are various signals that can be observed throughout the signal chain as depicted in Figure 4.1. Depending on which section of the DRFTx circuit is going to be described using nonlinear model, the effective memory depth observed will vary significantly. As depicted in Figure 2.13, there are are several sections that can be described using nonlinear models. Two of them have been found very illustrative for the capabilities of the individual models:

Firstly, a "booster" model describing the DRFTx from the digital input up to the booster stage's output. This sections shows significant nonlinear behavior, however, it requires only a relatively short effective memory depth. This section is well suited to test models for their maximum achievable performance since it is not necessary to truncate residual memory effects or having to deal with very large model parameter sets.

Secondly, the "full" model that describes the forward traveling waveform at the PA's output based on the excitation sequence. Doing so requires to capture relatively long memory effects, that are caused by the long time constants induced by the output filter. This section is the practical relevant section, since it allows to explain many nonlinear circuit effects. However, the relatively long nonlinear memory of this section poses a challenge for model estimation.

Estimating both of these sections using the same modeling approach and comparing them is very interesting since it uses the same hardware to gain measurements. As a consequence, it is possible to benchmark the performance of a model and test its performance for DUTs with long effective memory. The results are presented in sections 5.3, 6.5, and finally in chapter 7.

Thirdly, it is also possible to build a model that describes the output waveform based on the input sequence. Although, this description is, counter intuitively, inferior to the "full" model. On one hand it suffers from the relatively long memory length and strong nonlinearity like the "full" model, but it cannot be used to explain important performance measures, like power efficiency, since the waveform at the active device

cannot be calculated from the output signal. The narrow bandwidth of the output filter is obscuring many energy rich signal components that are relevant to the functionality of the circuit. However, if the time domain waveforms at the active circuit are known, the output waveform can be calculated in a straightforward manner since the output filter is linear. Therefore, this option has not been investigated since it is a special case of the "full" model.

CHAPTER 5

THE ARX MODEL

As discussed in chapter 4, having a nonlinear model for the circuitry in a DRFTx system makes it possible to analyze dependencies which are problematic to do solely using a hardware test setup. There are several types of models that can be used to describe the circuit, however, a time domain approach is the most suitable for the task at hand. Time domain models allow to analyze transient memory effects in a straight forward manner and can deal with the very broadband nature of the system.

Out of the group of time-domain models the Autoregressive Exogenous (ARX) model and consequently the Nonlinear Autoregressive Exogenous (NARX) model was chosen for detailed analysis. This model type is one of the most flexible concepts and it is well supported by software solutions like the MATLAB *system identification toolbox*. In fact all following calculations are based on MATLAB's *idnlarx* model. Deriving an ARX model is twofold. Firstly, the recorded output waveform is used to train a nonlinear ARX model. This process is well understood and even available as turnkey solution. Secondly, the model's predictions are compared to measured results gained from the test setup. The test signal in this case is a 20 MHz LTE signal generated by the PWM modulator.

5.1 Formal Definition of an ARX Model

The nonlinear ARX model identification process was investigated extensively in [20] and [45]. Therefore, the notation in this work is purposely chosen to be consistent. The

nonlinear version of the ARX model is an extension of the linear version, therefore, the simpler linear version is explained first.

The linear ARX model can be denoted as in eq. (5.1). Vectors X_n and Y_n denote the series of input and output samples at times index n. The A and B vectors correspond to the coefficients of a linear filter. These vectors are of length $na+1$ and $nb+1$ respectively. These terms are often reformulated as shown in eq. (5.2). Here the input and output samples are combined into a vector R forming a so called regressor, while the coefficients of the linear filter are combined in vector L.

$$\begin{aligned} X_n &= \begin{bmatrix} x_n & x_{n-1} & \cdots & x_{n-nb} \end{bmatrix}^T \\ Y_n &= \begin{bmatrix} y_n & y_{n-1} & \cdots & y_{n-na} \end{bmatrix}^T \\ A &= \begin{bmatrix} a_0 & a_1 & \cdots & a_{na} \end{bmatrix}^T, a_0 := 1 \\ B &= \begin{bmatrix} b_0 & b_1 & \cdots & b_{nb} \end{bmatrix}^T \\ A^T \cdot Y &= B^T \cdot X \end{aligned} \tag{5.1}$$

$$\begin{aligned} R_n &= \begin{bmatrix} y_{n-1} & \cdots & y_{n-na}, x_n & \cdots & x_{n-nb} \end{bmatrix}^T \\ L &= \begin{bmatrix} -\begin{bmatrix} a_1 & \cdots & a_{na} \end{bmatrix}, & \begin{bmatrix} b_0 & b_1 & \cdots & b_{nb} \end{bmatrix} \end{bmatrix}^T \\ y_n &= L^T \cdot R_n \end{aligned} \tag{5.2}$$

$$\begin{aligned} y_n &= L^T \cdot R_n + g(R'_n) \\ R'_n &= \begin{bmatrix} y_{n-1} & x_n & x_{n-1} & \cdots & x_{n-nc} \end{bmatrix}^T \end{aligned} \tag{5.3}$$

The linear version is extended to a nonlinear model by adding a term $g(R'_n)$ which provides the nonlinear component. The nonlinear term was chosen to be a *sigmoid network* of order *ns*, this type of nonlinear function is also part of the *MATLAB* framework. This nonlinearity is driven by the components of vector R'_n which is a subset of R_n. For this implementation R'_n was chosen to consist of the first $nc+2$ elements of

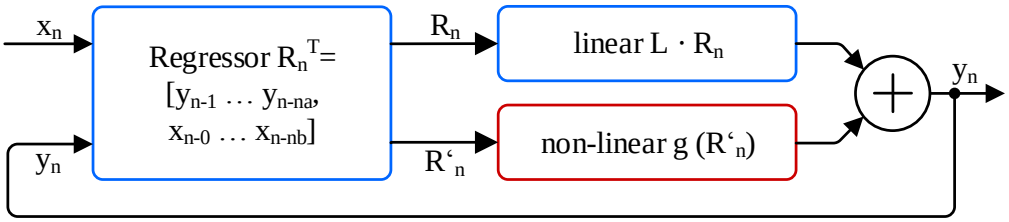


Figure 5.1: Block diagram of the used ARX model.

R_n. The structure of this model is also depicted in Figure 5.1.

5.2 Determining the Required ARX Model Complexity

As mentioned earlier the model's complexity needs to be chosen suitably for the problem at hand, which in this case is defined by the regressor length and the order of the nonlinearities.

To put this into perspective for the system presented in chapter 3, the model operates at an oversampling rate $r_{osr} = 4$ which results in an effective bandwidth BW_{model} of 20 GHz starting from DC. The value *nb*, defining the finite impulse response (FIR) component of the linear regressor, was varied from 1 to 284. The IIR component of the linear regressor is determined by *na* and was fixed to 4, these values in conjunction result in a linear filter component of up to 6 RF-bit.

The value *nc*, which determines the nonlinear regressor, can be varied from 1 up to *nb*. The IIR component of the linear regressor was intentionally excluded from the nonlinear regressor to improve the estimated model's stability. Otherwise longer models often show unstable behavior for long excitation sequences. The complexity of the nonlinearity is defined by *ns*. It specifies how many nonlinear basis functions are used to form the nonlinearity of the model. For the presented case *ns* was set to 5.

After the model parameters are set the model itself needs to be estimated. In this case the training step is optimizer based. In this step an algorithm tunes the model's coefficients in order to find a set that results in the least prediction error. In contrast to the linear model estimation, the nonlinear model estimation is based on a nonlinear optimizer. For all nonlinear model estimations carried out, the search method was

fixed to *lsqnonlin* which is a nonlinear least squares optimizer within MATLAB. Each model estimation is based on a figure of merit called *model fit*, as in Eq. 4.1.

Training an ARX model based on lengthy LTE modulator sequences showed to be inefficient. M-sequences have been found to be a more suitable excitation sequence for binary systems [39] since they provide long unique subsequences. To be able to capture the circuit's memory effects and to avoid conditioning problems during the model estimation procedure, the excitation sequence needs to be longer than the circuit's maximum effective memory depth. For this work sequences with lengths of roughly 1024 RF-bit have been used.

The model estimation process showed to produce more stable results when estimating a linear ARX model at first and finally expanding it to the nonlinear form. In this case the linear model is used as a start value for the nonlinear model optimization process. In this case the linear model determines the linear regressor R_n from which a subset is chosen to form the nonlinear regressor R'_n. The stability of this process tends to be problematic for higher complexities, however, this issue can be improved using a feature called *regularization* which introduces a penalty for model coefficients. Linear system theory offers methods to find for example optimal regularization parameters. For nonlinear systems this is not possible, however, the values used for the linear case have been found to be a good starting point also for the nonlinear case.

Experiments showed that nonlinear models sometimes show a lock-up behavior that is especially problematic when simulating long sequences. In these cases the nonlinear feedback in the model in conjunction with specific input sequences causes persistent oscillations. This phenomenon is problematic since the training step does not yield information if a given model is potentially unstable. Investigating the stability of NARX models is beyond the focus of this work and, therefore, has not been investigated. For the presented results this effect was mitigated by estimating a new model for the same regressor parameters if instability is detected during simulation. By tuning the regularization parameters, it was possible to gain stable models for the presented model complexities.

5.3 ARX Modeling Results

The results for these model estimations are depicted in Figure 5.2. In this depiction every pixel stands for a modeling result of a given complexity denoted in RF-symbols which is equivalent to 12 RF-bit. While this is somewhat arbitrary for ARX models which operate on a sample stream, it ensures consistency with the LUT-model results presented later in this work. Additionally, it also allows to compare the results to baseband models which use baseband-symbols as their native sampling rate. The model error shows that models with relatively short memory depth deliver up to about -32 dB of NMSE for the training sequence. These results are in good agreement with the results gained from the verification sequence depicted in Figure 5.3. Here a minimum model error of 35 dB is achieved, which computes to an SNR equivalent of 5.5 ENoB (Effictive Number of Bits) [46]. The oscilloscope can deliver around 6 ENoB, so this results can be seen as a benchmark of what is possible using this setup.

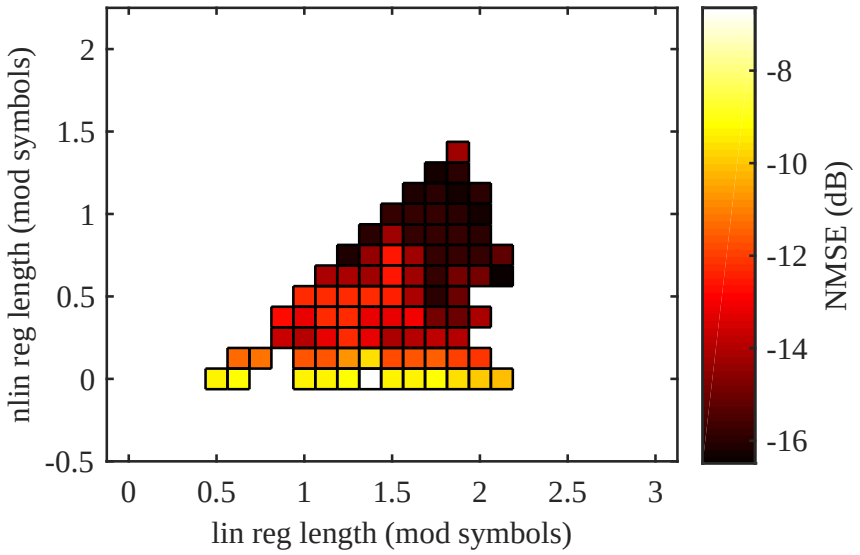


Figure 5.2: Booster model, NMSE error based on training sequence with respect to regressor length.

Observing a modeling error that is lower for verification sequence than for training is plausible since the M-sequence, used for the training, features a greater variety of unique subsequences than the verification sequence. In this time frame the nonlinear interaction in the model is also stronger and, therefore, causes higher error levels.

Models with a linear regressor longer than about 2 (modulator symbols) tend to show unstable behavior. In this case, deep memory effects in the circuit are negligible, however, the estimation algorithm tries to find model parameters that improve the overall result based on the measured waveforms. Although most of the unexplained contributions at this point are noise of the measurement setup. This leads to large coefficients in the model's parameter set which causes instability under certain excitation conditions. Gauging the stability of nonlinear models is a specialized field in its own right and was not further investigated in this work. In order to enhance stability it would be possible to optimize the model's *regularization* parameters, however, for the presented case the initial value was found to be close to optimum any way while having only little impact on the overall performance.

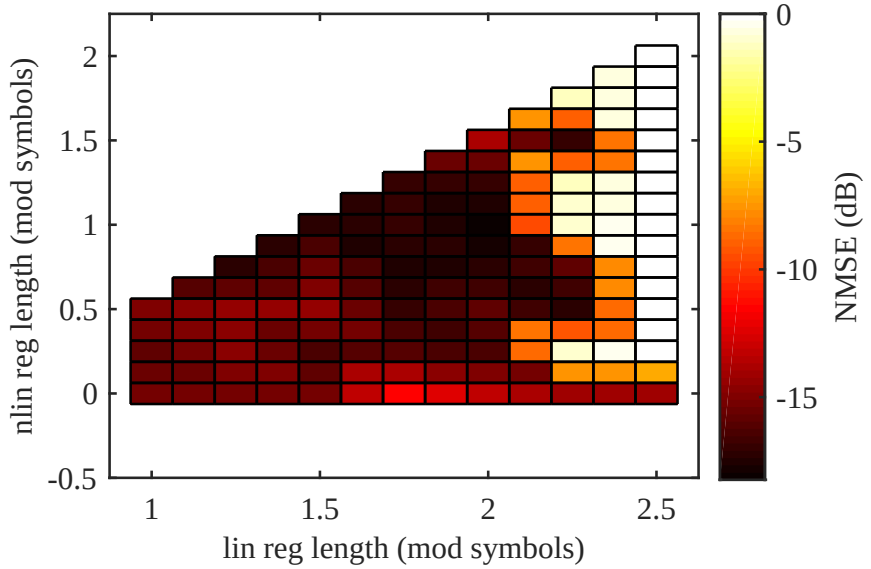


Figure 5.3: Booster model, NMSE error based on verification sequence with respect to regressor length.

When dealing with circuits that exhibit deep memory effects, M-sequences tend to overemphasize short memory effects over long ones because they do not reflect the burst-structure of the modulator's output signal. As a consequence the error measure gained from "full" model's training step (see Figure 5.4) does not agree with those gained from the verification sequence (see Figure 5.5).

These measurements allow to estimate models with deeper memory than for the

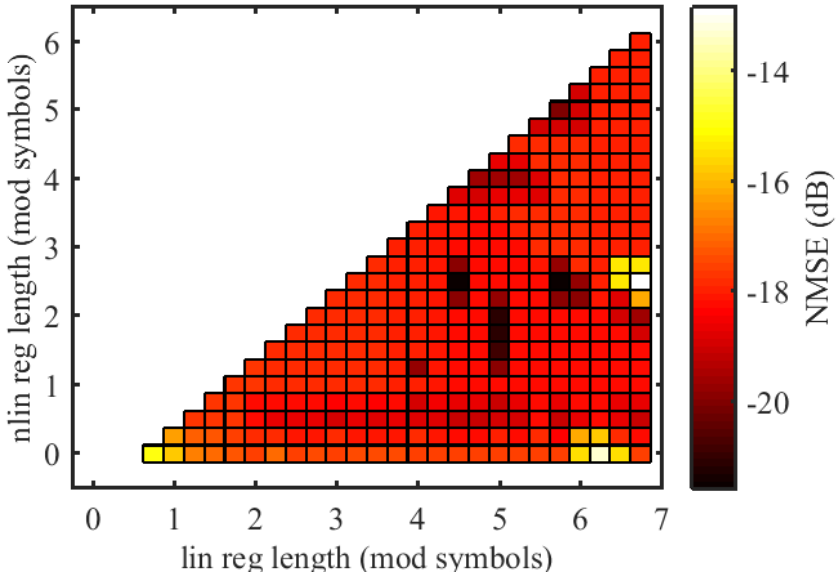


Figure 5.4: Full model, NMSE error based on training sequence with respect to regressor length.

"booster" model, however, the minimal achievable error is significantly reduced. This indicates that the nonlinear memory effects in the circuit are not captured to a sufficient extent. On the other hand, the complexity of the model cannot be expanded any more since its parameter space is already too large to be practical. Therefore, an alternative model type is proposed in this work that can make use of the limited signal space of the modulator in order to reduce the complexity of the model. This model is explained in chapter 6 and compared to this model in chapter 7.

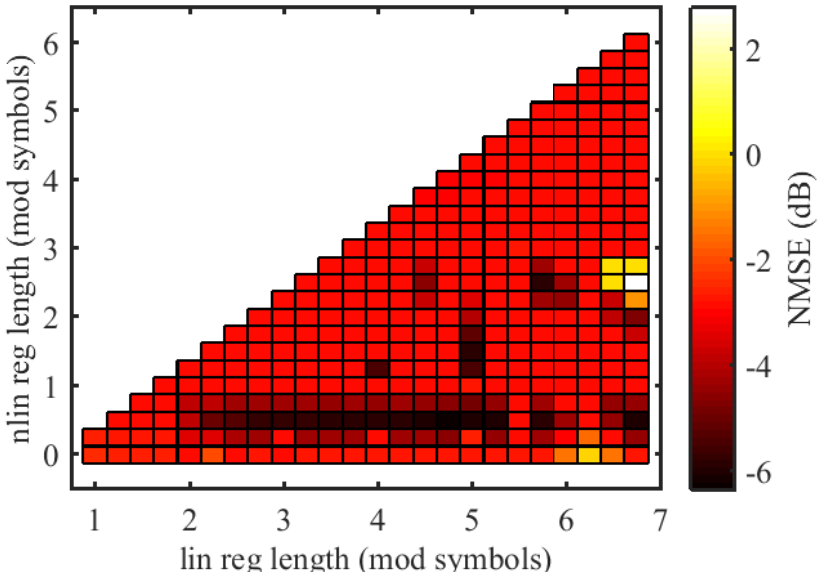


Figure 5.5: Full model, NMSE error based on verification sequence with respect to regressor length.

CHAPTER 6

LUT MODELS

In contrast to models that are based on describing systems using filters the idea behind the lookup table (LUT) model is to describe the circuit by tabulating all waveforms that a circuit can produce. This may look wasteful, however, when assuming that the memory effects in the circuit can be approximated as finite in time and the union of all excitation sequences forms a countable set, it is at least possible. Any LUT based model is based on the idea that a system is fully described by a limited set of parameters if it only depends on a limited set of inputs represented by a limited set of values. Otherwise the table would require an infinite number of entries. For the presented case, the signal space of the PWM-modulator that is used to drive the system is rather compact. This fact can be exploited to greatly reduce the complexity of the LUT model.

The system's outputs can be stored in individual tables, addressed by the evoking parameter combination. In this case, all output waveforms are indexed by a number of preceding excitation symbols, inherently modeling the circuit's memory. This clearly distinguishes the presented model from other LUT models in the literature, that describe the system as a memory affected channel [47] or as a set of system states [48].

For the task at hand tabulating waveforms is actually beneficial since these waveforms can be directly measured. Whereas finding an analytic description requires a parameter estimation step that can potentially result in an unstable description. In order to predict a DRFTx system's output waveform using a LUT model it is solely a matter of looking up waveform sections from the LUT and piecing them together.

Such operations can also be handled extremely efficient, especially, when compared to solving nonlinear equations for very many input samples.

The following sections elaborate on different aspects of this model type. Firstly, the nomenclature related to this model, which is strongly related to the modulator, is explained in section 6.1. Secondly, section 6.2 describes how the LUT waveform entries are generated from time domain measurements. Thirdly, section 6.5 discusses how the prediction quality of a LUT model can be quantified. This can be done in an entry by entry manner instead of an entire excitation sequence. As a consequence, the scalability of the model can be improved, as explained in the fourth subsection section 6.4. Finally, chapter 7 compares this modeling approach to the ARX model discussed earlier.

6.1 Formal Definition of LUT Model Parameters

Figure 6.1 depicts how the utilized modulator [4] composes an excitation signal from the input signal. In fact all these generated sequences are composed from a limited set of binary symbols.

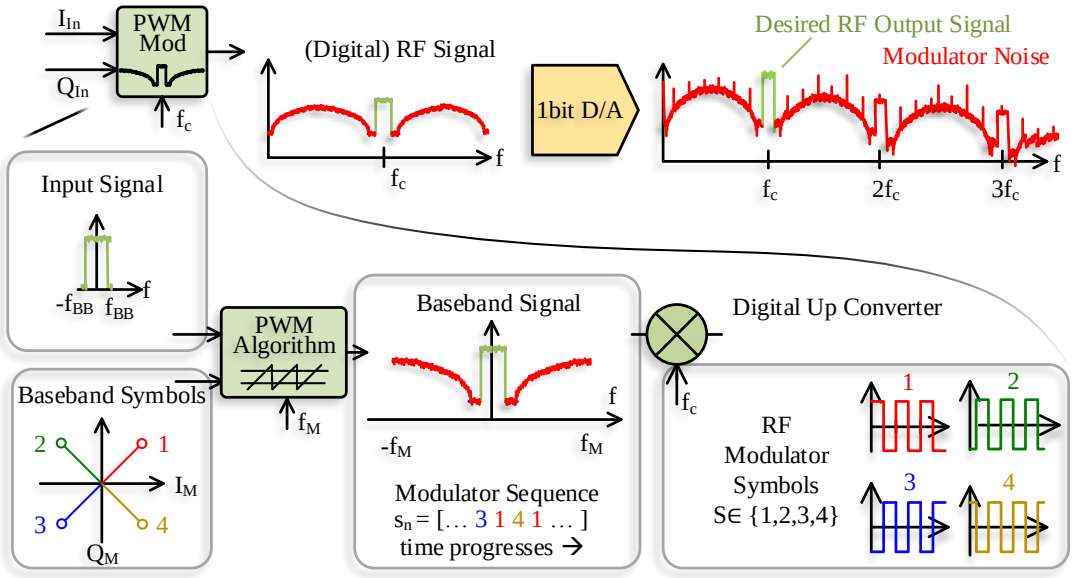


Figure 6.1: The inner workings of the PWM modulator.

The utilized modulator's set consists of $n_s = 4$ different symbols whereof each has a

fixed length of $l_s = 12$ bits. This simplifies the structure of the LUT since any possible input signal can be described equivalently by the sequence of baseband symbols s_n.

Since it is assumed that the circuit has a finite memory, only a certain number of n_t precursory baseband symbols will cause the currently observed sample at the circuit's output. Since there is no variety within a given input sequence for the duration of one modulator symbol, the same must hold true for the following output samples corresponding to the duration of the baseband symbol. To be directly comparable to the nonlinear ARX model, the model bandwidth has also been set to 20 GHz. This results in a waveform sampling rate of 40 GSPS or a sampling ratio r_s of 4 samples/RF bit. Therefore, one waveform entry in the LUT requires $l_q = r_s \cdot l_s \cdot n_s$ samples, which is considerably more than a baseband model that would require only a single coefficient. However, in contrast to the baseband model it captures the full bandwidth of the system.

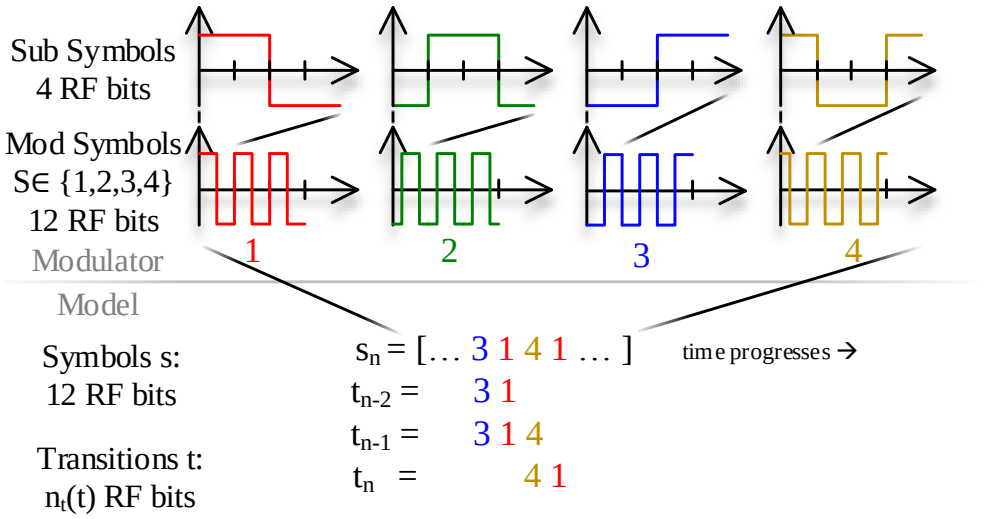


Figure 6.2: Dependencies between excitation signals, modulator symbols and transitions for hierarchical LUT models.

The groups of $n_t(t)$ input symbols indicating a waveform $w(t)$ stored in the LUT are called transitions t. Figure 6.2 depicts this consecutive grouping of symbols into transitions, which don't need to be of constant length as described in section 6.4. However, if they are the number of waveforms in the LUT is equal to $n_q = n_s^{n_t}$. This means that the number of parameters that need to be found in order to build this model is

$n_q \cdot l_q$ which is much more than for the presented ARX model. In this sense the LUT model requires a higher complexity compared to the ARX model, although, practically the LUT model generates these parameters directly from time-domain measurements by comparing waveforms. Such operations can be easily paralleled and are carried out highly efficient in modern computers, much in contrast to the nonlinear optimizer that is required to generate an ARX model. Additionally, the LUT model is unconditionally stable for any measurement derived from a stable circuit, which is an issue in analytic models like ARX models.

For a universal model it is necessary that it contains all possible transitions within the modulator's signal space. This is trivial for the constant transition length case $n_t(t) = n_t = const$, however, in the case of hierarchical LUTs a partitioning tree is used to ensure that T is complete. Since transitions of different length account for different sections withing the signal space they partition, the weighting function $a(t)$ is introduced to account for this fact in later calculations. All these relations are denoted in equation block 6.1.

$$\begin{aligned} S &= \{s_1, s_2, \ldots s_{n_s}\} \\ t &= s_n, s_{n-1}, \ldots, s_{n-n_t} \\ T &= \{t \,|\, |t| = n_t\} \\ 1 &= \sum_{t \in T} n_s^{-n_t(t)} \\ a(t) &= n_s^{-n_t(t)} \end{aligned} \tag{6.1}$$

6.2 Generation of LUT Model Entries

In order to generate a model for a given circuit it is required to build a LUT that provides an entry for any possible symbol combination. The simplest way is to use a LUT with constant length transitions $n_t = const$. However, as discussed in section 6.4, this is not optimal to gain a compact model and can be improved by using more sophisticated data structuring.

Nevertheless, in any case the LUT needs to be filled with waveforms to be used as a model. These waveforms are gained from measurements of a DRFTx lab setup using a training sequence that must contain all possible symbol combinations within

the expected memory length. In other words the training sequence needs to cover all of the modulator's signal space for the assumed memory depth. For LUTs of constant length ($n_t = const$), $n_q = n_s^{n_t}$ individual transitions are required to accomplish that.

The desired waveforms can be gained directly from the measurement by finding the corresponding transition in the excitation and extracting the corresponding waveform section from the measured waveform. This requires good synchronization between the measured channels as otherwise any delay would contribute to the memory demand of the model. Figure 6.3 depicts an exemplary result for one such a waveform segment.

For this kind of model, the training sequence must contain all transitions T at least once to populate all entries within the LUT. If the training sequence is built to contain exactly all transitions T of a memory depth n_t, there is only a single realization of the desired waveform. However, if the training sequence contains a superset of T this method allows to test whether the chosen memory depth is sufficient. In this case there are multiple realizations for the waveform corresponding to a given transition t which are contained in the set $W(t)$.

When extracting one LUT entry the set's average $\bar{W}(t)$ serves as the unified representation of the waveform. The set's standard deviation indicates how much all realizations, that fit the current transition, differ from each other. By definition, all memory effects spanning up to $n_t \cdot l_s$ RF-bits are captured by the extraction. Thus, the resulting errors are caused by noise in the measurement system, and more prominently, by residual uncaptured memory effects. The set's standard deviation corresponds to a root mean square error (RMSE) of the set with respect to the set's mean waveform. Since this is derived form a training sequence this value resembles a training error. However, for this type of model there is an individual training error for every transition t.

In order to be comparable with other systems, model orders, amplifier settings, and other transitions, the training error's energy is computed and normalized by the extracted waveform's energy. This is also depicted in Figure 6.3. For illustrative purposes the graph depicts a transition of order $n_t = 7$, although, only the section that is marked as "extracted transition" is utilized for the LUT entry as well as to compute the training error. It can be demonstrated that as long as there is some residual uncaptured memory the training error of LUT entries is steadily decaying [16] with its order n_t. So achieving a certain error limit is only a question of choosing n_t accordingly.

From a practical point of view, it is important to point out that by extracting more

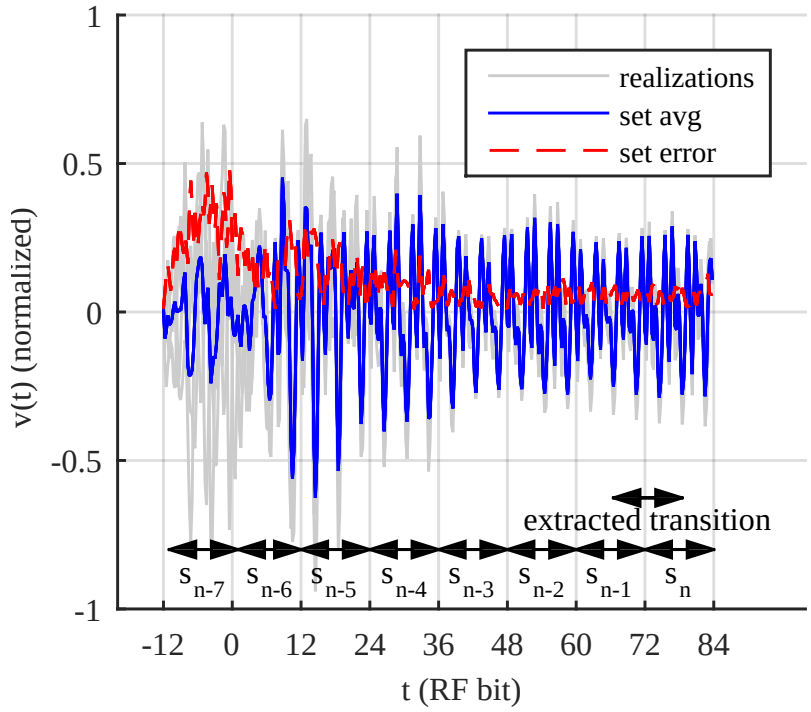


Figure 6.3: Waveform extraction

and more different waveform sets from the same measurement, the sets' cardinal number $|W(t)|$ needs to decrease accordingly. This can lead to a make-believe reduction of the extraction error since the deviation for every extracted waveform of the set is computed from lesser and lesser realizations as n_t increases. At some point they may not represent all possible variations due to uncaptured memory. This issue can be diminished by deriving a LUT of order $n_t + 1$ and using its LUT waveforms to compute the desired error measures for the LUT of order n_t. In this case all waveforms related to one LUT entry consist of exactly n_s realizations.

As a consequence, all n_s waveforms that make up a set $W(t)$ are caused by the same n_t symbols differing only in the $n_t + 1$th symbol. Doing so showed to produce more consistent error values, even when the chosen model order is clearly insufficient. Although, one has to keep in mind that this approach puts an emphasis on the variation caused by the $(n_t + 1)$th symbol.

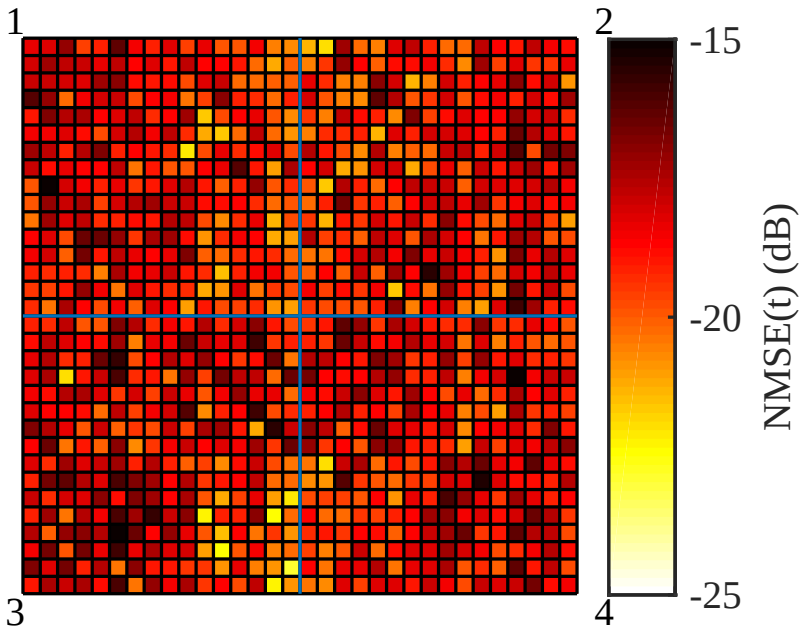


Figure 6.4: Normalized extraction error for a LUT derived from measurements of order $n_t = 5$.

6.3 Gauging LUT Models

A key observation for this class of circuits is that the training error varies significantly for different transitions, which has been described in [16] and [49] by measurements and simulations. Therefore, constant length LUTs are sub-optimal for robust prediction since their error varies significantly for different transitions. It indicates that some symbol combinations result in a longer lasting memory effect than others. An example for this behavior is visualized in Figure 6.4, where every pixel in the chart represents a LUT entry.

To depict transition specific parameters consistently over different LUT orders, all results are arranged in square groups as indicated by the bold lines. Each pixel that relates to a transition ending with the same symbol is located in the same quadrant. The numbers in the corners of the graph indicate the corresponding symbol number. For each of these subgroups, the same relation applies. Hence, the result for a specific transition can be found by iteratively dividing the chart into quadrants till a single pixel

remains. As a consequence, the outmost cornerstones of the graph correspond to the normalized extraction errors of the transitions containing n_t times the same symbol s. The outlining square indicates the modulator's entire signal space.

Having access to an error measure that is based on LUT entries instead of a single one for the whole model is useful to optimize the model. However, when comparing one model realization to another a cumulative model error is required. Although, it is important to accomplish this without anticipating a specific modulator statistic. By using a simple error measure like NMSE based on a verification sequence, the modulator's statistics of how often it picks a specific transition, may have a significant influence on the overall prediction error. As depicted in Figure 6.4, the individual entries result in significantly different errors.

The most general assumption for a modulator statistic is one that randomly picks symbols that are uniformly distributed over its signal space. In this case, the unified model error is the mean error from all entries of a LUT. This is applicable for the training error as well as for the verification error. For a constant length LUT this is trivial, however, for hierarchical LUTs the average must account for the individual length of the transitions. Longer transitions represent a smaller range within the modulator's signal-space than shorter ones. This is described by the weighting function $a(t)$ that weighs each LUT parameter according to the length of the transition t. These relations can be found in eq. (6.2).

$$\begin{aligned} a(t) &= n_s^{-n_t(t)} \\ 1 &= \sum_{t\in T} a(t) \\ nerr(t) &= \frac{\sum_{w\in W(t)} \|(\bar{W}(t) - w)\|}{n_w(t) \cdot \|\bar{W}(t)\|} \\ \overline{nerr} &= \sum_{t\in T} nerr(t) \cdot a(t) \\ \sigma_{nerr} &= \sqrt{\sum_{t\in T} (nerr(t) - \overline{nerr})^2 \cdot a(t)} \end{aligned} \tag{6.2}$$

To account for potential variations due to the modulator's statistic, the LUT's error is described with a first order statistic instead of stating only a single value per LUT. The results for different model complexities are presented in Figure 6.5. Each error bar depicts the achieved error of the LUT for a given order. The maximum, minimum,

confidence band, and mean extraction error of a set are indicated by markers. The significant difference between minimum and maximum extraction error indicates that the resulting modeling error varies greatly for different transitions. However, the mean training error steadily decreases with higher LUT order and the standard deviation stays about the same throughout for all LUT models.

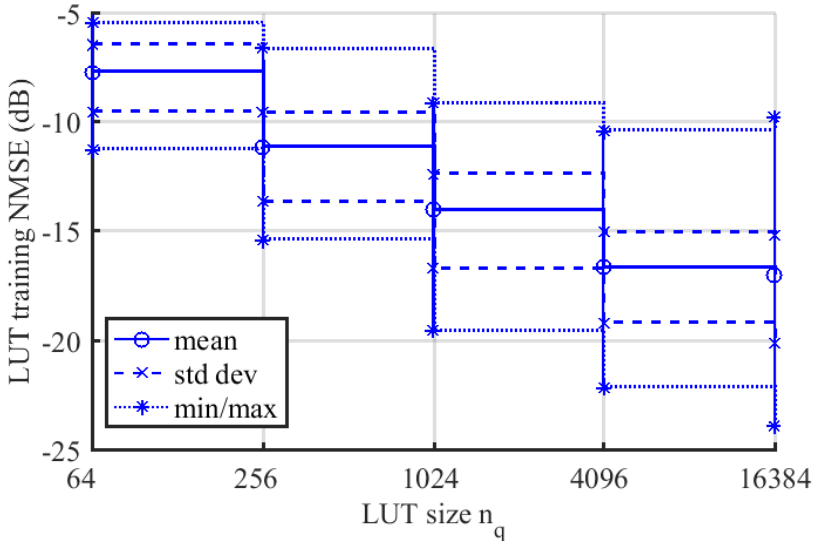


Figure 6.5: Normalized extraction error with respect LUT size n_q.

6.4 Hierarchical LUT Models

Generating a LUT model it is important to find the transition length $n_t(t)$ that is required to capture the memory effects present in the circuit. In contrast to many other model types, LUT models allow to do that entry by entry instead of specifying a universal memory depth for the model.

As demonstrated in section 6.2, the extraction error within a constant length LUT model may vary significantly. However, it is possible to exploit this property in order to gain more compact and more stable models. This is accomplished by introducing a more systematic approach to build, store, and interpret LUT models using trees. By introducing tree based LUTs it is possible to partition the modulator's signal space adaptively instead of uniformly. Ultimately, this results in a more scalable model that

can be generated from generic measurements without prior knowledge of the memory effects present in the circuit.

The tree consists of a root node symbolizing a transition of zero length and, therefore, covers the entire signal space. Every additional node is attached to the tree in a recursive manner. By augmenting the existing tree only at its leaf nodes with exactly n_s new leaf nodes this results in an exhaustive partitioning of the signal space. Any new branch subdivides the signal-space of the parent node in n_s new sections. This is equivalent to replace a transition of order n_t by n_s new transitions of order $n_t + 1$ that differ only in the s_{n-n_t+1}th symbol. This process is depicted in Figure 6.6 where the top white root node is branched to n_s new nodes at every layer. Each node stands for a different transition of a given order, corresponding to the layer that it is located in. Each leaf node corresponds to one LUT entry holding one waveform that is used for simulating the output of the system. So the LUT itself works the same like in the constant length case, however, the tree is needed to relate the transitions of individual lengths to the LUT entries.

The variable n_q is again used to denote the total number of LUT entries, however, here it corresponds to the number of leaf nodes. All transitions that address leaf nodes are contained in the set T. Therefore, the average extraction error for a hierarchical LUT computes according to the expressions in 6.2. Also the set's standard deviation needs to account for the length of the transition by applying the weighting function $a(t)$.

The variation of training errors in one LUT can be attributed to the composition of transients that makes up the individual waveforms stored in the LUT. All excitation sequences which are unique for each transition causes transients within the circuit. If one transition causes a more persistent transient, all transitions containing it will show a deeper memory than others. By increasing the order of the LUT entries that result in such comparably large extraction errors, the LUT's overall error can be significantly reduced while maintaining a compact set of LUT entries. However, the model still has to cover the modulator's whole signal space, which is not trivial when dealing with symbols of different length.

Partitioning the modulator's signal space was solved by using an elimination-tree. This graph based representation, as in Figure 6.6, allows for a unique and comprehensive structuring. Even if the orders of a LUT leaf nodes are diverse, the model still

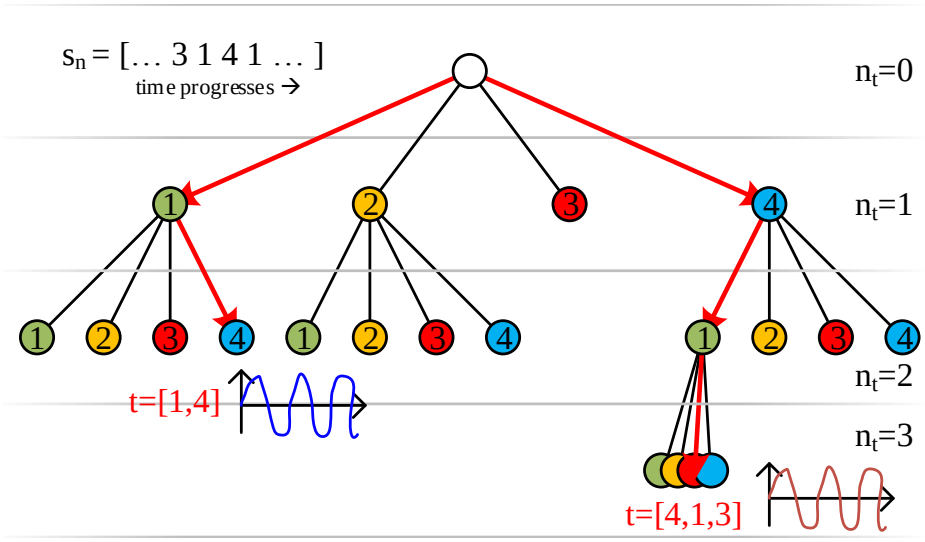


Figure 6.6: Graphical representation of traversing the tree of a hierarchical LUT to find the corresponding leaf node.

computes mapping from symbols to transitions efficiently. The implementation of this tree based LUT model was accomplished using the *Matlab* class *@tree* [50].

To gain a finite and, therefore, realizable model, the partitioning algorithm that adds branches to the tree requires a termination criterion. For this work a fixed threshold of the normalized extraction error was chosen to terminate the refinement of a tree. Starting from the root node, every leaf is refined until every resulting leaf node features a normalized extraction error below the threshold. Additionally, a secondary criterion has been applied to limit the maximum transition order to seven. The second criterion is solely implemented to ensure a termination of the tree within the transition combinations offered by the conducted measurements. However, the second criterion is of no practical relevance if the primary criterion is chosen properly. In contrast to constant length LUT models, this approach overcomes the need for identifying the circuit's effective memory length beforehand. It automatically grows from low to high transition orders until the desired error criterion is met.

The resulting error diagram for a hierarchical LUT model is depicted in Figure 6.7. It shows that the partitioning algorithm segments the signal space significantly nonuniform. In contrast to constant order LUT models, this reduces the maximum extraction error of transitions that show long memory effects and it scales far less granular with

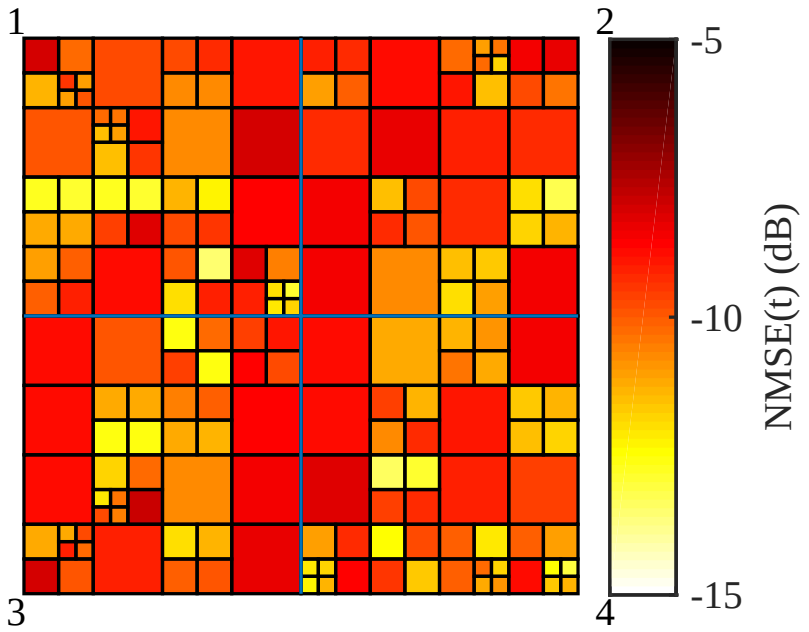


Figure 6.7: Normalized training error of a hierarchical LUT. The *nerr* threshold was set to -12 dB, resulting in a LUT size of $n_q = 181$.

respect to the size of the LUT. Leaf nodes can be added or dropped as long as the partitioning condition is meet.

The complexity of hierarchical LUTs can, therefore, be scaled arbitrarily while constant order LUTs need to scale exponentially. Figure 6.9 illustrates this behavior by comparing the extraction error for constant order and hierarchical LUTs. The error bars indicate the set's maximum and minimum values as well as the set's mean and standard deviation. The blue bars indicate constant length LUTs, the red bars represent hierarchical LUTs, and the star indicates the result gained from the verification set explained in section 6.5.

While the average error is about the same for constant order and hierarchical LUTs of a given model complexity, the hierarchical model shows significantly reduced maximum error as depicted in Figure 6.9. The expansion of the maximum error of the hierarchical LUT in the last two data points originates from the secondary termination criterion limiting the maximum memory depth. Because this is a result of the chosen implementation and not of the approach itself, it is considered an artifact. When com-

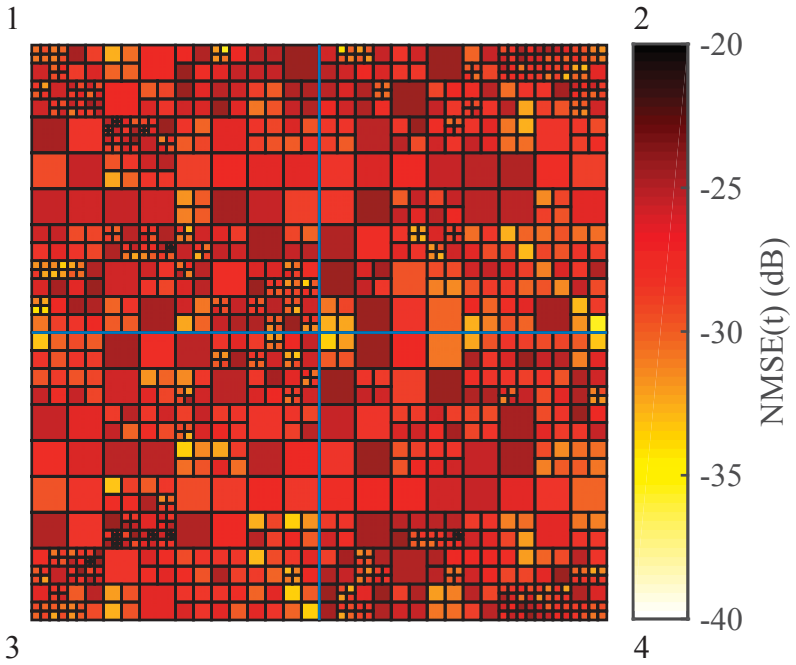


Figure 6.8: Normalized training error of a hierarchical LUT. The *nerr* threshold was set to -25 dB, resulting in a LUT size of $n_q = 1090$.

paring the confidence intervals of the two LUT approaches, it is important to note that the hierarchical LUT shows significantly smaller variations. This means it delivers a more stable performance for different input sequences.

6.5 Verification of LUT Models

In order to gauge a given LUT model, there are several options: On one hand, it is possible to simulate a measured sequence in order to compute an error measure like NMSE from the results. This sequence should resemble an RF-signal that is relevant for the targeted application. For the sake of comparison, the same 20 MHz LTE signal was used that already served as a verification sequence for the ARX Models.

Simulating the output of a system using a hierarchical LUT model requires to use the LUT model's tree to group the modulator symbols to the corresponding transitions of individual length. For any new input symbol the tree needs to be followed from the

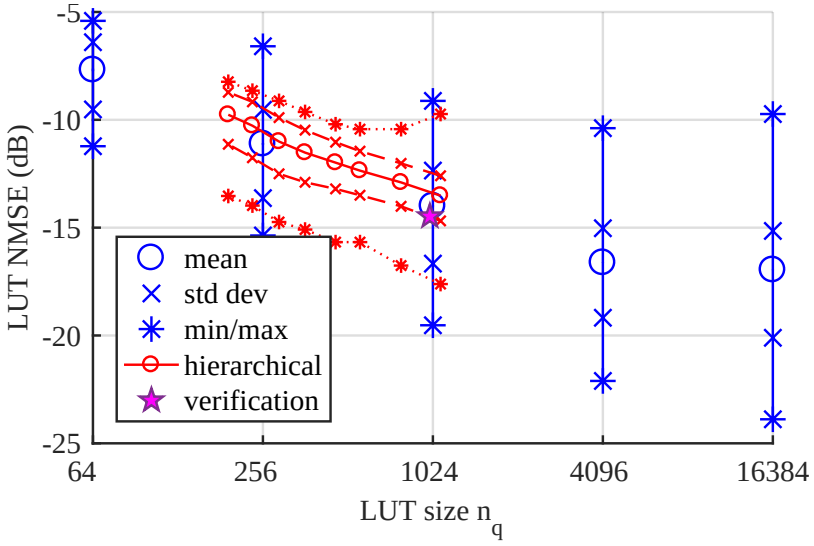


Figure 6.9: Normalized extraction error with respect LUT size n_q.

trunk up to a leaf-node to distinguish the number of prior symbols effecting the current waveform. The process is illustrated for a sample section of symbols in Figure 6.6. While in the first case ($t = [4,1,3]$) a leaf node is reached after three steps, which means that two precursory symbols are needed to capture the circuit's memory behavior of this branch. In the second time instance ($t = [1,4]$) only two nodes are sufficient to capture the memory in this case. Each of the so found leaf-nodes corresponds to exactly one LUT entry, which means that the complexity of the model is defined by the number of its leaf nodes. A modulator centric depiction of this process can be found in Figure 6.10.

Generating an error measure this way is a valid approach which is most comparable to other models. However, it only provides a single error measure that entangles the modulator's statistic with the model error on a transition level. The results may vary significantly depending on the frequency with which the modulator selects specific transitions and how well these transitions are captured by the model. This way is equivalent to deriving an implicitly defined, weighted average of all the LUT entries.

On the other hand, it is possible to break this entanglement by gauging a LUT model on a transition basis, this is similar to building a new model based on the measurements gained form the verification sequence. The resulting error values are com-

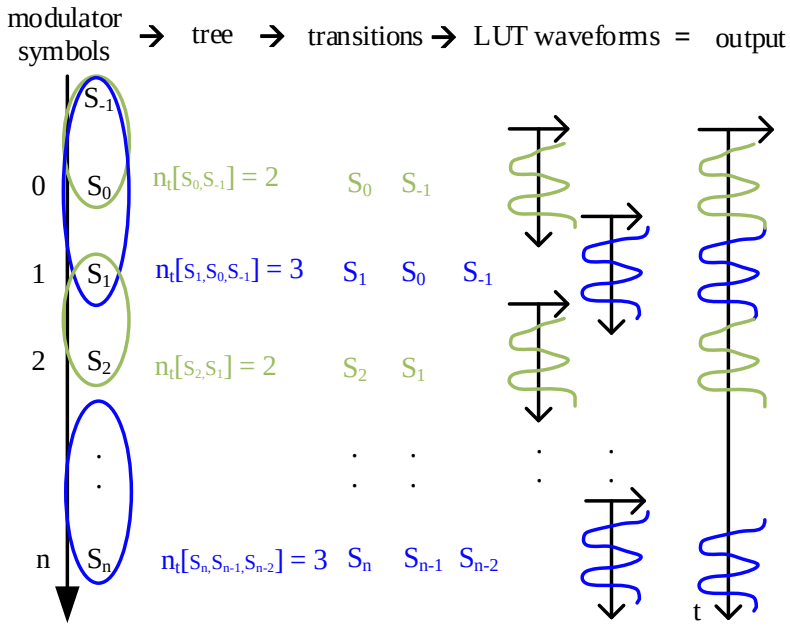


Figure 6.10: Simulation process for a general LUT model.

puted exactly like the *nerr* is computed. However, in this case the waveforms found in the LUT model are compared to all realizations found in the verification measurements. Since the verification sequence can be incomplete with respect to the signal space of the modulator, it is possible that it is missing some entries. The weighting function $a(t)$ needs to be re-normalized to account for this.

In a more formal notation: For a given LUT entry all realizations corresponding to a transition t form a verification waveform set $W_v(t)$. Since not every possible symbol combination needs to be present within the verification sequence, the verification set T_v is a subset of T. Each of the waveforms in the verification set is compared to the corresponding LUT entry to derive the corresponding prediction error. An example for such a comparison is depicted in Figure 6.11. Contrary to many other modeling approaches, the prediction error is less than the extraction error under the assumption that the model order n_t is chosen to be large enough. In this case the identification sequence contains fewer different symbol combinations than the verification sequence, and, therefore, is bounded by the extraction error. The results in Figure 6.11 underline the validity of choosing the training error as quality measure for building LUT models. Figure 6.9 provides a comparison between the performance estimated from the

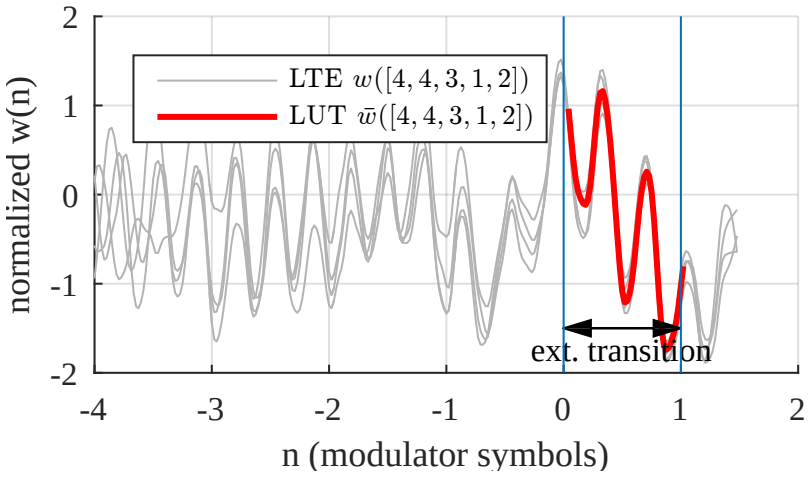


Figure 6.11: Comparison of a typical waveform predicted by the LUT compared to all realizations in the verification sequence.

extraction error and the results gained from a LTE frame verification signal.

CHAPTER 7

MODEL COMPARISON

This chapter discusses the difference between the two nonlinear modeling approaches described in chapters 5 and 6 and what can be gained from using behavioral models in DRFTx system design. As introduced in chapter 4 there are many options to model a DRFTx or sections of it. While models are often discussed solely based on their complexity and the resulting predictive capability, there is much more to it.

The estimation step for example can be equivalent important for the application. It is required to find the parameter set of the model that can be significantly different in terms of complexity and stability for different model types. Section 7.1 discusses that for the model types used in this work. It also shows how it is possible to gain information about the circuit itself in this step. However, the simulation will be used ultimately to predict the output of a circuit. The prediction error that was achieved for the two investigated model types is compared in section 7.2. Finally, a conclusion of this work is provided in 7.3.

In order to test how well different models can capture specific effects, it was found useful to use different sections of the DRFTx as DUTs. The section up to the booster amplifier shows strong nonlinear behavior, although, relatively short memory effects. The DRFTx circuit up to the transistor in the PA, on the other hand, shows much longer memory. Since the output waveforms of these sections can be probed in the DRFTx lab setup, they are valuable test cases. Figure 7.1 provides an overview of the DRFTx and the signals that are to be expected at the individual interfaces.

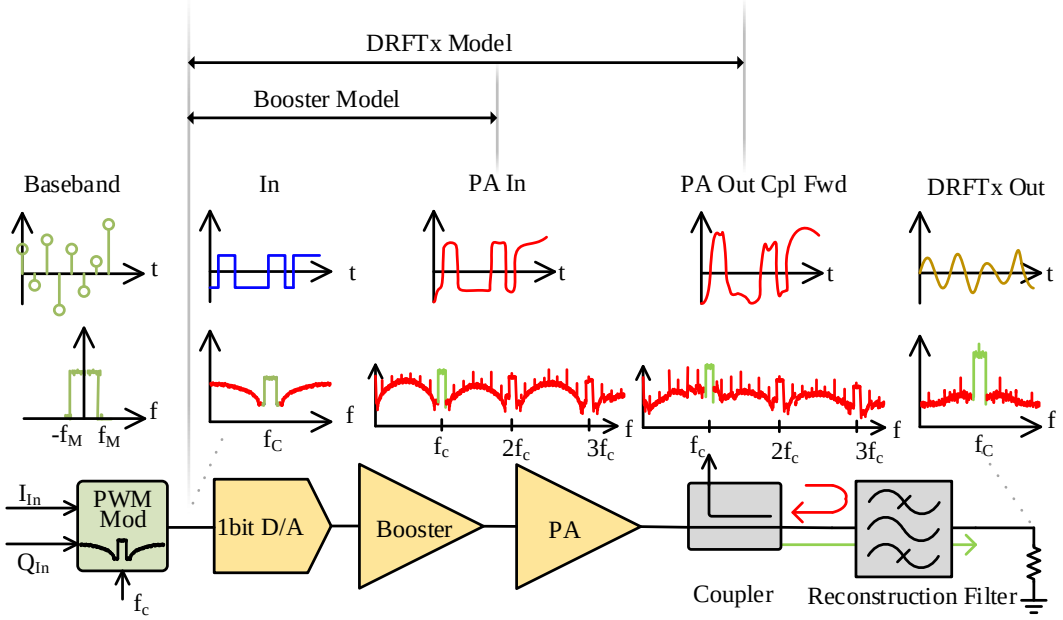


Figure 7.1: Sections of the DRFTx circuit and their corresponding models.

7.1 Model Estimation and Model Complexity

The complexity of a model is defined by a set of model parameters. They specify properties like the linear and nonlinear memory depth, the type of nonlinear feedback and so on, although, they are specific for a certain model type. In any case, there are usually many of them and they determine how many model coefficients need to be found during the model estimation step. Since these coefficients are also used to compute the output signal of the model, the computational cost for a simulation run of a model often scale with its complexity. Since typical signals in DRFTx systems are lengthy by nature, the computational cost for a simulation run of a model is relevant for its practicality. So a compact fast model is more beneficial than a slow overly complicated one with too many model coefficients.

Finding the model parameters required to capture all relevant effects within the DUT, is a major challenge for the overall modeling process. They can either be estimated beforehand, as for the ARX model described in chapter 5, Or they are estimated inherently within the model estimation, as for the LUT model described in chapter 6.

An example for one of these model parameters is the memory depth that can be

represented by the model. In order to capture the circuit behavior correctly it needs to cover the relevant memory depth of the circuit. While for linear circuits the memory depth can be measured using various methods, this is different for strongly nonlinear circuits. For the presented case, where some of the nonlinear effects are memory induced, the results will also depend on the given excitation. The training error resulting from the estimation step is an indicator if the parameter set of the model was chosen adequately.

Such an analysis is depicted in Figure 5.2 for an ARX model and in Figure 6.9 for the LUT model. Both models show that the resulting training error steadily decays with increased complexity. However, there are practical limits to the maximum complexity which are different for both model types.

The parameter space of the ARX model is mostly limited by the conditioning of the model estimation problem. As the number of model coefficients gets larger, the training sequence remains the same which reduces the conditioning of the estimation problem till the estimation algorithm fails. Using multiple different training sequences to overcome this issue showed little success as the utilized sequence is already optimized to be well conditioned. Another problem with large model parameter sets is the multitude of model parameter combinations which results in an equal complexity. Finding an optimum combination based on the estimation error is very time consuming and doesn't guarantee for good modeling results for other excitation sequences.

The LUT model avoids the parameter estimation problem since it just tabulates observations. If this is done in a structured manner, the resulting memory adaptively increases until the resulting training error has decayed sufficiently. Nevertheless, the memory demand for the table increases, in the worst case, exponentially with the resulting memory depth. Since all table entries need to be filled in the training step and, therefore, the training sequence must cover all possible signal combinations for the given memory depth, its length increases proportional with the complexity of the model. Ultimately, the memory demand and the memory of the instruments used to conduct the estimation step limits the maximum memory depth.

Gaining universal performance values for the presented models is difficult since they function very different. However, the time required to conduct one simulation for a typical verification sequence was chosen for this work. The ARX model requires a vast amount of floating point operations, while the LUT model purely relies on data

addressing. The LUT model can utilize parallel processing very well while the ARX model can't, however, for reasons of consistency a sequential implementation of both models is assumed.

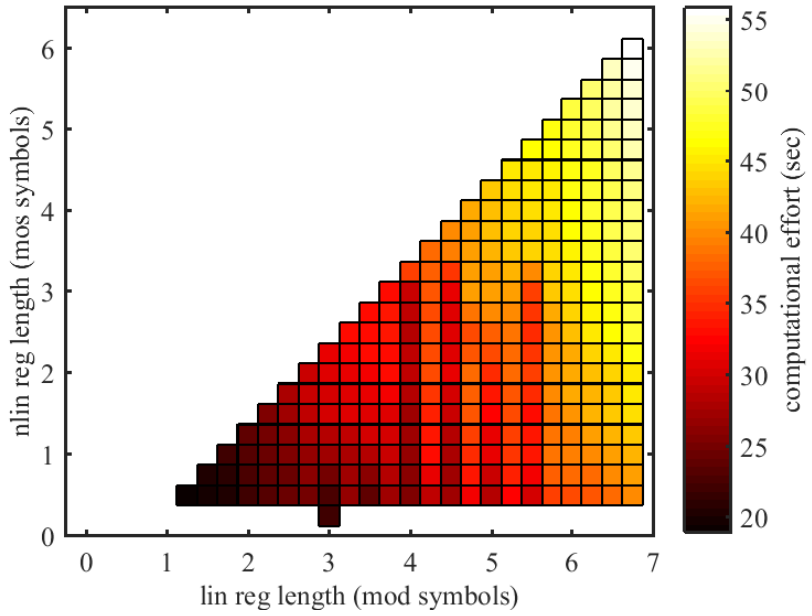


Figure 7.2: Computational effort of nonlinear ARX models with respect to regressor length.

In both cases, the simulation time is obviously proportional to the length of the input sequence. However, from the results summed up in Figure 7.2, it can be concluded that the simulation time is roughly proportional to the complexity of the model in the ARX case.

For the LUT model depicted in Figure 7.3 the computational effort grows exponentially, due to the number of table entries. Although it is important to note that this assumes the least sophisticated data structuring of the LUT, the model doesn't utilize parallelism in the simulation algorithm, and still provides about seven times higher simulation rates under the same conditions as its ARX counterpart.

In a hierarchical LUT the memory depth of individual table entries is set adaptively. Depending on the DUT, this can result in a significantly more compact model at equivalent prediction performance. Figure 7.4 depicts the relative frequency of transition lengths in a hierarchical LUT model as a function of its length. It shows that the

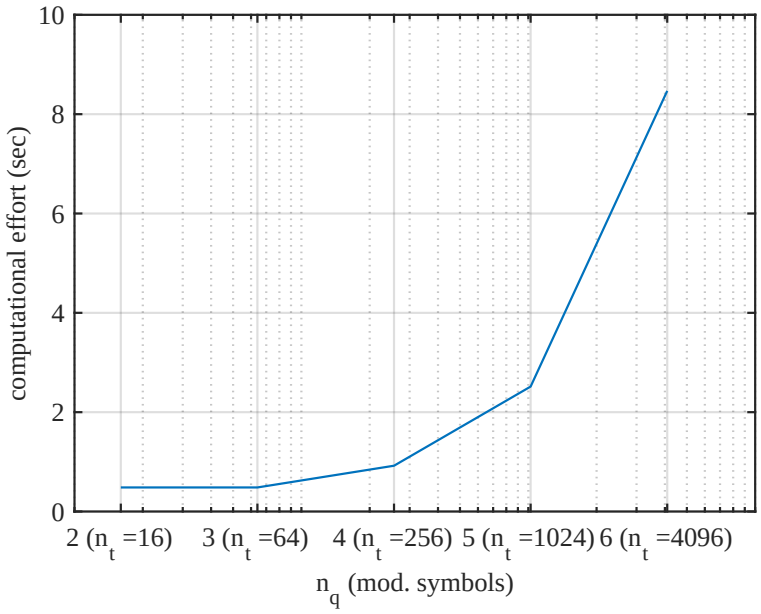


Figure 7.3: Computational effort of nonlinear LUT models with respect to regressor length.

algorithm utilizes the hierarchical structure of the LUT substantially, especially around a model complexity equivalent to a 4-5 modulator symbols fixed length LUT model. This trend would continue if the estimation would not be restricted to $n_t \leq 6$ by the training data-set.

7.2 Simulation and Prediction-Error

When estimating models based on recordings of training sequences, the estimation step results in a training error for the model. However, a model that results in small training error does not guarantee for good overall performance. Strictly speaking, the training error only indicates how well the model explains the training measurement. For any other excitation sequence the result will be different, although, a model that delivers similar performance for a wide range of input sequences is considered more robust.

To test the robustness of the models, a separate verification step is required. A sequence that resembles the target application of the circuit is of the greatest practi-

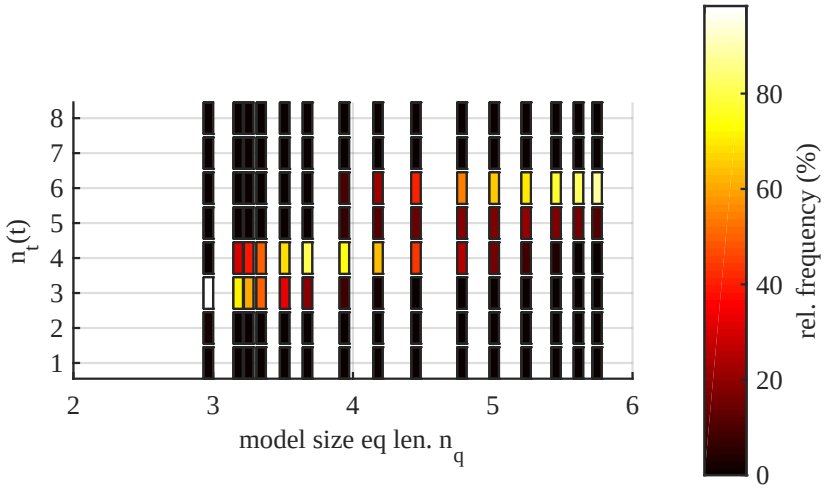


Figure 7.4: Variation of transition length for LUT models of the full DRFTx circuit for different model complexity.

cal relevance. Since the verification sequence is less constrained than a training sequence, any modulator sequence can be used. For this work a sequences that generates a 20 MHz LTE frame was used. The DRFTx lab setup was used to gain recordings from the waveforms that can be also predicted by the model and compared directly. The names of the signals and interfaces are according to Figure 7.1. The excitation sequence has a length of about 1 Mbit and the simulation uses an oversampling rate $r_{OSR} = 4$ resulting in a signal length of 4 MS.

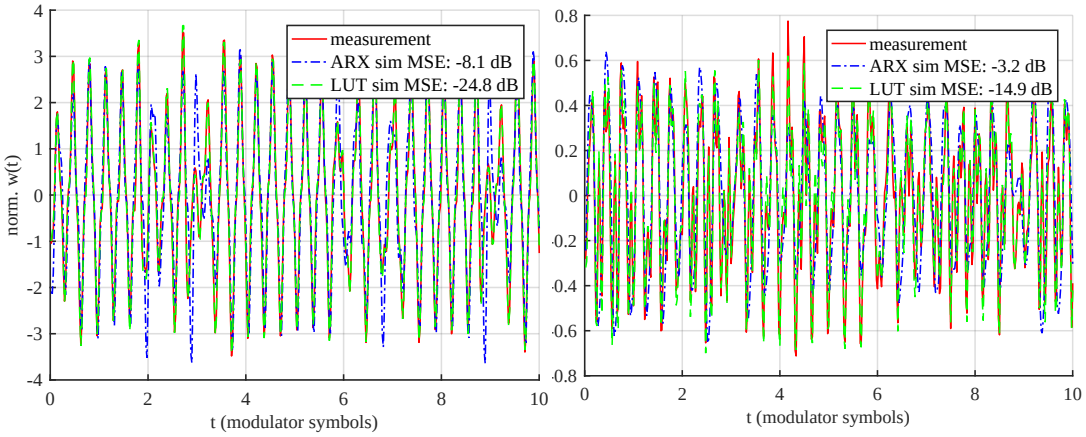
For the booster stage, the ARX model used a regressor length of 1.75 and 1.25 modulator symbols for the linear and nonlinear part, respectively. The LUT model is a non hierarchical model of order $n_t = 6$ modulator symbols.

For the full DRFTx circuit, the ARX model used a regressor length of 3 and 6.75 modulator symbols for the linear and nonlinear part. The LUT model is a non hierarchical model of order $n_t = 6$ modulator symbols.

Time Domian

As the LUT model and the recording method are based on time domain, the NMSE can be computed directly from these signals. For the booster stage the error of the ARX

model is significantly higher than the training predicted, however, the error resulting from the LUT model is only slightly higher than its training results. The results for the full DRFTx show similar tendencies and result in error values that are worse than what is expected from training.

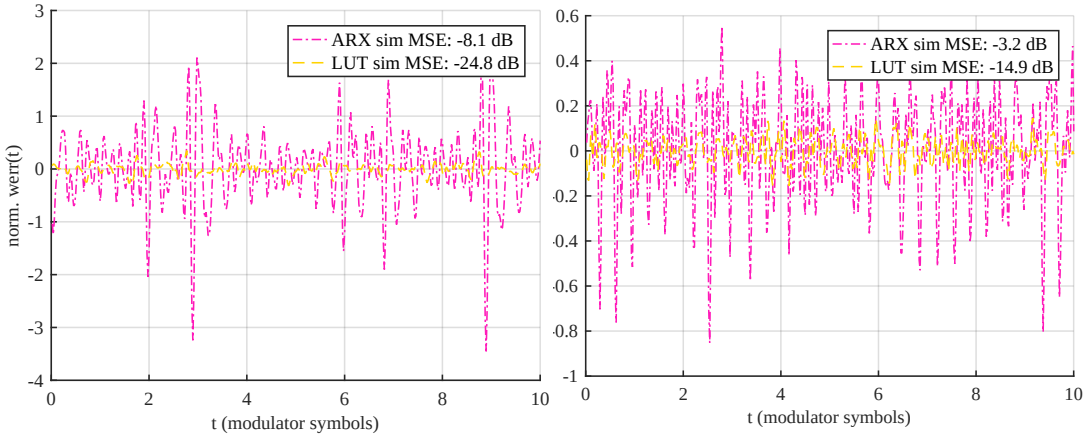


(a) booster (b) DRFTx

Figure 7.5: Comparison of the predicted output waveforms in time domain.

The prediction error of the models is exemplary depicted in Figure 7.6. The prediction by the LUT model is obviously significantly better than the one derived from the ARX model in both cases. It is important to note to that the error of the LUT model is not concentrated at boundaries of modulator symbols. This indicates that there is only little residual memory present in the circuit, otherwise, discontinuities at the interfaces of different table entries would be unavoidable.

Obviously, precise synchronization between the signals is necessary to gauge the prediction error in time domain. Due to the steep slopes of the waveforms, sub-sample synchronicity is required for precise error results. Although the system is synchronized by using the bit clock of the excitation, synchronizing the recordings in post-processing is required. One ongoing systematic source of error is the sampling oscilloscope itself, which introduces an arbitrary lag for each recording run. The lag is compensated in post-processing, although, due to the bad correlation properties of the signals this process is unreliable. Only if this lag can be considered static, this issues can be mitigated by brute force optimization of the synchronicity. Since the verification sequences are lengthy, this approach causes a lot of computational effort which makes its use im-



(a) booster (b) DRFTx

Figure 7.6: Comparison of the prediction errors in time domain.

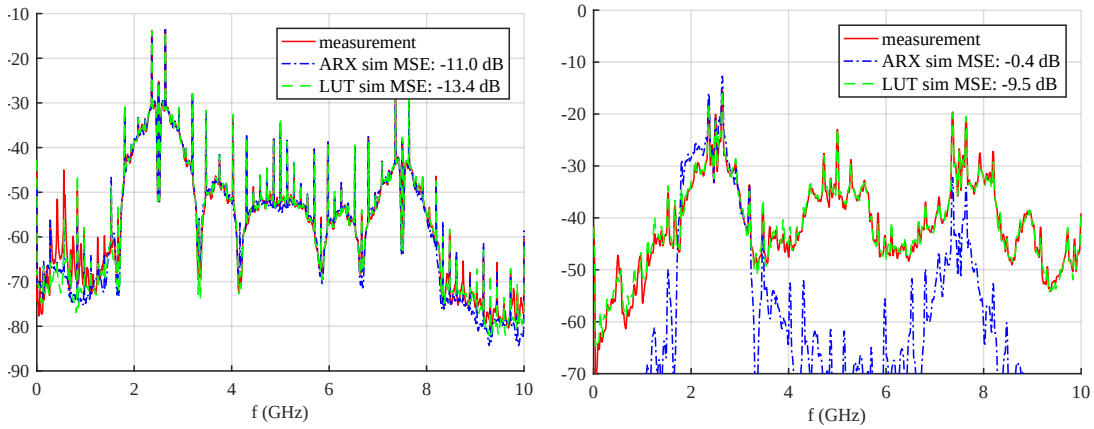
practical. Therefore, the resulting NMSE error is mainly determined by the imperfect synchronization and, therefore, insufficient for gauging the model.

Frequency Domain

In order to compare lengthy signals like these in frequency domain, a power spectral desity estimate (PSDE) was chosen instead of using a complete Fourier-transformation. Computing a full Fourier transform would result in an unwieldy detailed spectrum, given the lengthy signals, which is not useful for depicting the observed effects. Also some instruments, like spectrum-analyzers, provide only a PSDE of the measured signal. It would also be somewhat pointless as a comparison since the Fourier transform is a linear operation. So the prediction error gained form a complete time domain observation must be equal to that of a complete complex spectrum. On top of that, the PSDE of different signals can be compared without the need for synchronizing them, which makes it very interesting for practical applications. In this section a frequency representation based on a Blackman-Harris window of length NFFT $= 2^{16}$ is used. To calculate the PSDE, the windowing function selects NWIN $= 10^4$ subsections randomly from the recording using a normal distribution.

However, the shape of the waveforms is strongly influenced by the phase relations between different energy rich components present at a given time instance. Neglecting

the phase information altogether is therefore not suitable when gauging the prediction of such a model. However, in order to identify the individual error contributions it can be a useful tool.



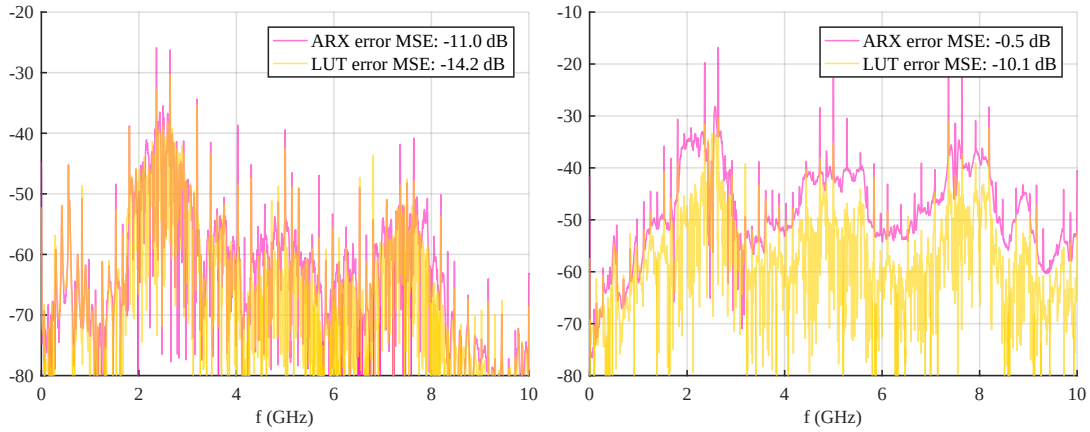
(a) booster

(b) DRFTx

Figure 7.7: Comparison of the predicted output signals in the frequency domain.

Figure 7.7 compares the simulated results from the ARX and the LUT models to the measurement from the lab setup in the frequency domain based on a PSDE. The diagrams show that the MSE estimate for the booster stage gained from the frequency domain is around -10 dB. These results are surprising since the model estimation process in chapters 5 and 6 resulted in training errors of up to -30 dB. Also the time domain analysis predicted better results for the LUT model. The error for the DRFTx section is around -0 dB which renders it useless for any practical application, although, the energy rich signal components around the center frequency and the second harmonic of the system has been modeled correctly. The LUT model predicts the overall frequency response much better, although, the error values are as good as predicted by the training step.

Figure 7.8 compares the spectral distribution of the error for the ARX and LUT model predictions. Since their power density is roughly proportional to the spectrum of the excitation signal, it is unlikely that the error is due to oscillations within the model.



(a) booster (b) DRFTx

Figure 7.8: Comparison of the prediction error in frequency domain domain.

Spectrogram

The discrepancy between frequency- and time-domain results, as well as the significant difference between training and verification suggest that residual measurement errors are masking these results. Since the synchronization of the recording showed to have a significant influence on the error results, a compromise between time and frequency analysis was used to gain more stable results. This method is similar to the synchronization used in orthogonal frequency division multiplex (OFDM) systems [51].

At first the sequences are synchronized using a correlative delay estimate. In a second step, a time frequency representation of both signals is generated using a Blackman-Harris window of length NFFT $= 2^{16}$. The larger bin sizes allow to get a more reliable delay estimate based on the phase of both signals. After compensating the remaining delay by adding a linear phase term the NMSE is computed in a straight forward manner.

In contrast to operating solely in the frequency domain, the waveforms need to be synchronized as the windowing in the time domain follows a sequence which ensures that the short-time spectra of corresponding waveform sections are compared. But, since they are frequency domain results, a slight and almost constant lag between the signals only affects the phase information. The experimental results for the booster

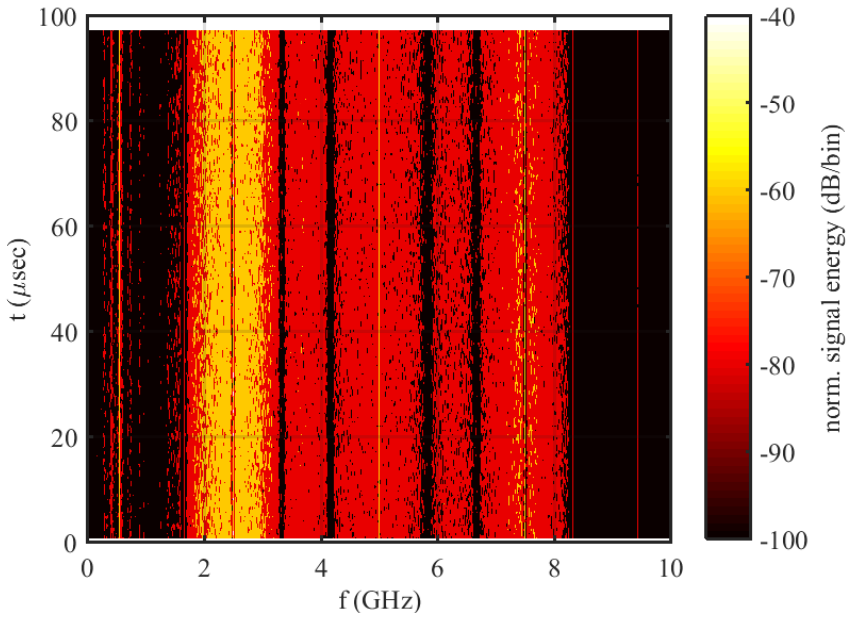
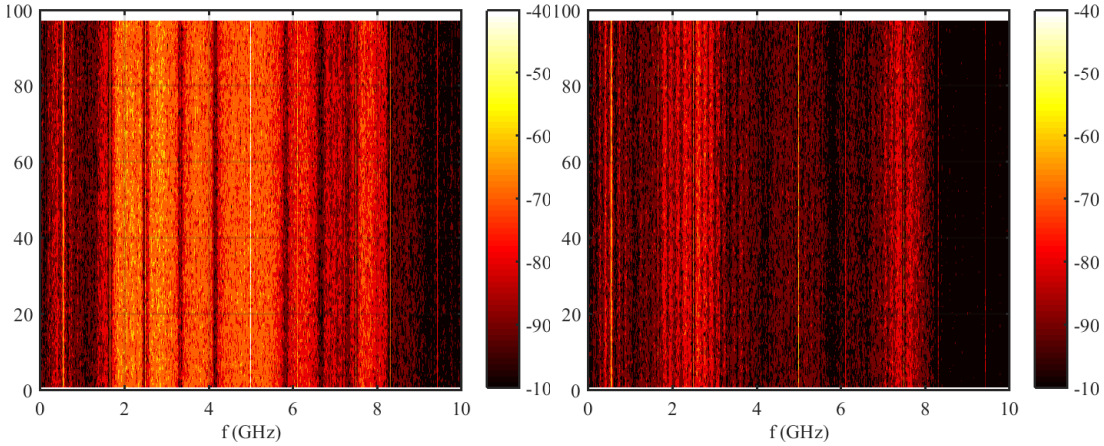


Figure 7.9: Spectrogram of the booster's output signal using $NFFT = 2^{16}$.

stage are depicted in Figure 7.10 and show good agreement with these theoretical considerations. The MSE for this method resulted in -11.4 dB for the ARX model and -20.9 dB for the LUT model.

The results for the full DRFTx circuit are summed up in Figure 7.11 and are surprising. The MSE derived using this method resulted in -5.0 dB for the ARX model which is in good agreement with the value computed in the time domain. However, the error value for the LUT model resulted in a MSE of -23.8 dB which is about 10 dB better than the time domain result. Although, when compared to the modeling error presented in section 6.5 they are within a plausible range.

Thirdly, the MSE error measures computed from these signals in the time domain are very sensitive to synchronization errors. Even small variations influence the resulting MSE significantly without affecting the quality measures gained from magnitude spectra. This effect was considered during the design of the test bench and is mitigated by synchronizing the measurement setup using the bit clock. Estimating the overall sampling clock jitter for the test bench was hindered due to limited information about the jitter-contribution of the individual instruments. However, the discrepancy between the spectrogram error and the time domain error are most likely caused by



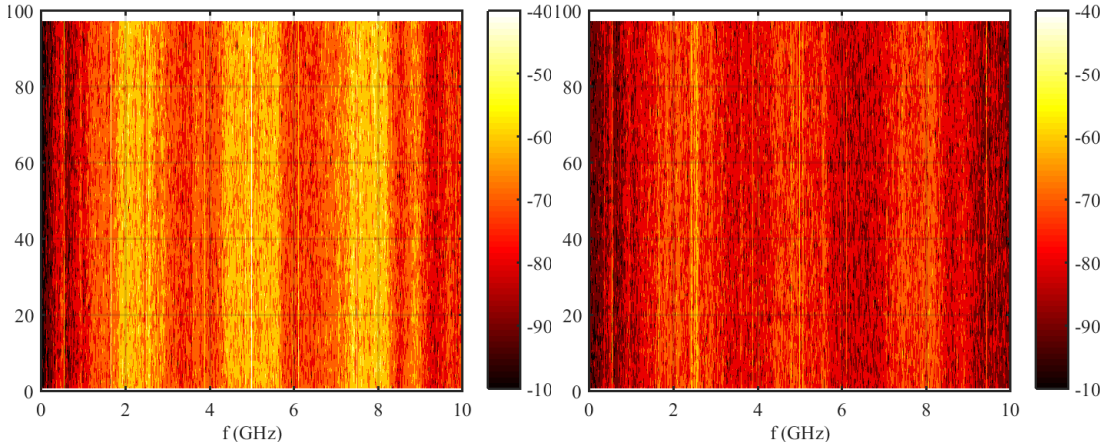
(a) ARX model (b) LUT model

Figure 7.10: Spectrogram of the prediction errors for the booster stage of the DRFTx.

a slowly varying time lag between the channels. The spectrogram representation of the signals would not be affected by this while the time domain representation would. However, this hypothesis is beyond the focus of this work.

Finally, every measurement run is delayed by an arbitrary lag introduced by the sampling oscilloscope. The lag needs to be compensated in post processing using correlative methods, which not always yield perfect results. The resulting timing error can lead to misleadingly bad error measures.

To illustrate these effects, the results are summed up in Figure 7.12. The training error (trn.) of the LUT models are depicted by the error bars showing that there is a dynamic of about 5 dB NMSE within one LUT model. The error is steadily decaying for higher LUT sizes which shows that the residual error due to uncaptured memory decreases for longer memory. The star-shaped marker indicates a single verification based on the LUT extraction process. Due to the long computation time and unreliable results, only one of the longer LUTs was tested. The result shows good agreement with the training error, yet, it is slightly lower than predicted. The red solid line represents the verification results gained from the simulation based approach. It also shows good agreement with the training error, although, it is significantly lower for very compact LUT models. This effect is caused by the utilized PWM modulator which prefers to generate long runs of the same modulator symbols over frequently



(a) ARX model (b) LUT model

Figure 7.11: Spectrogram of the prediction errors for the DRFTx.

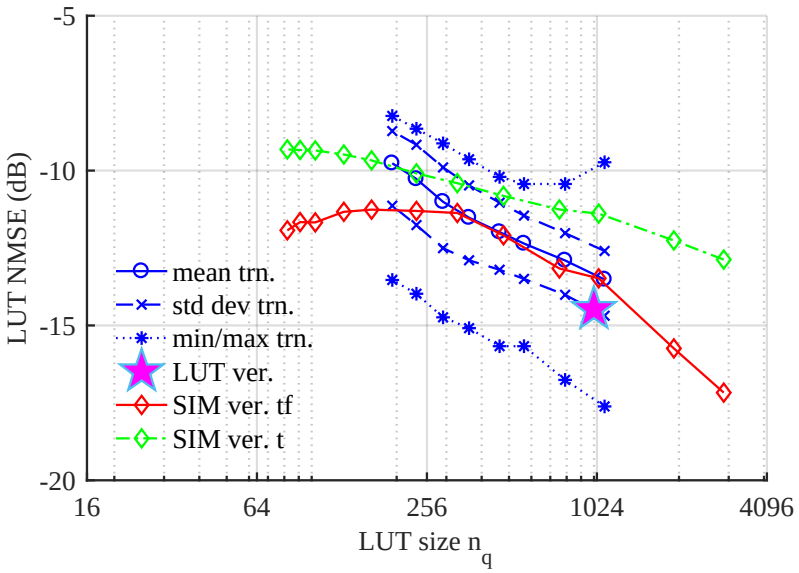


Figure 7.12: Comparison of training- and simulation errors of different hierarchical LUT models depending on their size.

changing them. However, memory effects are more dominant if frequent symbol transitions occur. Consequently, the training sequence that has a higher content of such subsequences also has a slightly higher overall NMSE error. For the special case of restricting the LUT model to values shorter than the memory of the circuit, the waveforms stored in the model mostly represent a quasi-stationary state of the circuit which is the best trivial estimate for its output waveform. Finally, the green dash-dotted line depicts the simulation error without time-frequency lag compensation. These results show that residual measurement errors mask the prediction performance of the model if not handled properly.

7.3 Conclusion and Outlook

This work is based on experimental results from a laboratory test bench for a DRFTx which was built to produce a 20 MHz LTE signal at a center frequency of 2.5 GHz. Since the DRFTx in this setup was designed specifically for pursuing an experimental approach, it values ruggedness and a flexible design over a compact solution, or one that would require highly complex processes to produce its functional units.

From a circuit design aspect the design could be improved in two major, yet mutually exclusive ways:

Firstly, taking circuit design concepts that require a circuit to reach a quasi-equilibrium and adapting their results to accommodate heavily modulated signals, as described in chapter 3 is probably sub-optimal. While these results are a good starting point for a design, these design issues are generally better addressed using simulations of the transient processes in time domain as in [16] than forcing it into a frequency domain framework. It would allow to directly investigate important circuit parameters based on the waveforms using available computer aided RF design solutions.

Secondly, the decision on using a binary excitation source makes the design more complicated than it needs to be. By limiting the design to a single bit DAC the coding efficiency of the source is very limited. The most tricky circuit design aspects, like the very selective reconstruction filter and the long memory effects have their ramifications in this design aspect. The nonlinear amplifier in the booster amplifier could be adapted relatively simple, to have a controllable envelope. The bandwidth constraints on the envelope control are in the range of some multiples of the target system bandwidth.

Achieving such an envelope control would boost the coding efficiency by orders of magnitude, relaxing the constraints on the circuit design significantly. In order to be compatible with the LUT model purposed in this work, the envelope signal would need to take on discrete values which is not a major limitation. In both cases the modeling techniques presented in this work would still be applicable.

A rather simple, however, important contribution to the DRFTx circuit is the addition of a compact, yet broadband, coupler into the matching network of the PA. This allows to probe all signal components that are present at this point in the circuit. Practically, this makes a big difference since most of them are strongly rejected by the output filter and, given the limited dynamic range of the measurement instruments, wouldn't be observable at the output of the system. This is necessary to be able to explain the interaction between excitation signal, nonlinear distortion, and power consumption.

Since DRFTx systems are niche systems, modeling these circuits is even more exotic. At this point it is important to note that this work puts an emphasis on circuit centric modeling of a DRFTx. This distinction is important as one can think of a DRFTx as the combination of modulator and transmitter in an RF communication system.

From a telecommunications perspective, an equivalent baseband description of the DRFTx is sufficient to gauge its performance. Therefore, it is very popular and also allows to reduce the sample rate of the model significantly to cope with deep memory effects more easily. As a result, the majority of works on DRFTx models falls into that scheme. The prediction error of these approaches is outstandingly low, although, they only cover signal components which fall into the equivalent baseband of the system.

From a circuit design perspective, one could also think of a DRFTx, as a power amplifier driving a highly reflective load. In order to compute PA related measures like power efficiency. The presented approach requires to capture any energy rich signals at the active device of the PA. This has been solved by predicting the RF-waveforms of this unit. It allows to co-design the modulator and circuit without the need to build all modulator circuit combinations to test them. Doing so balances modulator complexity, circuit alternatives, and noise shaping parameters to gain an overall optimal combination. This approach and aims at complementing equivalent-baseband models as both approaches excel in their individual domain.

From a modeling perspective the nonlinear ARX models, alongside all filter based models, are one of the most general. These models are not limited to a specific class of excitation signal and they describe the modeled system by a relatively compact set of parameters. However, estimating these parameters is generally computationally expensive and can lead to unstable models. Additionally, it is not possible to take advantage of the limited variety of excitation sequences generated by the modulator.

Exploiting the limited signal space of the modulator is the most important step to make a LUT model feasible. It can be used only for a signal composed of predefined modulator symbols for given rate settings. However, for typical DRFTx applications there is anyway a very limited set of modulator symbols and modulator parameters that make sense [4]. Although, this limits a LUT model to a specific set of PWM parameters, it does not anticipate a specific baseband encoder/modulation algorithm. So with respect to DRFTx design applications the LUT model is as universal as its counterparts. But, since it only tabulates waveforms gained from measurements, the model cannot be unstable, the implementation can be computationally very efficient, and memory is the only limiting factor with respect to complexity.

The verification of nonlinear RF-models has also been studied within this work. While the theory behind comparing predicted waveforms to measurements is pretty straight forward, distinguishing between effects that are caused by the measurement setup and those caused by the model is not. The large bandwidth and high frequencies of the signals involved makes it tricky to mitigate drift and jitter effects within one measurement run and even more problematic when comparing different measurements. As demonstrated in chapter 7, comparing complete LTE signals can be problematic because of their lengthy nature. Additionally, it inherently entangles the statistic of the modulator with the resulting error measure. The more general approach is to segment the waveforms into typical subsections and compare these separately. Luckily this functionality comes naturally with the use of LUT models, since their table structure can be used to segment any possible waveform in a DRFTx in a unified way. Doing so separates the error caused by residual memory effects from the algorithm used in the modulator and provides a more holistic error description.

By utilizing more sophisticated data structuring methods for building the table of a LUT model, the memory demand of a LUT model can be improved. This work does

so by using an elimination-tree representation of LUT entries. It allows to deal with entries of different memory depth while ensuring that any possible excitation signal combination is covered by the model. The adaptive branch length of the tree allows to truncate individual branches based on the error caused by residual memory. As a consequence, the tree can either be adjusted for a given memory size or a given modeling error requirement. While the average error of this modeling approach remains comparable to a LUT with constant length, it is demonstrated that the error bounds of the modeling error could be dramatically reduced.

The results show that moving from a non-hierarchical model to a tree based, adaptive, description is practically and scientifically interesting. However, trees offer much more flexible topologies than those utilized in this work. Developing a tree model that branches only specific nodes in each layer could improve the performance of the model for sparse delay scenarios. Implementing and testing such more advanced tree topologies could expand the usability of LUT models for problems with a significantly bigger signal space, beyond only DRFTx circuits.

Time-frequency methods, similar to what is used in today's OFDM systems, have been found to be a robust option for gauging verification errors of lengthy signals. By using this method, it was possible to demonstrate that the error gained during the model training procedure is in good agreement with verification measurements. For future work, it seems promising to analyze the dependencies between residual errors calculated by the time-frequency method and memory effects with time constants beyond the capability of the model. A promising example would be to compare the effects of self heating by measurements and what can be explained through the power dissipation as predicted by the LUT model.

References

[1] C. Roff, J. Benedikt, and P. J. Tasker, "Design approach for realization of very high efficiency power amplifiers," in *IEEE/MTT-S International Microwave Symposium (IMS)*, pp. 143–146, 2007. 3

[2] S. C. Cripps, *Advanced techniques in RF power amplifier design*. Artech House microwave library, Boston: Artech House, 2002. 3, 18

[3] J. Wonhoon, N. Silva, A. Oliveira, and N. Borges Carvalho, "Analysis on in-band distortion caused by switching amplifiers," *IET Microwaves, Antennas & Propagation*, vol. 8, no. 5, pp. 351–357, 2014. 3

[4] H. Ruotsalainen, H. Arthaber, and G. Magerl, "A new quadrature pwm modulator with tunable center frequency for digital rf transmitters," *IEEE Trans. Circuit and Systems II: Express Briefs*, vol. 59, no. 11, pp. 756–760, 2012. 4, 56, 77, 107

[5] H. A. Ruotsalainen and H. Arthaber, "Evaluation of quadrature pwm modulator performance for digital wideband transmitters," in *European Microwave Conference (EuMC)*, pp. 1308–1311, 2012. 4, 5, 8

[6] R. A. R. van der Zee and E. A. J. M. van Tuijl, "A power-efficient audio amplifier combining switching and linear techniques," *IEEE Journal of Solid-State Circuits*, vol. 34, pp. 985–991, July 1999. 4

[7] J. Keyzer, J. Hinrichs, A. Metzger, M. Iwamoto, I. Galton, and P. Asbeck, "Digital generation of RF signals for wireless communications with band-pass delta-sigma modulation," in *IEEE MTT-S International Microwave Symposium (IMS)*, vol. 3, pp. 2127–2130 vol.3, May 2001. 4

[8] W. Yuanxun, "An improved kahn transmitter architecture based on delta-sigma modulation," in *IEEE MTT-S International Microwave Symposium Digest*, vol. 2, pp. 1327–1330, 2003. 4

[9] M. Nielsen and T. Larsen, "A transmitter architecture based on delta–sigma modulation and switch-mode power amplification," *IEEE Trans. Circuit and Systems II: Express Briefs*, vol. 54, no. 8, pp. 735–739, 2007. 4

[10] J. Ketola, J. Sommarek, J. Vankka, and K. Halonen, "Transmitter utilising bandpass delta-sigma modulator and switching mode power amplifier," in *International Symposium on Circuits and Systems (ISCAS)*, vol. 1, pp. I–633–6, 2004. 4

[11] D. Seebacher, W. Bosch, P. Singerl, and C. Schuberth, "Highly efficient carrier bursting rf transmitter employing direct band pass filter connection," in *European Microwave Conference (EuMC)*, pp. 692–695, 2013. 4, 20, 21, 56

[12] F. M. Ghannouchi, S. Hatami, P. Aflaki, M. Helaoui, and R. Negra, "Accurate power efficiency estimation of ghz wireless delta-sigma transmitters for different classes of switching mode power amplifiers," *IEEE Trans. Microwave Theory and Techniques*, vol. 58, no. 11, pp. 2812–2819, 2010. 4, 8, 18, 32

[13] H. A. Ruotsalainen, H. Arthaber, and G. Magerl, "Quantization noise cancelation scheme for digital quadrature rf pulse encoding," in *IEEE MTT-S International Microwave Symposium (IMS)*, pp. 1–4, 2013. 5, 18

[14] H. A. Ruotsalainen, *Encoding and modeling techniques for center frequency agile digital RF transmitters*. PhD thesis, Vienna University of Technology, 2015. 5, 8, 11, 63

[15] L. Xiaoguang, "Tunable rf and microwave filters," in *IEEE Wireless and Microwave Technology Conference (WAMICON)*, pp. 1–5, 2015. 7

[16] N. Leder, B. Pichler, H. Ruotsalainen, T. Faseth, and H. Arthaber, "On nonlinear memory effects in all digital rf-transmitters," in *IEEE International Microwave and RF Conference (IMaRC)*, pp. 300–303, Dec. 2015. 7, 8, 56, 63, 80, 82, 105

[17] M. Helaoui, S. Hatami, R. Negra, and F. M. Ghannouchi, "A novel architecture of delta-sigma modulator enabling all-digital multiband multistandard rf transmitters design," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 55, no. 11, pp. 1129–1133, 2008. 8

[18] A. K. Gupta, J. Venkataraman, and O. M. Collins, "Measurement and reduction of isi in high-dynamic-range 1-bit signal generation," *IEEE Trans. Circuits and Systems I: Regular Papers*, vol. 55, no. 11, pp. 3593–3606, 2008. 13, 20, 57, 58

[19] N. Leder, H. Arthaber, and H. Ruotsalainen, "Characterization and optimization of pulse drivers for switched mode power amplifier measurements," in *IEEE International Workshop on Integrated Nonlinear Microwave and Millimetre-wave Circuits (INMMiC)*, pp. 1–3, Apr. 2014. 14, 57, 60

[20] N. Leder, T. Faseth, H. A. Ruotsalainen, and H. Arthaber, "Nonlinear modeling and model-validation for digital switched mode rf power amplifiers," in *IEEE Mediterranean Microwave Symposium (MMS)*, pp. 1–4, 2014. 14, 60, 68

[21] V. Carrubba, A. L. Clarke, M. Akmal, J. Lees, J. Benedikt, P. J. Tasker, and S. C. Cripps, "The continuous class-f mode power amplifier," in *Microwave Conference (EuMC), 2010 European*, pp. 1674–1677, 2010. 18, 20

[22] B. M. Merrick, J. B. King, and T. J. Brazil, "A novel continuous class-f mode power amplifier," in *IEEE Topical Conference on Power Amplifiers for Wireless and Radio Applications (PAWR)*, pp. 19–21, 2014. 18

[23] C. Kenle and P. Dimitrios, "Design of broadband highly efficient harmonic-tuned power amplifier using in-band continuous class-f mode transferring," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 12, pp. 4107–4116, 2012. 18

[24] N. Tuffy, A. Zhu, and T. J. Brazil, "Novel realisation of a broadband high-efficiency continuous class-f power amplifier," in *European Microwave Integrated Circuits Conference (EuMIC)*, pp. 120–123, Oct. 2011. 20

[25] B. Pichler, "Continuous mode class f amplifier," Master's thesis, Technische Universität Wien, jun 2014. 20, 42

[26] B. Pichler, N. Leder, T. Faseth, H. Arthaber, and G. Magerl, "Efficiency enhanced switched mode power amplifier for digital rf transmitters," in *IEEE Wireless and Microwave Technology Conference (WAMICON)*, pp. 1–3, 2015. 21, 42

[27] S. C. Pires, P. M. Cabral, and J. C. Pedro, "Impact of the amplifier-bandpass reconstruction filter interaction on the linearity of carrier amplitude-burst transmitters," in *IEEE Wireless and Microwave Technology Conference (WAMICON)*, pp. 1–4, 2013. 21

[28] B. Pichler, N. Leder, T. Faseth, and H. Arthaber, "Calibration method for coupler based time domain waveform measurements," in *IEEE Integrated Nonlinear Microwave and Millimetre-wave Circuits Workshop (INMMiC)*, pp. 1–3, 2015. 24, 52

[29] F. M. Ghannouchi and M. S. Hashmi, *Load-Pull Techniques with Applications to Power Amplifier Design*, vol. 32 of *Springer Series in Advanced Microelectronics*. Dordrecht: Springer Netherlands, 2013. 29, 30

[30] B. Pichler, N. Leder, T. Faseth, and H. Arthaber, "Design of a pwm driven continuous mode class f amplifier using harmonic load pull measurements," in *IEEE Mediterranean Microwave Symposium (MMS)*, pp. 1–3, 2014. 32, 42

[31] S. Y. B. and J. C. H., "A simplified real frequency technique appliable to broadband multistage microwave amplifiers," in *IEEE MTT-S International Microwave Symposium (IMS)*, pp. 529–531, 1982. 42

[32] B. Siddik Yarman, "Modern techniques to design wide band power transfer networks and microwave amplifiers on silicon rf chips," in *International Conference on Recent Advances in Microwave Theory and Applications*, pp. 1–4, 2008. 42

[33] V. Carrubba, J. Lees, J. Benedikt, P. J. Tasker, and S. C. Cripps, "A novel highly efficient broadband continuous class-f RFpa delivering 74% average efficiency for an octave bandwidth," in *IEEE MTT-S Int. Microwave Symposium (IMS)*, pp. 1–4, June 2011. 42

[34] N. Tuffy, L. Guan, A. Zhu, and T. J. Brazil, "A simplified broadband design methodology for linearized high-efficiency continuous class-f power amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, pp. 1952–1963, June 2012. 42

[35] D. M. Pozar, *Microwave engineering*. Hoboken, NJ: Wiley, 4. ed. ed., 2012. Includes bibliographical references and index. 43

[36] B. S. Yarman, *Design of Ultra Wideband Power Transfer Networks*. JOHN WILEY & SONS INC, 2010. 43, 45

[37] R. Cameron, R. Mansour, and C. Kudsia, *Microwave Filters for Communication Systems: Fundamentals, Design and Applications*. Wiley, 2007. 43

[38] H. R. F., "Intermodulation distortion in kahn-technique transmitters," *IEEE Transactions on Microwave Theory and Techniques*, vol. 44, no. 12, pp. 2273–2278, 1996. 57

[39] H. Harada, H. Kashiwagi, Y. Toyozawa, N. Sonoda, and T. Yamaguchi, "Identification of Volterra kernels by using m-sequence which consists of arbitrary two values," in *SICE Annual Conference (SICE)*, pp. 1156–1159, Aug. 2012. 60, 71

[40] L. Ding, G. T. Zhou, D. R. Morgan, M. Zhengxiang, J. S. Kenney, K. Jaehyeong, and C. R. Giardina, "Memory polynomial predistorter based on the indirect learning architecture," in *IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 1, pp. 967–971, 2002. 62, 63

[41] K. Hyunchul and J. S. Kenney, "Behavioral modeling of nonlinear rf power amplifiers considering memory effects," *IEEE Transactions on Microwave Theory and Techniques*, vol. 51, no. 12, pp. 2495–2504, 2003. 62, 63

[42] H. Ruotsalainen, N. Leder, B. Pichler, H. Arthaber, and G. Magerl, "Equivalent complex baseband model for digital transmitters based on 1-bit quadrature pulse encoding," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 62, no. 11, pp. 2739–2747, 2015. 63

[43] W. Jiangfeng and L. R. Carley, "A table-based time-domain simulation method for oversampled microelectromechanical systems," in *IEEE/ACM International Workshop on Behavioral Modeling and Simulation*, pp. 37–41, 2000. 64

[44] M. Younes, O. Hammi, A. Kwan, and F. M. Ghannouchi, "An accurate complexity-reduced "plume" model for behavioral modeling and digital predistortion of rf power amplifiers," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 4, pp. 1397–1405, 2011. 64

[45] N. Leder, T. Faseth, H. Ruotsalainen, and H. Arthaber, "Characterization and modeling of pulse drivers for switch mode power amplifier measurements," in *IEEE Topical Conference on Power Amplifiers for Wireless and Radio Applications (PAWR)*, pp. 43–45, 2015. 68

[46] W. Kester, *Taking the Mystery out of the Infamous Formula, "SNR = 6.02N + 1.76dB," and Why You Should Care*. Analog Devices, 2009. 72

[47] O. Hammi, F. M. Ghannouchi, S. Boumaiza, and B. Vassilakis, "A data-based nested lut model for rf power amplifiers exhibiting memory effects," *IEEE Microwave and Wireless Components Letters*, vol. 17, no. 10, pp. 712–714, 2007. 76

[48] Y. Guo and L. Peng, "Lookup table based simulation and statistical modeling of sigma-delta adcs," in *ACM/IEEE Design Automation Conference*, pp. 1035–1040, 2006. 76

[49] N. Leder, B. Pichler, H. Ruotsalainen, T. Faseth, and H. Arthaber, "Nonlinear simulation of digital rf transmitters under modulated excitation," in *IEEE International Workshop on Integrated Nonlinear Microwave and Millimetre-wave Circuits (INMMiC)*, pp. 1–3, 2015. 82

[50] J.-Y. Tinevez, "A simple tree data structure in a matlab class." GitHub, 6 2014. 86

[51] J. E. Kleider and M. E. Humphrey, "Robust time-frequency synchronization for OFDM mobile applications," in *Proc. Fifth Int. Symp. Signal Processing and Its Applications ISSPA '99*, vol. 1, pp. 423–426 vol.1, 1999. 101

Appendices

Abbreviations

ADC	...	Analogue to digital converter
ARX	...	Auto regressive exogenous
BPG	...	Bit pattern generator
CPL	...	Coupled
CW	...	Continuous wave
DAC	...	Digital to analog converter
DBW	...	Design bandwidth
DC	...	Direct Current
DRFTx	...	Digital radio frequency transmitter
DSO	...	Digital storage oscilloscope
DUT	...	Device under test
EVM	...	Error vector magnitude
FET	...	Field-effect transistor
FFT	...	Fast fourier transform
FOM	...	Figure of merit
FWD	...	Forward
GaN	...	Gallium Nitrite
HEMT	...	High-electron mobility transistor
HFSS	...	High frequency structural simulator
HPR	...	Harmonic phase reference
IF	...	Intermediate frequency
IFFT	...	Inverse fast fourier transform
IIR	...	Infinite impulse response
LNA	...	Low noise amplifier
LO	...	Local oscillator
LP	...	Load pull
LTE	...	Long term evolution
LTI	...	Linear time-invariant
LUT	...	look-up table
LVDS	...	Low voltage differential signaling
MOD	...	Modulator
MWO	...	Microwave office
NL	...	Nonlinearity
NLIN	...	Nonlinear
NMSE	...	Normalized mean square error
NVNA	...	Nonlinear vector network analyser
OFDM	...	Orthogonal frequency division multiplex
PA	...	Power amplifier

PAE	...	Power added efficiency
PCB	...	Printed circuit board
PSD	...	Power spectral density
PWM	...	Pulse width modulation
Q	...	Quality factor
RET	...	Return
RF	...	Radio frequency
RMSE	...	Root mean square error
SDR	...	Software defined radio
SFDR	...	Spurious free dynamic range
SIM	...	Simulation
SMPA	...	Switched mode power amplifier
SNR	...	Signal to noise ratio
SPS	...	Samples per second
SRFT	...	Simplified real frequency technique
TRL	...	Through, reflect, line
TX	...	Transmitter
VNA	...	Vector network analyzer
VSS	...	Visual system simulator