Nonlinear Characterization of Wideband Microwave Devices and Dispersive Effects in GaN HEMTs

Sebastian Gustafsson
Nonlinear Characterization of Wideband Microwave Devices and Dispersive Effects in GaN HEMTs

SEBASTIAN GUSTAFSSON

© Sebastian Gustafsson, 2018

ISBN 978-91-7597-758-4

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 4439
ISSN 0346-718X

Technical report MC2-391
ISSN 1652-0769

Microwave Electronics Laboratory
Department of Microtechnology and Nanoscience – MC2
Chalmers University of Technology
SE-412 96 Gothenburg, Sweden
Phone: +46 (0) 31 772 1000

Printed by Chalmers Reproservice
Gothenburg, Sweden 2018
Abstract

Measurements play a key role in the development of microwave hardware as they allow engineers to test and verify the RF performance on a system, circuit, and component level. Since modern cellular standards employ complex modulation formats with wider signal bandwidths to cope with the growing demand of higher datarates, nonlinear characterization using wideband stimuli is becoming increasingly important. Furthermore, III-N semiconductor technologies such as gallium nitride (GaN) are to a larger extent utilized to enable higher performance in microwave circuits. However, GaN is highly frequency-dispersive due to trapping phenomena and thermal effects. This thesis deals with the development of nonlinear measurement instruments as well as characterization of dispersive effects in GaN high-electron-mobility transistors (HEMTs).

A measurement setup for wideband, nonlinear characterization of microwave devices has been designed and verified. The setup allows for simultaneous acquisition of low-frequency and radio-frequency signals from DC up to 4 GHz through the use of wideband signal generators and measurement receivers. This enables measurement scenarios such as multi-band load-pull and large-signal characterization of IQ-mixers, which can give useful insight into how to optimize the performance in a RF transmitter.

Electrothermal characterization of GaN devices has been carried out using conventional measurement methods such as pulsed I-V, and it is shown that trapping phenomena and thermal effects due to self-heating or mutual coupling are challenging to separate. Multiple methods must be utilized to fully characterize both the large-signal and small-signal impact on device performance. A new characterization method has been developed, for extraction of thermal transfer functions between mutually coupled devices on e.g. a semiconductor wafer. The method does not require any DC-bias on the measured devices, which can potentially reduce the influence of trapping during characterization of thermal properties in materials.

Keywords: large-signal, microwave, wideband, GaN, HEMT, dispersive effects, electron trapping, thermal coupling.
List of Publications

Appended Publications

This thesis is based on the following papers:


Other Publications

The following papers have been accepted for publication but are not included in the thesis. The content partially overlaps with the appended papers or is out of the scope of this thesis.


**Thesis**


As part of the author’s doctoral studies, some of the work presented in this thesis has previously been published in [l]. Figures, tables and text from [l] may therefore be fully or partly reproduced in this thesis.
Notations and abbreviations

Notations

A_k \quad \text{Amplitude coefficients}
\nA_i \quad \text{Traveling voltage wave at port } i
\nA_b \quad \text{Traveling voltage wave at port } i
\nc_i \quad \text{Polynomial coefficients}
\non \quad \text{Drain-source capacitance}
\nC_{Ts} \quad \text{Trap capacitance in admittance dispersion model}
\nD_1 \quad \text{Activation energy of surface-related trap}
\nD_2 \quad \text{Activation energy of buffer-related trap}
\ne_{ij} \quad \text{Error coefficients}
\nf_c \quad \text{Carrier frequency}
\nf_{IF} \quad \text{Intermediate frequency}
\nf_k \quad \text{Frequency of the } k^\text{th} \text{ tone}
\nf_{RF} \quad \text{Radio frequency}
\ns \quad \text{Sampling frequency}
\ng_{ds} \quad \text{Output conductance}
\nH_{TH} \quad \text{Thermal transfer function}
\ID \quad \text{Drain-to-source current}
\i_{LP} \quad \text{Low-frequency current}
\kj \quad \text{Complex correction factor for channel } i \text{ at voltage range } x
\nl \quad \text{Harmonic index}
\nm_i \quad \text{Measured voltage at port } i
\ns \quad \text{Number of samples}
\nP_{DC} \quad \text{DC power}
\nP_{diss} \quad \text{Dissipated power}
\nP_{in} \quad \text{Input power}
\nP_{out} \quad \text{Output power}
\nP_{ref} \quad \text{Reference power}
\nr_{mod} \quad \text{Resistance modulation}
\nRO \quad \text{ON-resistance}
\nRT \quad \text{Trap resistance in admittance dispersion model}
\nRT_{TH} \quad \text{Thermal Resistance}
\nr_{T0} \quad \text{Nominal sensor resistance}
$R_{T1}$  Linear temperature coefficient
$S_{ij}$  S-parameters
$S_{LF}$  Apparent low-frequency power
$T_{DUT}$  T-parameter matrix of DUT
$T_i$  T-parameter matrix of error model at port $i$
$T_m$  Measured T-parameter matrix
$T$  Temperature
$T_A$  Ambient temperature
$T_P$  Pulse period
$T_{PW}$  Pulse width
$V_{DG}$  Drain-to-gate voltage
$V_{DS}$  Drain-to-source voltage
$V_{DS,Q}$  Quiescent Drain-to-source voltage
$V_{GS}$  Gate-to-source voltage
$V_{GS,Q}$  Quiescent Gate-to-source voltage
$v_{LF}$  Low-frequency voltage
$v_{sens}$  Sensor voltage
$x(t)$  Input signal, time domain
$X(f)$  Input signal, frequency domain
$y(t)$  Output signal, time domain
$Y(f)$  Output signal, frequency domain
$Y_{ij}$  Y-parameters
$Z_0$  Characteristic impedance
$Z_{TH}$  Thermal impedance
$\alpha$  Absolute magnitude and phase error coefficient
$\alpha_r$  Resistance transfer function
$\gamma$  Current scaling factor
$\Gamma_L$  Load reflection coefficient
$\Gamma_m$  Measured load reflection coefficient
$\Delta f$  Frequency resolution
$\phi_k$  Phase of the $k^{th}$ tone
$\omega$  Angular Frequency
$\omega_0$  Fundamental angular frequency
$\omega_n$  Angular frequency of the $n^{th}$ tone
$\omega_T$  Characteristic angular frequency of trap
$\omega_{LF}$  Angular frequency of low-frequency signal
$\omega_{RF}$  Angular frequency of radio frequency signal
$\omega_{LO}$  Angular frequency of local oscillator
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>5G</td>
<td>5th Generation Wireless Systems</td>
</tr>
<tr>
<td>AC</td>
<td>Alternating Current</td>
</tr>
<tr>
<td>ACPR</td>
<td>Adjacent Channel Power Ratio</td>
</tr>
<tr>
<td>ADC</td>
<td>Analog-to-Digital Converter</td>
</tr>
<tr>
<td>AMPS</td>
<td>Advanced Mobile Phone System</td>
</tr>
<tr>
<td>AWG</td>
<td>Arbitrary Waveform Generator</td>
</tr>
<tr>
<td>C</td>
<td>Carbon</td>
</tr>
<tr>
<td>CAD</td>
<td>Computer-Aided Design</td>
</tr>
<tr>
<td>CDMA</td>
<td>Code Division Multiple Access</td>
</tr>
<tr>
<td>CMOS</td>
<td>Complementary-Metal-Oxide-Semiconductor</td>
</tr>
<tr>
<td>CW</td>
<td>Continuous Wave</td>
</tr>
<tr>
<td>DAC</td>
<td>Digital-to-Analog Converter</td>
</tr>
<tr>
<td>DC</td>
<td>Direct Current</td>
</tr>
<tr>
<td>DC-HSDPA</td>
<td>Dual Cell High Speed Downlink Packet Access</td>
</tr>
<tr>
<td>DCT</td>
<td>Drain Current Transient</td>
</tr>
<tr>
<td>DUT</td>
<td>Device Under Test</td>
</tr>
<tr>
<td>IF</td>
<td>Intermediate Frequency</td>
</tr>
<tr>
<td>Fe</td>
<td>Iron</td>
</tr>
<tr>
<td>FFT</td>
<td>Fast Fourier Transform</td>
</tr>
<tr>
<td>FOM</td>
<td>Figure of Merit</td>
</tr>
<tr>
<td>GaAs</td>
<td>Gallium Arsenide</td>
</tr>
<tr>
<td>GaN</td>
<td>Gallium Nitride</td>
</tr>
<tr>
<td>GSM</td>
<td>Global System for Mobile Communications</td>
</tr>
<tr>
<td>HEMT</td>
<td>High Electron Mobility Transistor</td>
</tr>
<tr>
<td>HSDPA</td>
<td>High Speed Downlink Packet Access</td>
</tr>
<tr>
<td>HSPA</td>
<td>High Speed Packet Access</td>
</tr>
<tr>
<td>IM3</td>
<td>Third Order Intermodulation</td>
</tr>
<tr>
<td>I-V</td>
<td>Current-Voltage</td>
</tr>
<tr>
<td>LF</td>
<td>Low-Frequency</td>
</tr>
<tr>
<td>LO</td>
<td>Local Oscillator</td>
</tr>
<tr>
<td>LCM</td>
<td>Least Common Multiple</td>
</tr>
<tr>
<td>LRM</td>
<td>Line-Reflect-Match</td>
</tr>
<tr>
<td>LSNA</td>
<td>Large Signal Network Analyzer</td>
</tr>
<tr>
<td>LTE</td>
<td>Long Term Evolution</td>
</tr>
<tr>
<td>LTI</td>
<td>Linear and Time-Invariant</td>
</tr>
<tr>
<td>MTA</td>
<td>Microwave Transition Analyzer</td>
</tr>
<tr>
<td>NVNA</td>
<td>Nonlinear Vector Network Analyzer</td>
</tr>
<tr>
<td>OFDM</td>
<td>Orthogonal Frequency-Division Multiplexing</td>
</tr>
<tr>
<td>op-amp</td>
<td>Operational Amplifier</td>
</tr>
<tr>
<td>PA</td>
<td>Power Amplifier</td>
</tr>
<tr>
<td>PAPR</td>
<td>Peak-to-Average Power Ratio</td>
</tr>
<tr>
<td>PISPO</td>
<td>Periodic In Same Periodic Out</td>
</tr>
<tr>
<td>PRF</td>
<td>Pulse Repetition Frequency</td>
</tr>
<tr>
<td>PSU</td>
<td>Power Supply Unit</td>
</tr>
<tr>
<td>RF</td>
<td>Radio Frequency</td>
</tr>
<tr>
<td>RTO</td>
<td>Real Time Oscilloscope</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>SDD</td>
<td>Symbolically Defined Device</td>
</tr>
<tr>
<td>Si</td>
<td>Silicon</td>
</tr>
<tr>
<td>SiC</td>
<td>Silicon Carbide</td>
</tr>
<tr>
<td>SiGe</td>
<td>Silicon Germanium</td>
</tr>
<tr>
<td>SINAD</td>
<td>Signal-to-Noise and Distortion Ratio</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>SOI</td>
<td>Silicon-on-Insulator</td>
</tr>
<tr>
<td>SOL</td>
<td>Short-Open-Load</td>
</tr>
<tr>
<td>SOLR</td>
<td>Short-Open-Load-Reciprocal</td>
</tr>
<tr>
<td>SOLT</td>
<td>Short-Open-Load-Through</td>
</tr>
<tr>
<td>TRL</td>
<td>Through-Reflect-Line</td>
</tr>
<tr>
<td>VGA</td>
<td>Variable Gain Amplifier</td>
</tr>
<tr>
<td>VNA</td>
<td>Vector Network Analyzer</td>
</tr>
<tr>
<td>WCDMA</td>
<td>Wideband Code Division Multiple Access</td>
</tr>
<tr>
<td>WLAN</td>
<td>Wireless Local Area Network</td>
</tr>
</tbody>
</table>
# Contents

Abstract iii  
List of Publications v  
Notations and Abbreviations ix

## 1 Introduction 1

1.1 The Evolution of Wireless Communication Systems 1  
1.2 Measurement-aided Design of Microwave Hardware 2  
1.3 Semiconductor Materials for Next Generation Applications 3  
1.4 Thesis Contributions and Outline 4

## 2 Nonlinear RF Characterization Techniques 7

2.1 Two-port RF Networks 7  
2.2 RF Characterization of Nonlinear Two-port Networks 10  
   2.2.1 Sampler-based Architectures 11  
   2.2.2 Mixer-based Architectures 13  
   2.2.3 Direct RF Sampling Architectures 13  
   2.2.4 Discussion 14  
2.3 Wideband Nonlinear Measurement Setup 15  
   2.3.1 Signal Acquisition and Sensing 15  
   2.3.2 Signal Generation 17  
   2.3.3 Calibration Procedure 18  
   2.3.4 Oscilloscope Correction Algorithms 21  
   2.3.5 Measurement Setup Verification 24  
2.4 Chapter Summary 27

## 3 Applications of Wideband Nonlinear Characterization 29

3.1 Multi-band Load-pull for PA Design 30  
3.2 The Effect of LF Impedance Termination on Transistor Linearity 32  
3.3 Large-signal IQ-mixer Characterization 34  
3.4 Chapter Summary 37

## 4 Characterization of Dispersive Effects in GaN HEMTs 39

4.1 Sources of Dispersion 40  
   4.1.1 Electrical Effects 40  
   4.1.2 Thermal Effects 41  
4.2 Pulsed I-V 42
CONTENTS

4.2.1 Pulsed I-V for Trap Characterization ............... 42
4.2.2 Pulsed I-V for Thermal Characterization .......... 44
4.3 Transient Methods ................................... 45
  4.3.1 Drain Current Transient Measurements for Trap Characterization .............. 45
  4.3.2 Transient Lateral Thermal Coupling ............... 46
4.4 Small-signal AC measurements ........................ 47
  4.4.1 Admittance Dispersion .......................... 47
  4.4.2 Thermal Impedance Extraction from Z-parameters ................ 49
4.5 Characterization of Thermal Coupling using a Modified 3ω-method ......... 49
  4.5.1 Theory ......................................... 51
  4.5.2 Method Verification .............................. 52
  4.5.3 Thermal Coupling in GaN-on-SiC .................. 52
4.6 Chapter Summary ...................................... 54

5 Conclusions and Future Work .......................... 55
  5.1 Conclusions ....................................... 55
  5.2 Future Work ...................................... 56

6 Summary of appended papers .......................... 59

Acknowledgments ........................................ 61
Chapter 1

Introduction

Wireless technology has revolutionized how we interact with each other and the world, and has brought us many technical innovations over the past 100 years. Navigation systems such as the global positioning system (GPS) enable accurate cartography, faster responses of emergency services, and high-efficiency agriculture among others. Object-detection systems based on radar are instrumental for the future vision of fully automated vehicles, and emerging wireless applications include wireless power transfer and medical implants with wireless health monitoring. With regards to communication, wireless information transmission began as simple telegraph messages, but now enable us to access the global system of interconnected computers we call internet from anywhere in the world.

1.1 The Evolution of Wireless Communication Systems

Portable wireless communication has evolved from being a simple link between individuals, to a global inter-connected network where people can share their entire life in real-time. With the fifth generation (5G) communication systems currently being standardized, the world is once again entering a new phase of mobile communication. Although the specifications and goals of 5G are slightly broader than previous cellular standards, higher data throughput remains one of the driving forces, with peak data rate requirements of 20 Gb/s [1]. Such data rates are partly enabled by the utilization of wideband and spectrally efficient communication signals.

As shown in Fig. 1.1, the signal bandwidths of cellular standards have increased dramatically over the years in order to meet data rate requirements, with a roughly ten-fold increase every six years. Due to the scarcity of non-allocated spectrum around conventional carrier frequencies, new frequency bands are being defined at mm-wave frequencies and beyond [1, 2]. Multiple frequency bands are also being combined at several different carrier frequencies through carrier aggregation techniques [3]. With regards to the increased spectral efficiency, modern cellular standards rely heavily on higher order modulation formats and frequency division multiplexing techniques such as
orthogonal frequency-division multiplexing (OFDM), which tend to increase the peak-to-average power ratio (PAPR) of the communication signals [2, 4]. With such signals, in order to preserve the average output power in a microwave transceiver, the linearity in the hardware must be improved to leave headroom for high power swings. One of the other demands put forward by the 5G standardization is that the energy consumption in the networks should be lower [5], which in turn leads to requirements of energy-efficient hardware. This, in combination with the increase in signal complexity, results in the fact that wireless systems need to be very linear and efficient across a large frequency range.

1.2 Measurement-aided Design of Microwave Hardware

The design of high-performance radio frequency (RF) hardware is reliant on comprehensive component models in computer-aided design (CAD) tools. Models are required to reduce the time-to-market and to narrow product design cycles. For development and verification of these models, input data in the form of accurate measurements is needed [13].

Microwave measurements are challenging from many aspects. Parasitic effects distort the measured signals, transmission lines shift measurement reference planes, and different applications require characterization of the device performance under various operating conditions. Take a critical transmitter component such as the power amplifier (PA) as an example. The design of a high-efficiency PA requires a comprehensive analysis of its nonlinear time-domain behavior and how the performance is affected by input and output loading conditions at fundamental, harmonic, and intermodulation frequencies [14]. Following the trends in communication signals, this analysis needs to be carried out using wideband stimuli. The utilization of wideband communication signals can trigger dynamic phenomena, which originate both from surrounding electrical networks and from transistors, and can cause spectral regrowth and time-varying performance.

Measurement instruments such as the nonlinear vector network analyzer (NVNA) are becoming more and more vital in microwave measurement labs due
to the aforementioned needs. These instruments have traditionally focused on multi-harmonic nonlinear continuous-wave (CW) measurements, which provide great insight into high efficiency operation modes of PAs, model extraction, and model validation [14–16]. However, since the trend in wireless systems is pointing towards communication signals with GHz-bandwidths, there is also a need for measurement instruments with wideband capabilities, something which is ongoing research [17–20].

In addition to pure microwave measurements, there is also a need for methods which focuses on the characterization of electrical and thermal effects in transistors. These effects are typically referred to as memory effects or dispersion. They degrade the RF performance and reliability, and determination of underlying causes requires characterization methods which can e.g. separate electrical effects due to trapping and thermal effects due to mutual- and self-heating. In certain semiconductor materials, such as gallium nitride (GaN), memory effects are more pronounced due to high power densities and still a high density of material defects. Characterization of such effects in GaN are therefore of particular importance.

1.3 Semiconductor Materials for Next Generation Applications

The selection of device technology is critical in order to meet requirements of next generation wireless systems. Silicon (Si) technologies such as silicon-on-insulator (SOI), complementary-metal-oxide-semiconductor (CMOS), and silicon germanium (SiGe) are preferable for low-cost applications and high integration due to compatibility with other Si-circuitry. Output power at high-frequency operation is however rather lacking due to the low electric breakdown field and low electron mobility in these technologies. For applications requiring high output power at high frequencies, such as base stations or radars, III-V semiconductor compounds such as gallium arsenide (GaAs) and GaN are instead interesting candidates. Of the two, GaN offers superior performance due to a higher breakdown voltage and that the electron velocity saturates at a higher electric field [21,22].

Although GaN often is considered ideal for high-power circuits, other applications and use areas can also benefit from GaN-integration. Due to the high power density, smaller transistor sizes can be utilized in comparison to Si or GaAs, translating to more compact circuits and higher integration [23]. Smaller transistors also imply higher device impedances and lower parasitic capacitances, which increase the bandwidth and efficiency of the implemented circuit. GaN-based circuits can therefore also result in lightweight systems with lower power consumption [23]. This is not only of interest for communication systems, but also space [24] and defense [25,26] applications.

Despite being considered a relatively mature technology by now, GaN faces a number of challenges. High power densities and miniaturization of circuits inevitably lead to problems with thermal management. Although the efficiency and maximum operating temperature for GaN-based transistors is relatively high, thermal throttling is widely considered as one of the limitations on the achievable output power [27,28]. Another issue is the formation of so-called
trapping centers throughout the material, which causes dynamic variations in the transistor performance on both fast and slow time-scales, with lower-than-expected RF output powers and unreliable performance as a result [29]. These time-varying effects can be enhanced when being subject to wideband stimuli, and can be challenging to compensate for [30]. Both the thermal performance and trapping effects highly depend on the overall epitaxial structure of a GaN high-electron mobility transistor (HEMT). In order to provide epitaxial suppliers with valuable design feedback, and for implementation in models, characterization of the various electrical and thermal effects in GaN HEMTs is of high importance.

To summarize, the research and development of emerging circuit and device solutions depend on characterization tools which provide accurate measurements under relevant operating conditions. This includes methods for quantification of dispersive effects and nonlinear characterization using wideband stimuli.

1.4 Thesis Contributions and Outline

This thesis focuses on nonlinear characterization of microwave circuits and components. These nonlinearities are studied using a custom-built measurement setup capable of wideband signal injection. Furthermore, the difficulty of characterizing electrical and thermal effects in GaN HEMTs is highlighted through the use of common measurement techniques. A new method for measuring the mutual thermal coupling between adjacent devices has been developed, which potentially minimizes the influence of electrical effects due to trapping phenomena during thermal measurements.

In Chapter 2, the limitations of conventional RF characterization methods such as S-parameters are discussed. To know the true operation of the device under test (DUT), the nonlinear time-domain information is needed. Furthermore, wideband characterization is of increasing importance in order to study DUT behavior using communication-like signals. Wideband nonlinear characterization is in this thesis enabled by a custom-built measurement setup equipped with simultaneous low-frequency (LF) and RF measurement capabilities. The proposed setup is introduced towards the end of Chapter 2, and comparisons with other nonlinear measurement setups are made. The calibration of the setup is discussed, as well as a correction algorithm for the RF receivers, which increases the usable dynamic range (paper [B]).

In Chapter 3, the wideband nonlinear measurement setup is used for characterization of common RF transmitter-chain components. The wideband capabilities of the proposed setup enables interesting measurement applications. This is exemplified with multi-band active load-pull (paper [A]), LF impedance termination on transistor linearity (paper [D]), and large-signal characterization of IQ-mixers (paper [E]).

In Chapter 4, electrical and thermal effects, commonly referred to as memory effects or dispersion, are studied in GaN HEMTs. An overview of common measurement techniques for dispersion characterization is given. Limitations and challenges related to the separation of electrical and thermal effects are discussed. The measurement methods are exemplified using the material studies presented in paper [C] and paper [F]. Furthermore, a new measurement
method, for characterization of thermal coupling between adjacent devices, is presented. This method is used in paper [F] together with the wideband nonlinear measurement setup to extract transfer functions between thermally coupled semiconductor resistors on three different GaN materials.

The thesis is finally summarized in Chapter 5, where important conclusions from previous chapters are detailed. A future outlook of the research field is given as well.
Chapter 2

Nonlinear RF Characterization Techniques

RF measurements are needed for performance evaluation of microwave circuits and components as well as for model extraction and verification. The requirements on the measurement instrumentation highly depend on what type of circuit is being analyzed and under what conditions it should operate, for example what type of input signal is used. In this chapter, the limitations of linear network parameters such as scattering parameters (S-parameters) are discussed. For a better description of how circuits and components actually behave, nonlinear measurements are needed and therefore an overview of nonlinear characterization setups is given. A new nonlinear measurement setup capable of wideband device characterization is presented towards the end, enabling the measurement applications presented in Chapter 3.

2.1 Two-port RF Networks

The primary goal of an electrical measurement is to gain insight about the operation of a certain network. Consider the two-port network shown in Fig. 2.1, labeled device under test (DUT). The output signal, \( y \), will depend on the input signal, \( x \), and how it is being transformed by the network, described by a

![Figure 2.1](image)

**Figure 2.1:** Illustration of how the linear frequency response of a two-port network introduces a magnitude offset and phase shift on a sinusoidal stimulus signal.
function, $F(X)$. A relatively straight-forward way to measure the input-output characteristics of this network would be to connect a signal generator to the input, and measure the output power with a power meter. Although this would in some cases work perfectly fine, some questions arise: How much power is actually delivered to the DUT, what is the phase delay of the circuit, and how does it behave when the source and load impedance diverge from the system impedance? These questions can be answered by measuring the full two-port network parameters of the DUT. In the RF domain, S-parameters are a well-established concept of describing the characteristics of an electrical network. In contrast to measuring ratios of currents and voltages, as done with impedance (Z) and admittance (Y) parameters, S-parameters describe ratios of travelling voltage waves frequency-by-frequency, as illustrated in Fig. 2.2, and are defined as follows

\[
\begin{align*}
S_{11}(\omega) &= \frac{b_1(\omega)}{a_1(\omega)} \bigg|_{a_2(\omega) = 0} \\
S_{12}(\omega) &= \frac{b_1(\omega)}{a_2(\omega)} \bigg|_{a_1(\omega) = 0} \\
S_{21}(\omega) &= \frac{b_2(\omega)}{a_1(\omega)} \bigg|_{a_2(\omega) = 0} \\
S_{22}(\omega) &= \frac{b_2(\omega)}{a_2(\omega)} \bigg|_{a_1(\omega) = 0} ,
\end{align*}
\]

where $a_1$ and $b_1$ are the forward and backward travelling waves at port 1, respectively, $a_2$ and $b_2$ are the forward and backward travelling waves at port 2, respectively, and $\omega$ is the angular frequency. The forward travelling waves, $a_1$ and $a_2$, can be seen as stimuli signals and the backward travelling waves, $b_1$ and $b_2$, can be seen as responses of the network. When measuring the S-parameters of a circuit, the non-excited port is terminated in the system impedance, so that $a_i = 0$. This can be advantageous compared to measuring the Z- or Y-parameters, which instead require either an open or short termination on the non-excited port. Due to parasitics, an open or short condition can be hard to maintain across a large RF bandwidth. Furthermore, active devices such as transistors are more prone to oscillations with an open or shorted input/output [31].

Now that a framework for characterizing the two-port DUT is in place, its validity must be considered. S-parameters work under the assumption that the DUT is linear and time-invariant (LTI). Consequently, the network response of the DUT can be extracted from combinations of subsequent measurements at different frequencies, and the stimulus signal power level can mostly be ignored. However, consider the following DUT, with the output signal, $y(t)$, described
by a polynomial expression as
\[ y(t) = c_1 x(t) + c_2 x(t)^2 + c_3 x(t)^3, \] (2.2)
where \( c_1, c_2, \) and \( c_3 \) are arbitrary but static first, second and third order coefficients. The first order component, \( c_1 x(t) \), corresponds to the linear amplification of the input signal. The second order term, \( c_2 x(t)^2 \), will, with a sinusoidal stimulus, generate a direct current (DC) component and a frequency component at two times the stimulus frequency. Similarly, the third order term, \( c_3 x(t)^3 \), will generate frequency components at three times the stimulus frequency, but also components at the stimulus frequency.

This nonlinear DUT will not be represented correctly with S-parameters, as illustrated in Fig. 2.3. As the input power is increased, the measured \( S_{21} \) will deviate from its linear approximation. To solve this, one can introduce power-dependent S-parameters, but this will also result in an unsatisfying description of the DUT since the DC and harmonic content is not taken into account. In addition to a power-dependency, ratios of travelling waves at different frequencies or nonlinear figure-of-merits (FOMs) such as 1 dB compression point can be introduced, but will still not result in a satisfying description of the two-port network. This is largely due to the fact that the network will show very different responses to different kinds of stimuli, making FOMs a fairly useless metric for describing the network behavior [31]. Instead, the traveling waves should be measured directly, in contrast to ratios, and the two-port networks we want to characterize should be limited to a certain class of systems. This thesis focuses on the characterization of periodic in, same periodic out (PISPO) systems, which means that the response of a network will have the same periodicity as the stimulus signal. Chaotic systems are an example of non-PISPO systems and are therefore not covered. Some dispersive effects such as electron charging and discharging phenomena will however be discussed in later chapters.

Unlocking the ability to measure un-ratioed travelling wave quantities requires in principle two things; absolute quantification of the magnitude of the travelling wave, and relative cross-frequency phase information in order to relate different frequencies to one another. While the former can intuitively be considered necessary to be able to measure the correct power at each frequency, the latter is equally important for accurately characterizing a nonlinear DUT, as they together enable the possibility to obtain the time-domain equivalent waveform of the measured quantities. In Fig. 2.4, two signals with different relative phases, but the same magnitude spectrum and thus same average power, are shown, resulting in time domain waveforms with vastly different
CHAPTER 2. NONLINEAR RF CHARACTERIZATION TECHNIQUES

Figure 2.4: Two signals having the same magnitude spectrum but different relative cross-frequency phases, resulting in significantly different time domain equivalents.

Figure 2.5: Illustration of how dynamic effects in an amplifier can affect the output spectrum. As an example, supply dynamics can cause sideband distortion asymmetry.

PAPRs. With time-domain measurement capabilities, the true operation of the DUT can be analyzed, which is of utmost importance when for example designing and characterizing high efficiency amplifiers, where the output power performance and efficiency is to a large extent determined by the shape of the current and voltage waveforms [14].

In addition to measuring the time-domain waveforms, it is increasingly important to characterize RF circuits using wideband stimuli. Dynamic effects in the circuits can cause spectral regrowth or distortion that varies depending on the input signal type and frequency content. The origin of such effects can be on a transistor level, in the form of electron trapping and discharging phenomena, or bias modulation due to varying temperature in the device [29]. It can also be on a circuit level if the bias network is not properly decoupled, as illustrated in Fig. 2.5, or due to the frequency response in matching networks [29]. Dynamic effects can easily be underestimated or invisible when using CW or two-tone stimuli, and therefore a measurement scenario using communication-like signals should always be considered [31].

2.2 RF Characterization of Nonlinear Two-port Networks

With the conclusion that time-domain waveforms are important to measure in order to understand the true operation of the DUT, RF instruments and measurement tools must be adapted accordingly to provide such information. At low frequencies, time domain voltage and current waveform measurements are relatively straight-forward, by means of oscilloscopes and I/V probes. Higher up in frequency, however, parasitics and phase delays will eventually distort the measured signals to such an extent that the acquired data is unusable [31].
2.2. RF CHARACTERIZATION OF NONLINEAR TWO-PORT NETWORKS

RF engineers have therefore over the years developed measurement hardware, in combination with extensive calibration routines, which enables accurate characterization of nonlinear two-port networks at microwave frequencies and beyond. Two-port characterization setups have more or less converged to a common base-architecture, as depicted in Fig. 2.6 [31,32]. The essentials of this architecture are: signal sources for signal injection, couplers for signal sensing, an optional down-conversion stage, and analog-to-digital converters (ADCs) for signal digitization. Since the time-domain waveforms are of interest, the obvious choice would be to directly digitize the signals at the coupled coupler ports and apply calibration techniques to obtain the waveforms at the DUT reference planes, thereby bypassing the need of any downconversion. Depending on the requirements of the measurements, this approach can be sufficient. For example, if the nonlinear network response to a wideband (>500 MHz) stimuli around a 2 GHz carrier frequency is to be measured, direct acquisition architectures can be the proper choice. On the other hand, if the third harmonic of a 20 GHz fundamental signal is of interest, such architectures lack in performance with today's ADC technology. The next sections will therefore give an overview of three different categories of nonlinear RF measurement setups: sampler-based, mixer-based and direct RF sampling setups.

2.2.1 Sampler-based Architectures

Sampler-based RF receivers were pioneered by the Hewlett-Packard microwave transition analyzer (MTA), which provided vectorial measurements with preserved relative cross-frequency phases up to 40 GHz [33]. Although many variants of two-port measurement setups were built around the MTA [34–36], nonlinear RF characterization was popularized by the large-signal network analyzer (LSNA) [32,37]. The general architecture for a sampler-based setup is shown in Fig. 2.7. The coupled RF signals are down-converted to IF by either a sampling gate or a track-and-hold amplifier, which is triggered by a train of short pulses, where the pulse repetition frequency (PRF) commonly is several orders of magnitude lower than the frequency of the coupled RF signal. This technique therefore also known as sub-sampling. By knowing the frequency content of the measured signal and by choosing the PRF of the pulse train carefully, a multi-harmonic RF signal can be converted to IF and later digitized by the ADCs in one acquisition. In time-domain it can be seen as
a successive sampling of the waveform, where the sampling instance moves forward relative to the RF signal for each triggering of the sampling gate, visualized in Fig. 2.8. Since the entire frequency content of the RF signal is captured in one measurement, the relative cross-frequency phases are in theory preserved, although a phase calibration is in practice often needed [38]. In order for the sub-sampling to work properly, the trigger pulses must be very short so that they can be approximated as a train of Dirac pulses which have a uniform amplitude spectrum. The spectrum of the Dirac pulses is convoluted with the measured signal spectrum, and the down-converted multi-harmonic RF signal appears at the following frequencies

$$f_{IF,l} = l \text{mod}(f_{RF}, f_s),$$  \hspace{1cm} (2.3)

where $l$ denotes the harmonic index of the RF signal, $f_{RF}$ contains the signal frequencies, mod is the modulo operator, and $f_s$ is the PRF of the Dirac pulse train [32]. The modulo operation acts as a form of spectral compression, allowing multi-GHz bandwidth signals to be captured by ADCs with MHz-sampling rates. The above expression is only valid for multi-harmonic RF signals, but different techniques exist to cover measurements using multi-tones which are not harmonically related [39, 40]. However, wideband communication signals can be harder to fit within the limited ADC bandwidth of a typical sampler-based setup, if they are not approximated by sparse multi-tones [31].
2.2. RF CHARACTERIZATION OF NONLINEAR TWO-PORT NETWORKS

Mixer-based nonlinear two-port measurement setups can be seen as an extension of the conventional vector network analyzer (VNA), since they are both based on the same down-conversion principle. As depicted in Fig. 2.9, the coupled RF signals are mixed with a local oscillator (LO) to generate an IF signal which can be digitized by the ADCs. Since the mixing process is ideally a linear translation of the frequency content from RF to IF, and due to the narrow bandwidth of the ADCs, multiple harmonics of a signal must be captured by subsequent measurements [38]. This is in contrast to sampler-based systems which compress the spectrum of the measured signal, allowing for a single acquisition. Without any control or knowledge of the LO phase, every time the down-conversion is performed, the frequencies in the resulting IF signal will be contaminated with an unknown phase. The added phase from the LO will not affect the magnitude spectrum, but the time-domain information is corrupted. This is commonly addressed by using a fifth reference receiver which relates all of the measurements to the known phase of a multi-harmonic signal. This phase reference is often a comb-generator which outputs a very narrowband pulse, resulting in a harmonically-rich spectrum with ideally flat amplitude [41].

The majority of commercially available nonlinear characterization setups are nowadays built around the mixer-based architecture, such as the Keysight PNA-X [42], Maury Microwave MT2000 [19], Rohde & Schwarz ZVA [43] and Mesuro RAPID [18]. The dynamic range in mixer-based setups are generally better than in sampler-based setups [38], and the lack of commercial development in sampler-based setups has made them less suitable for measurements using modern communication signals.

2.2.3 Direct RF Sampling Architectures

Direct RF sampling removes the need of any intermediary frequency-conversion, as shown in Fig. 2.10, where the digitization often is done with a four-port oscilloscope. This can either be a real-time oscilloscope (RTO) or an equivalent-time oscilloscope. The latter has a similar working principle as the sampler-based measurement architecture and is therefore considered redundant and
Figure 2.10: Generalized schematic of a direct RF sampling measurement setup. The coupled RF signals are directly digitized by the measurement receivers. Under the assumption that the sampling rate is at least twice the highest frequency content of the signal, the time-domain signal can be directly retrieved.

not covered in this thesis. RTOs capture the entire time-domain signal in a single measurement, leading to preserved cross-frequency phases and large instantaneous bandwidths, with high-end models being able to measure up to 100 GHz at sampling rates of up to 240 GSa/s [44]. The single-shot acquisition of the entire bandwidth makes the direct RF sampling architecture suitable for non-periodic stimuli signals, although periodic signals are often preferred for waveform averaging and for less leakage in the fast Fourier transform (FFT) processing [45]. The large bandwidth comes however at the cost of increased noise power, limited dynamic range in the ADCs, and potential problems related to timebase jitter [46]. The dynamic range can be improved by either waveform averaging or by increasing the measurement time, but will in the end be limited by spurious signals and the linearity of the ADC.

2.2.4 Discussion

The presented measurement architectures all come with advantages and disadvantages, and even though there are techniques to improve on shortcomings in each architecture, there exists fundamental differences that need to be discussed.

First of all, the user of the measurement setup must define the goal and objectives with the measurement. In other words, the measurement scenario must be well-defined beforehand. If the intention is to design an efficient power amplifier at e.g. X-band (8-12 GHz), which requires very precise waveform engineering through specific termination impedances, the higher order harmonics must be captured with high accuracy [47, 48]. The natural choice would therefore be to use a sampler-based or a mixer-based setup. The mixer-based setup would have a slightly better dynamic range, because of noise folding in the sampler-based setup. However, the sampler-based setup does not require a fifth receiver for phase tracking, which relaxes the system complexity slightly.

If the intention on the other hand is to design a very wideband power amplifier, or to investigate dynamic effects across a large bandwidth, the choice becomes less trivial. Sampler-based systems typically have narrowband ADCs and the spectral compression properties of the subsampling operation could limit the choice of input stimulus [31]. Methods for wideband operation of sampler-based setups exist [39], but nonlinear characterization using a wideband multitone signal could suffer from limited dynamic range due to the noise folding. Mixer-based setups could be a better choice since the dynamic range
would be sufficient for wideband nonlinear characterization, and techniques exist for overcoming limitations in ADC bandwidth [49, 50]. For very wide signal bandwidths, direct RF sampling setups can be a suitable choice. This architecture can also offer the user a lot of flexibility, since both periodic and non-periodic behavior can be investigated.

For the commercial aspect of developing and selling an RF measurement instrument, it makes sense to base it on a mixer-based architecture since it offers good linearity and dynamic range, with relatively wide signal bandwidths at high carrier frequencies. Direct RF sampling architectures are nevertheless of interest to use academically for the flexibility they offer. Furthermore, the fast-paced development in ADCs could in the future open up for direct RF sampling architectures with very high dynamic range [51, 52], or low-cost instruments due to the decreased complexity compared to e.g. a mixer-based solution.

### 2.3 Wideband Nonlinear Measurement Setup

As touched upon in the discussion in the previous section, direct RF sampling architectures offer wide instantaneous bandwidths, which can be of interest for wideband circuit design and for studying dynamic effects in microwave devices. Although the dynamic range can be limited, the flexibility in a direct RF sampling architecture opens up for interesting measurement scenarios. With regards to commercial instruments, software or hardware limitations in such setups can potentially restrict their use for some measurement cases. A nonlinear measurement setup based on a direct RF sampling architecture has therefore been designed for characterization of components and circuits using wideband stimuli. This section is dedicated for describing the measurement setup in terms of hardware, calibration and verification.

#### 2.3.1 Signal Acquisition and Sensing

The wideband nonlinear measurement setup is based on a direct RF sampling architecture, as shown in Fig. 2.11. The measurement setup can be split up into two independent parts; one RF path and one low-frequency (LF) path, with each path having its own set of measurement receivers and calibration procedures. For the RF path, the measurement receiver is a four-channel Rohde & Schwarz RTO1044 real-time oscilloscope with 4 GHz bandwidth and up to 10 GSa/s sampling rate. For the LF path, two Spectrum M3i.4121 acquisition boards are used, which allow for a sampling rate of 250 MSa/s with 14-bit resolution. The specifications for both the LF and RF receivers are summarized in Table 2.1. The signal-to-noise ratio (SNR) can be improved for both receivers by waveform averaging using periodic signals, although spurious signals and distortion may limit the signal-to-noise and distortion ratio (SINAD). The frequency resolution, or frequency bin size, of the measured spectrum is determined by the sampling frequency and number of samples in the measurement receivers as

$$\Delta f = \frac{f_s}{N_s},$$  \hspace{1cm} (2.4)
Figure 2.11: The wideband nonlinear measurement setup. The RF signals are sensed with couplers and digitized with a four-port oscilloscope and the LF signals are sensed with low-frequency sense boards and digitized with PCIe-based acquisition boards. (a) Schematic. (b) Photograph.

Table 2.1: Receiver specifications of the wideband nonlinear measurement setup, taken from datasheets [53–55]. SNR and SINAD are calculated from the number of bits and the effective number of bits, respectively [56].

<table>
<thead>
<tr>
<th>Path</th>
<th>Receiver</th>
<th>Bandwidth</th>
<th>Samplerate</th>
<th>SNR</th>
<th>SINAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>LF</td>
<td>Spectrum M3i.4121</td>
<td>140 MHz</td>
<td>250 MSa/s</td>
<td>86 dB</td>
<td>67 dB</td>
</tr>
<tr>
<td>RF</td>
<td>R&amp;S RTO1044</td>
<td>4 GHz</td>
<td>10 GSa/s</td>
<td>50 dB</td>
<td>44 dB</td>
</tr>
</tbody>
</table>

where $\Delta f$ is the frequency resolution, $f_s$ is the sampling frequency, and $N_s$ is the number of acquired samples. The frequency bin size of the measurement receivers and the signal generator are preferably aligned within integer multiples of each other to reduce FFT leakage.

In order for the incident and reflected signals at the DUT reference planes to be digitized by the measurement receivers, they must be sensed with limited interference of the DUT performance. For the RF path, this is done using commercially available couplers. For the LF path, however, the options are somewhat limited. In [57,58] couplers are used, which are in practice viable down to a few kHz. In [59,60] voltage and current probes are used, which could potentially suffer from limited dynamic range. In this thesis, low-frequency sense boards which actively sense the voltages and currents have been designed and manufactured, as described in paper [D].

The LF sense boards are designed using operational amplifiers (op-amps). The voltage sensing is based on a single op-amp in non-inverting configuration
2.3. WIDEBAND NONLINEAR MEASUREMENT SETUP

The stimuli signals are generated by a Keysight M8190A arbitrary waveform generator (AWG). It has a maximum sampling frequency of 8 GSa/s with 14 bit digital-to-analog (DAC) resolution, 5 GHz analog bandwidth and a memory depth of 2 GSa per channel. The large memory depth and high sampling frequency enables direct RF synthesis of wideband modulated signals without the need of an upconverting stage. The AWG is synchronized with the rest of the system using a 10 MHz reference clock and a trigger. The AWG also has built-in amplifiers to increase the signal power level [54]. External driver amplifiers are however needed for measurements with high power requirements, such as transistor load-pull.

Figure 2.12: Schematic of the low-frequency sense board, with a single operational amplifier for voltage sensing and an instrumentation amplifier for current sensing.

and the current sensing is based on three op-amps in an instrumentation amplifier configuration, as shown in Fig. 2.12. In addition to the op-amps, there are a number of resistors which set parameters such as: input impedance, gain, bandwidth, output impedance and common-mode rejection ratio. For the voltage sensing there is an input resistor \((R_{in})\), pull-down resistor \((R_{pd})\), feedback resistor \((R_f)\), gain resistor \((R_{g1})\) and an output resistor \((R_{out})\). The current sensing is based on differential voltage amplification across a sense resistor \((R_{sense})\). The current sense resistor should be small enough in order to minimize insertion loss and self-heating, with the latter being particularly important as it can introduce calibration errors. The first two op-amps in the instrumentation amplifier are high gain-bandwidth op-amps, of which the outputs are buffered using a differential op-amp. The current sensing branch employs some additional resistors compared to the voltage sensing branch, where \(R_{g2}\), \(R_1\), \(R_2\) and \(R_3\) set the gain, bandwidth and common-mode rejection ratio.

2.3.2 Signal Generation

The stimuli signals are generated by a Keysight M8190A arbitrary waveform generator (AWG). It has a maximum sampling frequency of 8 GSa/s with 14 bit digital-to-analog (DAC) resolution, 5 GHz analog bandwidth and a memory depth of 2 GSa per channel. The large memory depth and high sampling frequency enables direct RF synthesis of wideband modulated signals without the need of an upconverting stage. The AWG is synchronized with the rest of the system using a 10 MHz reference clock and a trigger. The AWG also has built-in amplifiers to increase the signal power level [54]. External driver amplifiers are however needed for measurements with high power requirements, such as transistor load-pull.
2.3.3 Calibration Procedure

In order to measure the actual performance at the reference planes of the DUT in an accurate and reliable manner, a calibration of the system is needed. Even though the measurement receivers might have a factory calibration, systematic errors arising from for example external cabling and signal sensing must be taken care of. Calibration techniques for conventional VNAs have been continuously developed since the 1950s [61], and by adding some extra calibration steps, these techniques can be adapted to nonlinear measurement setups as well. Although each path of the wideband nonlinear measurement setup, LF and RF, require their own separate calibration, they utilize similar procedures. This section is therefore generalized so it can be applied to both paths. The voltages and currents acquired from the LF sense boards are converted to travelling voltage waves in order to follow the same convention as is done in calibration with conventional RF couplers.

The calibration can be divided into three parts. First, a relative calibration is performed at the DUT ports to remove systematic errors. It does not involve any absolute quantification of the travelling waves, hence only ratioed, in other words relative, quantities of the waves can be measured. For a conventional VNA, this is the only step required to be able to accurately measure the S-parameters of an electrical network. The second step in the calibration procedure is to perform an absolute measurement of the travelling waves in reference to a power meter. The third step and last step involves the phase calibration where typically a harmonic phase reference in the form of a comb generator is used. In the wideband nonlinear measurement setup, however, one of the measurement receivers of the oscilloscope is used as a reference receiver whilst a wideband stimuli is applied. As previously discussed, the power and phase calibrations are necessary in nonlinear characterization in order to properly reconstruct the time-domain equivalent of the waveforms at the DUT reference planes. The calibration techniques discussed here originate from the procedures discussed in [62–64].

The calibration of the measurement setup works under the assumption that the measurement receivers operate in a linear region, both during calibration and during the DUT measurements. The user must ensure that the signal levels are sufficiently attenuated before being acquired by the measurement receivers. For the RF receivers, oscilloscopes often have variable gain amplifiers (VGAs) in the front-end for gain adjustment. These can be used to make sure that the incident signals are digitized within a linear region of the ADCs. As discussed in paper [B] however, VGAs can introduce magnitude offsets and phase shifts when the VGA gain is varied. A correction technique has been developed in paper [B], and is later discussed in this chapter. Other assumptions that are made during calibration is that ambient temperature variations are negligible, and that the isolation between the measurement channels is sufficient.

**Relative calibration**

Various algorithms for the relative calibration exist: short-open-load-through (SOLT), short-open-load-reciprocal (SOLR), through-reflect-line (TRL), line-reflect-match (LRM), etcetera [32, 61]. While some algorithms yield more accurate calibrations than others, the SOLR algorithm is presented here due to
its mathematical clarity. The systematic errors of the measurement system in each frequency path, LF and RF, can be modeled as two error boxes, one for each measurement port as seen in Fig. 2.13. This is also known as an 8-term error model \[31,65\]. The acquired signals of the measurement receivers are here denoted as \(m_1, m_2, m_3,\) and \(m_4\). The calibration procedure is carried out one frequency at a time, and combined to obtain the complete frequency dependence of the error boxes. The error model can mathematically be described as follows

\[
\begin{bmatrix}
\tilde{a}_1 \\
\tilde{b}_1 \\
\tilde{a}_2 \\
\tilde{b}_2
\end{bmatrix} =
\begin{bmatrix}
\frac{1}{e_{10}e_{01}} \begin{pmatrix}
-\Delta x & e_{11} \\
e_{00} & 1
\end{pmatrix} & 0 \\
0 & \frac{1}{e_{10}e_{32}} \begin{pmatrix}
-\Delta y & e_{22} \\
e_{33} & 1
\end{pmatrix}
\end{bmatrix}
\begin{bmatrix}
m_1 \\
m_2 \\
m_3 \\
m_4
\end{bmatrix},
\]

where \(\Delta x = e_{00}e_{11} - e_{10}e_{01}\) and \(\Delta y = e_{22}e_{33} - e_{32}e_{23}\). The tilde-sign above each traveling wave quantity, \(\tilde{a}_1, \tilde{b}_1, \tilde{a}_2, \tilde{b}_2\), indicates that absolute calibration is not yet applied. All of the error terms except \(e_{10}e_{32}\) are extracted from two separate short-open-load (SOL) calibrations at each DUT port, as shown in Fig. 2.14a. The \(e_{10}e_{32}\)-term is determined by an unknown-through measurement, as shown in Fig. 2.14b.

The SOL calibration at each port can be set up in the following way to extract the error terms, exemplified here at port 1 as

\[
\begin{bmatrix}
\Gamma_{m,S} \\
\Gamma_{m,O} \\
\Gamma_{m,L}
\end{bmatrix} =
\begin{bmatrix}
1 & \Gamma_{L,S}\Gamma_{m,S} & -\Gamma_{L,S} \\
1 & \Gamma_{L,O}\Gamma_{m,O} & -\Gamma_{L,O} \\
1 & \Gamma_{L,L}\Gamma_{m,L} & -\Gamma_{L,L}
\end{bmatrix}
\begin{bmatrix}
e_{00} \\
e_{11} \\
\Delta x
\end{bmatrix},
\]

where \(\Gamma_{L,S}, \Gamma_{L,O}, \Gamma_{L,L}\) are the reflection coefficients of the short, open and load terminations, respectively, and \(\Gamma_{m,S}, \Gamma_{m,O}, \Gamma_{m,L}\) the measured reflection coefficients of the corresponding terminations. The reflection coefficients of the terminations are ideally \(\Gamma_{L,S} = -1, \Gamma_{L,O} = 1,\) and \(\Gamma_{L,L} = 0\). They are, however, often described with equivalent circuit models or tabulated for increased calibration accuracy.

The \(e_{10}e_{32}\)-term can be determined from a measurement using a reciprocal transmission line, also known as an unknown-through calibration. By describing the signal flow graph in Fig. 2.13 with scattering transfer parameters (T-parameters [66]), the following equation can be set up

\[
T_m = \frac{1}{e_{10}e_{32}} T_1 T_{\text{DUT}} T_2
\]
where $T_m$ is the measured T-parameter matrix, $T_1$ is the T-parameter matrix for error box 1, $T_2$ is the T-parameter matrix for error box 2 and $T_{DUT}$ is the T-parameter matrix for the through transmission line. Since the through standard is assumed to be reciprocal, the determinant of $T_{DUT}$ is equal to unity. As such, taking the determinant of the left hand side and right hand side of (2.7) leads to

$$\det(T_m) = \det\left(\frac{1}{e_{10}e_{32}}T_1 T_{DUT} T_2\right) = \frac{1}{(e_{10}e_{32})^2} \det(T_1) \det(T_2), \quad (2.8)$$

and solving for $e_{10}e_{32}$ finally results in

$$e_{10}e_{32} = \pm \sqrt{\frac{\det(T_1) \det(T_2)}{\det(T_m)}}. \quad (2.9)$$

The sign ambiguity requires only rough knowledge of the unknown through $S_{21}$-phase, and the selection procedure is described in detail in [62].
Absolute calibration

The next stage in the calibration procedure is to perform the absolute power and phase calibration, which determines the $\alpha$-term in the equation below

$$\begin{bmatrix} a_1 \\ b_1 \\ a_2 \\ b_2 \end{bmatrix} = \alpha \begin{bmatrix} \tilde{a}_1 \\ \tilde{b}_1 \\ \tilde{a}_2 \\ \tilde{b}_2 \end{bmatrix},$$  \hspace{1cm} (2.10)

where the tilde-denoted variables are the wave quantities with relative calibration applied from (2.5). The magnitude of $\alpha$ is determined by comparing the measured input power by the system at the DUT reference planes to a reference power meter, as shown in Fig. 2.14c. Mathematically this can be described as follows

$$P_{\text{ref}} = \frac{|a_1|}{2Z_0} - \frac{|b_1|^2}{2Z_0} \Rightarrow |\alpha| = \sqrt{2Z_0P_{\text{ref}}} \frac{|a_1|^2 - |b_1|^2}{|\tilde{a}_1|^2 - |\tilde{b}_1|^2},$$  \hspace{1cm} (2.11)

where $P_{\text{ref}}$ is the power measured by the power meter and $Z_0$ is the system impedance. Because of the frequency limitations of most RF power meters, which are typically valid down to 10 MHz, a power meter cannot be used as a reference for the absolute magnitude calibration of the LF path. Instead, one of the ports of the oscilloscope is used as a measurement reference. Since the frequencies of the LF signals are fairly low in comparison to the analog bandwidth of the oscilloscope, the measured power at the ports of the oscilloscope is assumed to be reasonably accurate.

The phase calibration is lastly made by connecting the port of the first reference plane to one of the measurement receivers of the oscilloscope, as shown in Fig. 2.14d. This is based on the assumption that all measurement receivers of the oscilloscope are fully synchronized. The phase of $\alpha$ can therefore be calculated as follows

$$\alpha = \arg\left(\frac{m_3}{e_{10e_{01}}(m_2(1 + e_{11}) - m_1(e_{00} + \Delta_x))}\right),$$  \hspace{1cm} (2.12)

where $e_{10e_{01}}, e_{11}, e_{00},$ and $\Delta_x$ are the error coefficients in (2.5). The system is now calibrated and ready to perform measurements. However, some additional correction algorithms are needed to correct for discrepancies introduced by the oscilloscope, as discussed in paper [B].

2.3.4 Oscilloscope Correction Algorithms

To extend the usable dynamic range of a nonlinear measurement setup, pre-calibrated step attenuators are often used for leveling of the coupled signals, placed before the measurement receivers [37,42]. The performance of the ADC can therefore always be maximized even though measurements are carried out across a large range of power levels. In the wideband nonlinear measurement setup presented in this thesis, no step attenuators are used. For the LF measurement receivers, the dynamic range is typically enough to avoid adjusting the attenuation of the sensed signals whilst measuring. For the RF path
CHAPTER 2. NONLINEAR RF CHARACTERIZATION TECHNIQUES

Figure 2.15: Typical block diagram of the front-end of a real-time oscilloscope. The input signal to the ADCs can be adjusted in two steps; course adjustments with an attenuator, and fine adjustments with VGAs. The low-pass filter has a cut-off frequency slightly above the analog bandwidth of the oscilloscope to prevent aliasing.

![Block Diagram](image)

Figure 2.16: Measurement setup for extracting the correction factors. (a) Magnitude discrepancy measurement versus a power meter. (b) Phase discrepancy measurement with one port used as reference and one port as a measurement port.

however, the internal VGAs of the oscilloscope are used for signal leveling. A generalized schematic of a typical front-end of an oscilloscope is shown in Fig. 2.15 [67]. A variable attenuator handles the course adjustment of the signal levels, followed by a VGA for finer leveling adjustments. Since adjustment of the input attenuator could have a significant impact on the port match of the oscilloscope, thus invalidating the calibration of the measurement setup, the correction algorithms presented in paper [B] are aimed to handle errors only due to adjustment of the VGA.

In paper [B], it is shown that range adjustment of the VGA introduces a complex gain to the measured signal. Depending on the frequency of the input signal and the voltage range adjustment of the VGA, the magnitude offset and phase shift can be up to 0.6 dB and 15.7 °, respectively, for the particular oscilloscope used in the measurement setup. A correction method based on look-up tables is presented in paper [B]. The correction factors in the look-up tables are extracted from reference measurements using a power meter and differential measurements between oscilloscope ports. The extraction process is simple since it only requires a CW signal generator, a power splitter, a power meter and the oscilloscope itself.

The correction factors for the look-up table are extracted as follows. First, the magnitude offset error is determined by monitoring the discrepancy in measured power of the oscilloscope and measured power of a calibrated power meter. As shown in Fig. 2.16a, the output of a CW signal generator is connected to a two-way power splitter, of which the outputs are connected to a power meter and one of the oscilloscope ports. The extraction procedure for the
Define frequency grid and voltage ranges
Set input power and voltage range to lowest setting
Measure oscilloscope discrepancy vs. power meter

Signal clipping in scope?
Yes
Increase input power
No
Power sweep done?
Yes
Step input signal frequency
No
Freq. limit reached?
Yes
Finished

Increase voltage range

(a)

Define frequency grid and voltage ranges
Set input power and voltage range to lowest setting
Measure & calculate relative phase shift between channels

Toggle voltage range of non-reference channel

Measure & calculate relative phase shift between channels

Calculate phase discrepancy between (1) and (2)

All voltage ranges measured?

Yes

(b)

Figure 2.17: Flow diagram describing the measurement procedure for extracting the correction factors. (a) Magnitude measurement. (b) Phase measurement.

Magnitude is summarized in a flow diagram in Fig. 2.17a. For the phase extraction of the correction factors, two oscilloscope ports are referenced to each other as shown in Fig. 2.16b, and the extraction procedure is summarized in a flow diagram in Fig. 2.17b.

Once the magnitude and phase of the correction factors for each voltage range and frequency are obtained, the raw measurements from the oscilloscope are corrected as

\[
\begin{bmatrix}
\hat{m}_1 \\
\hat{m}_2 \\
\hat{m}_3 \\
\hat{m}_4 \\
\end{bmatrix} =
\begin{bmatrix}
k_{1,x} & 0 & 0 & 0 \\
0 & k_{2,y} & 0 & 0 \\
0 & 0 & k_{3,z} & 0 \\
0 & 0 & 0 & k_{4,w} \\
\end{bmatrix}
\begin{bmatrix}
m_{1,x} \\
m_{2,y} \\
m_{3,z} \\
m_{4,w} \\
\end{bmatrix},
\]

where \(k_{1,x}, k_{2,y}, k_{3,z}, k_{4,w}\) are the complex correction factors for each channel at voltage ranges \(x, y, z, w\). The corrected oscilloscope measurements are denoted as \(\hat{m}_1, \hat{m}_2, \hat{m}_3, \hat{m}_4\).

The validity of the obtained correction factors was verified by measuring the S-parameters of a reflection amplifier. The results were benchmarked against a calibrated VNA and are shown in Fig. 2.18. Without the correction enabled, the measured phase of \(S_{11}\) shows a large discrepancy versus the VNA, whereas the measured magnitude is fairly acceptable. With correction enabled, the
mean discrepancy of both the magnitude and phase tend towards zero, with largest improvements seen in the phase. The standard deviation of both metrics are improved as well.

### 2.3.5 Measurement Setup Verification

To ensure accurate measurements with the wideband nonlinear measurement setup, verification of the LF and RF calibrations must be carried out either on known references or standards, or benchmarked against a trusted measurement instrument.

The LF verification was carried out with two sets of measurements. In the first measurement, a periodic signal with a pulse period, $T_P$, of 10 ms containing two voltage pulses were injected into a 50 Ω load. The voltage and current measured by the LF sense boards were referenced against a voltage and current probe. As seen in Fig. 2.19, the measurements obtained with the LF sense boards agree very well with the voltage and current probes. The second measurement was made using slightly faster pulses, where the periodicity of the two voltage pulses were set to 100 μs. Since the voltage and current probes were not calibrated to provide accurate readings of high-speed signals, one of the ports of the RTO1044 oscilloscope was used as reference instead. The load was in this case set by the port impedance of the oscilloscope, which nominally
2.3. WIDEBAND NONLINEAR MEASUREMENT SETUP

Figure 2.19: LF verification of two pulses, $T_P = 10$ ms, across a 50 Ω load. (a) Measured voltage, voltage probe as reference. (b) Measured current, current probe as reference.

Figure 2.20: LF verification of a two pulses, $T_P = 100$ µs, across a 50 Ω load. (a) Measured voltage, oscilloscope port as reference. (b) Zoomed in view on the second pulse.

is 50 Ω. As seen in Fig. 2.20, the measurement results agree also in this case.

For verification of the RF calibration, the wideband nonlinear measurement setup was benchmarked against a calibrated Keysight N5247A PNA-X. Two verification measurements were carried out, one S-parameter verification measurement to ensure proper relative calibration, and one absolute verification of the power and phase calibrations. Measured S-parameters of a 10 dB attenuator and a 25 Ω mismatch airline are shown in Fig. 2.21. The wideband nonlinear measurement setup performs fairly well over the entire measurement bandwidth, with slightly noisier traces and some minor discrepancies which are most prominent in the 10 dB attenuator measurement. The measured behavior of the mismatch airline is on par with the PNA-X, with slightly more noise.

For the verification of the absolute calibration, the response from a comb generator with 10 dB attenuation was measured. The PNA-X and the wideband nonlinear measurement setup were synchronized with a 10 MHz reference clock, and the comb generator was fed with a squared version of the reference clock. The PNA-X was calibrated on a 10 MHz frequency grid with 300 harmonics using the NVNA software. Since no triggering of the comb generator was done, the measured phases of the tones had a linear progression versus frequency due to a random time delay in both measurement setups. In order to compare the results, a linear de-trending in the measured phase of both measurement setups...
Figure 2.21: S-parameter RF verification of the wideband nonlinear measurement setup (WNMS) benchmarked against a Keysight PNA-X. (a) Measurements of a 10 dB attenuator. (b) Measurements of a 25 Ω mismatch airline.

Figure 2.22: RF verification of the absolute power and phase of the wideband nonlinear measurement setup (WNMS) benchmarked against a Keysight PNA-X. A comb generator fed with a pulse-shaped 10 MHz clock used as DUT. (a) Measured absolute power from the comb generator. (b) Measured relative phases between the comb generator tones with a linear phase de-trend.

was applied. The measured power and relative phases of the comb generator are shown in Fig. 2.22. The measured power shows good agreement with the PNA-X, although spurious signals generated internally in the oscilloscope result in slightly erroneous values at some frequency bins. The relative phases show a slightly larger discrepancy versus the PNA-X, with a ±3 degree variation across the measured bandwidth, if the spurious bins are ignored. Since the output from the comb generator is a very high PAPR pulse, the variation across the bandwidth could be due to a linearity limitation in the receivers of the oscilloscope.

To conclude the verification, the calibration of the measurement setup removes the systematic errors due to e.g. cables and couplers, leading to satisfactory measurement results which are in good agreement with commercial measurement instruments such as the Keysight PNA-X. This is achieved despite the lower dynamic range in the oscilloscope compared to the PNA-X.
2.4 Chapter Summary

In this chapter, the importance of nonlinear characterization has been outlined. Furthermore, it has been shown that wideband operation of circuits and components is important to consider when communication-like signals are used as stimuli. Three different nonlinear measurement setup architectures have been presented, and differences and limitations in each architecture has been discussed. The wideband nonlinear measurement setup proposed in this thesis has been thoroughly discussed and verified. This setup enables the measurements presented in the next chapter, which focuses on nonlinear characterization of RF transmitter chain components using wideband stimuli.
Chapter 3

Applications of Wideband Nonlinear Characterization

Nonlinear measurements enable engineers to study circuits and components under large-signal operating conditions. While the linear S-parameter framework is in many cases sufficient for passive components, some active components require large-signal operation in order to be efficient and to utilize their full potential performance [68]. A simplified RF transmitter block diagram is shown in Fig. 3.1, where a low-frequency information-carrying signal is up-converted to an RF carrier and subsequently amplified. Ideally, these operations should be linear in order to maintain the fidelity of the transmitted signal, and to minimize spectral regrowth which can violate regulatory spectrum masks. The utilization of wider-bandwidth communication signals in modern cellular standards necessitates nonlinear characterization using wideband stimuli. Measurement techniques such as load-pull and investigation of linearity-degrading dynamic effects can give useful insight into how wideband RF PAs should be designed [69, 70]. Furthermore, distortion in the upconverting stage can impair the overall system performance since it precedes the amplification stage [71].

In this chapter, the wideband nonlinear measurement setup presented in Chapter 2 has been used for: 1) transistor performance optimization for concurrent multi-band operation; 2) investigation of dynamic effects in the bias circuitry and its impact on transistor linearity; 3) analyzing the large-signal operation of IQ-mixers using wideband stimuli.

Figure 3.1: Simplified schematic of an RF transmitter. System performance depends highly on the upconverting and power amplification stages.
3.1 Multi-band Load-pull for PA Design

High power signal amplification is necessary to provide sufficient output power in order for the transmitter to meet requirements of a specified link budget. As a consequence, depending on the link distance, the power amplifier tends, in many cases, to be the main power-consuming component of a transmitter [72]. Power amplifier efficiency is, therefore, necessary to consider from a cost, environmental and thermal management perspective, and is often in trade-off with its other properties such as linearity [73]. Furthermore, limited availability of non-allocated frequencies for communication signals has driven the need for fragmented spectrum allocation techniques such as carrier aggregation [3]. For optimal circuit performance, the design procedure must include analysis of concurrent operation with multiple band-limited signals, i.e. concurrent multi-band operation. In this section results from paper [A] are presented, where the wideband nonlinear measurement setup is used for multi-band load-pull characterization of a GaN HEMT.

In contrast to small-signal amplifier design, where a conjugate match results in maximum gain, large-signal performance optimization of amplifiers requires other kinds of performance metrics which depend on multiple parameters of the transistor. Instead of practically realizing different circuit designs for a certain transistor, the networks surrounding the transistor can be synthesized using tunable source and load impedances. The systematic variation of these impedances is known as source-pull and load-pull, and can be achieved using a passive impedance tuner [74], by active signal injection [75], or using a hybrid solution [76]. Passive tuners have a frequency-dependent reflection coefficient which limits the applicability for wideband signals, whereas active injection can guarantee proper termination across wide signal bandwidths [77].

For power amplifier design, the tuning of the load impedance is of most importance, and the working principle of active load-pull can be seen from the following equation

\[
\Gamma_L(f_0) = \frac{a_2(f_0)}{b_2(f_0)} = \frac{Ae^{(2\pi f_0 + \phi)}}{b_2(f_0)}, \tag{3.1}
\]

where \(\Gamma_L(f_0)\) is the load impedance seen by the device at the frequency \(f_0\) and \(a_2(f_0), b_2(f_0)\) the propagating waves at the output port of the device. By monitoring the outwards propagating wave generated by the device, \(b_2(f_0)\), and controlling the amplitude, \(A\), and phase, \(\phi\), of the injected signal on port 2, \(a_2(f_0)\), the load reflection coefficient can be tuned as desired. The injection of \(a_2\) can either be based on a closed-loop or open-loop configuration. With closed-loop injection, the output wave from the device \(b_2\) is coupled, adjusted in amplitude and phase, and then re-injected to the device. Load impedances can in this way be synthesized very quickly, although filtering requirements could potentially limit the bandwidth [78]. Open-loop injection synthesizes the \(a_2\)-wave using a second signal generator on the output, and does therefore not need any filtering. On the other hand, open-loop injection requires an iterative optimization to reach the intended load impedance [77]. An open-loop signal injection was used in paper [A].

In conventional load-pull, a single fundamental tone is provided as stimulus for the device, and the load impedances at the fundamental tone and harmonics
3.1. MULTI-BAND LOAD-PULL FOR PA DESIGN

![Diagram of DUT and Stimuli](image)

**Figure 3.2:** Illustration of the multi-band active load-pull concept. Two high-power tones are injected to the DUT, which generate third order intermodulation products on the output. The load impedances at each frequency is controlled by open-loop active injection on the output. An iterative optimization is needed for obtaining the desired load impedance.

are controlled. By engineering the waveforms, different amplifier classes can be realized and very high efficiencies can be obtained [14]. As discussed in [79] however, controlling the load impedances at intermodulation frequencies can aid in understanding phenomena such as spectral regrowth asymmetry. In paper [A], load-pull at the fundamental and third-order intermodulation (IM3) frequencies was carried out using simultaneous injection of two tones ($f_1 = 2.14$ GHz and $f_2 = 2.65$ GHz), depicted in Fig. 3.2. Although this is a very simplified case of multi-band operation, the wideband capability of the nonlinear measurement setup is demonstrated.

Measurements were carried out on a commercial 6-W GaN HEMT, with a quiescent bias corresponding to Class-AB operation. Due to bandwidth limitations in the measurement setup and in the pre-driver amplifiers, the load impedances at harmonics and second order intermodulation frequencies were not controlled, and were instead passively kept close to 50 Ω. A comparison between single-tone load-pull at each fundamental frequency and simultaneous two-tone load-pull was carried out to investigate whether concurrent operation changes the loading requirements in the transistor. The load impedances were tuned for maximum total efficiency of the transistor, which is defined as follows for the concurrent case

$$
\eta_{\text{total}} = \frac{P_{\text{out},2.14} + P_{\text{out},2.65}}{P_{\text{DC}} + P_{\text{in},2.14} + P_{\text{in},2.65}}, \quad (3.2)
$$

where $P_{\text{out},2.14}$ and $P_{\text{out},2.65}$ are the RF output powers at the first and second fundamental frequencies (2.14 GHz and 2.65 GHz), respectively, $P_{\text{DC}}$ is the DC power consumption, and $P_{\text{in},2.14}$ and $P_{\text{in},2.65}$ are the RF input powers at the first and second fundamental frequencies, respectively. The average input power was set to roughly 31 dBm for both the non-concurrent and concurrent load-pull measurements. As seen in Fig. 3.3a, the optimal load impedances are slightly rotated in the Smith chart between non-concurrent and concurrent operation. Hence, designing an amplifier using single-tone load-pull at each frequency would yield slightly lower efficiency when operating the amplifier concurrently with both frequency bands, compared to a multi-band load-pull design case.

To optimize the transistor efficiency even further, the load impedances at the IM3 frequencies (1.63 GHz and 3.16 GHz) were tuned. The load impedances
at the fundamental tones were set to the previously obtained optimal values for concurrent operation. The IM3 load impedances were tuned purely reactive, i.e. along the periphery of the Smith chart, to minimize power dissipation. The efficiency contours are shown in Fig. 3.3b, where the x-axis and y-axis corresponds to the phase of the load reflection coefficient at 1.63 GHz and 3.16 GHz, respectively. The total efficiency varies from 41.5% to 46% with the optimal impedances found at $\Gamma_{L, 1.63} = 1\angle 80^\circ$ and $\Gamma_{L, 3.16} = 1\angle 140^\circ$. The efficiency enhancement due to optimally terminating the IM3 products is not significant, and a circuit implementation using this type of termination might exhibit even lower efficiency enhancement due to extra losses associated with the IM3 termination network. Nevertheless, these measurements demonstrate the feasibility of using the wideband nonlinear measurement setup for multi-band load-pull applications. PA design using measured load-pull data can, with this measurement setup, be carried out using GHz-bandwidth signals, opening up for studies on e.g. trade-offs between linearity, efficiency, and bandwidth.

### 3.2 The Effect of LF Impedance Termination on Transistor Linearity

As shown in the previous section, when designing power amplifiers for high efficiency and/or high linearity, it is important to consider harmonic terminations as well as terminations at intermodulation frequencies. These intermodulation products will be generated around RF, but will also be generated at frequencies close to DC, causing slow modulation of the voltages and currents in the bias networks. If the bias circuitry is not properly designed or if low-frequency resonances exist, different types of input signals will yield very different output results and could potentially cause severe linearity problems in e.g. base-stations [30]. Furthermore, as shown in [70, 80], low-frequency impedances can be tuned to improve the device linearity. In paper [D], the simultaneous
LF and RF measurement capabilities of the wideband nonlinear measurement setup was used to investigate the influence of the power supply impedance on transistor linearity.

The study was carried out on a CEL NE3512S02 heterojunction field effect transistor. A schematic of the measurement setup is shown in Fig. 3.4. The supply impedance at LF was varied by interchanging the power supply units (PSUs), denoted PSU1 and PSU2. The RF output of the transistor was terminated to 50 Ω. A random-phase multi-tone signal was used as input stimuli, described by

\[
x_{\text{in}}(t) = \sum_{k=1}^{N} A_k \cos(2\pi f_k t + \phi_k),
\]

where \(N\) is the number of tones, \(A_k\) is the amplitude of the \(k^{th}\) tone, \(f_k\) is the frequency of the \(k^{th}\) tone, and \(\phi_k\) is the phase of the \(k^{th}\) tone. Multi-tone signals are advantageous to use since they are periodic by nature, which in turn minimizes FFT leakage issues. Furthermore, the phase of each tone can be randomized to obtain a certain PAPR of the stimuli signal [81].

The phase of the multi-tones were randomized to obtain a PAPR of 10 dB of the input signal, and 10 random phase realizations were measured to obtain an averaged nonlinear response. A measurement using 41 tones with 50 kHz tone spacing, resulting in a total bandwidth of 2 MHz, centered at a carrier frequency of 2.145 GHz is shown in Fig. 3.5. As seen in the RF spectrum, for this particular input signal, there is a significant difference of roughly 9 dB in adjacent channel power ratio (ACPR) between measurements with the two PSUs. Furthermore, when PSU2 is used, the intermodulation distortion around the input signal appears to be slightly asymmetric and occupies a slightly narrower bandwidth compared to when PSU1 is used. Asymmetric distortion can arise from electrical effects in the bias circuitry, and as shown in Fig. 3.5b, the difference in LF impedance between the two PSUs is quite significant. PSU1 presents a short circuit with slightly inductive behavior, whereas PSU2 presents a lossy impedance which in this case translates to notably different RF distortion. The exact cause for the difference in termination impedance is hard to determine without knowing the internal schematics of the PSUs. It could be related to how the voltage and current regulation loops are designed.
3.3 Large-signal IQ-mixer Characterization

Nonlinearities in a microwave transmitter chain arise not only from the high-power amplification stages, but also from upconversion stages such as the mixer. IQ-mixers are commonly used, which upconvert the combination of two orthogonal LF input signals, denoted in-phase (I) and quadrature (Q). The wanted result from this upconversion is a linear frequency-translation from LF to RF, but in practice distortion and other non-idealities will contaminate the RF spectrum. It is therefore of interest to study where these non-idealities may originate from, in order to optimize the performance in a transmitter chain. Previous studies have shown the need of mixer characterization through CW measurements [82], model identification [83], vector mixer measurements [84], and formulation of S-parameters for mixers [85].

In paper [E], the wideband nonlinear measurement setup was reconfigured for four-port device characterization. Two of the ports measure the LF signals, i.e. the I- and Q-inputs, and the remaining two ports measure the RF signals, i.e. the RF and LO ports, as shown in Fig. 3.6. A 10 MHz reference clock and a trigger signal was used for synchronization of the signal generators.
3.3. LARGE-SIGNAL IQ-MIXER CHARACTERIZATION

RF Port 1 RF Port 2
a\textsubscript{1,m}, m b\textsubscript{1,m}, m b\textsubscript{2,m}, m a\textsubscript{2,m}, m
V\textsubscript{1,m}, m I\textsubscript{1,m}, m V\textsubscript{2,m}, m I\textsubscript{2,m}, m
LF Port 1 LF Sense LF Port 2
I Q
LO        ... 69087B
10 MHz Reference clock
Trigger

**Figure 3.6:** Schematic of the nonlinear characterization setup configured for four-port measurements. The I- and Q-signals are generated by the arbitrary waveform generator, and the LO is sourced from a CW signal generator.

As previously discussed, multi-tone signals are often advantageous to use for characterization purposes since they are periodic and can be tailored to provide specific signal properties. As described in [81], multi-tone signals can be further expanded upon by exciting specific frequency bins, allowing for a better analysis and estimation of nonlinear distortion generated in a circuit or device. Under the assumption that the circuit being analyzed belongs to the PISPO-class of systems, odd frequency grids can be utilized to separate odd and even order distortion products. An odd frequency grid means that the multi-tone frequency grid is an odd integer multiplied by a base tone spacing. Odd and even order distortion products can generally be separated when measuring RF circuits since only odd order distortion terms fall in-band. For frequency-translating devices such as mixers on the other hand, this can be a bit problematic since the exact origin of the nonlinearities might not be known; the nonlinearities might be generated at LF and then upconverted to RF, consequently allowing second order distortion products to fall in-band at RF [83].

The excitation signals for the IQ-mixer characterization are based on the odd-order random-phase multi-tone signal concept, but adapted for multi-port excitation, similar to what is discussed in [81]. The I- and Q-signals are synthesized as follows

\[
x_I(t) = \sum_{k=1}^{N} A_{I,k} \cos(2\pi(2k+1)f_{I,k}t + \phi_{I,k})
\]

\[
x_Q(t) = \sum_{k=1}^{N} A_{Q,k} \cos(2\pi(2k+1)f_{Q,k}t + \phi_{Q,k}),
\]

where \(N\) is the number of tones, \(A_{\bullet,k}\) is the amplitude of the \(k^{th}\) tone, \(f_{\bullet,k}\) is the frequency of the \(k^{th}\) tone, and \(\phi_{\bullet,k}\) is the phase of the \(k^{th}\) tone. If the frequency grids for the I- and Q-branches are properly selected, and if the
Figure 3.7: Measured RF spectra from the IQ-mixer. Data points resulting from pure measurement setup noise are excluded for readability. (a) Even order nonlinearities. (b) Odd order nonlinearities. (c) Even and odd nonlinearities, resulting from mixing products between the I and Q frequency grids.

PISPO assumption is kept, distortion products will only coincide on frequencies which share a least common multiple (LCM). In other words, this means that distortion products generated by the I-branch can be separated from distortion products generated by the Q-branch in the upconverted RF spectrum, and vice versa. The selection of the input signal frequency grids for each branch should therefore preferably be made such that the LCM frequency falls out-of-band. The frequency grids must, however, still be compatible with the signal generator and receiver frequency grids in the measurement setup, which are set by the frequency resolution (\(\Delta f = f_s/N_s\)), to avoid FFT leakage.

Measurements were carried out on an Eclipse Microwave IQ2040N IQ-mixer, which has an RF range of 2-4 GHz and an LF range of DC-500 MHz. A carrier frequency of 2.14 GHz was chosen and the RF port of the mixer was passively terminated with a 50-\(\Omega\) load. A signal bandwidth of roughly 10 MHz was excited using 17 tones on each of the I- and Q-branches, and the phases of the tones were randomized to achieve a PAPR of 10 dB. The frequency spacing was chosen to 99 kHz for the I-branch and 101 kHz for the Q-branch, resulting in an LCM frequency of 9.99 MHz. In Fig. 3.7, the upconverted RF spectrum is shown for an average LF drive power of 1.4 dBm. The even order distortion products generated in the I-branch is roughly 10 dB higher than the
Q-branch across the measured bandwidth. In terms of odd order distortion products, however, the distortion products generated in each branch are similar in amplitude, although they are about 10 dB to 20 dB higher than the even order distortion products. Furthermore, the inter-branch distortion products, shown in Fig. 3.7c, are very low in amplitude and close to the measurement noise floor of approximately $-80$ dBm. The measured results indicate that: 1) even order distortion products located in-band in the upconverted RF spectrum are not negligible; 2) for this particular IQ-mixer, there is some branch imbalance in terms of generated even order distortion products; 3) distortion at the very output of the mixer is limited, since the inter-branch distortion products are relatively insignificant.

To investigate the compression characteristics of the IQ-mixer, LO and LF power sweeps were carried out, and the measured results are displayed in Fig. 3.8. At the recommended LO drive power level, 13 dBm, the conversion gain for both branches agree well with data sheet specifications, roughly $-5.5$ dB. As the LF input power is increased, both branches exhibit similar characteristics, with the I-branch compressing at slightly lower input powers. However, the conversion gain of the Q-branch is significantly more sensitive to the LO drive power. Similar trends can be seen in the input impedance, with the I-branch slightly more sensitive to LF input power and the Q-branch more sensitive to the LO drive power. Better input matching is obtained at lower LO drive powers. The average power of the nonlinearities generated in the I-branch is more or less constant versus LO power, whereas the linearity in the Q-branch increases when the LO drive power is increased. For certain LF input powers, the average nonlinear distortion drops roughly 10 dB when the LO drive power is increased from 10 dBm to 13 dBm.

Studies on the large-signal operation of IQ-mixers, enabled by the wideband nonlinear measurement setup, give useful insight into generation of distortion at an early stage in the transmitter. These nonlinearities propagate through the transmitter chain and are eventually amplified by the PA. Characterization of this distortion can facilitate modeling and compensation to enable better overall system performance.

### 3.4 Chapter Summary

As shown in this chapter, variations of load impedances at both in-band and out-of-band intermodulation frequencies allow for fine-tuning of transistor efficiency and linearity, whereas concurrent multi-band operation changes the fundamental frequency loading requirements. For upconversion using IQ-mixers, significant in-band distortion due to even order nonlinearities is present, and I- and Q-branch imbalances result in differences in generated distortion and conversion gain. The I- and Q-branches also exhibit different compression characteristics. The measurements presented in this chapter demonstrate the importance of wideband characterization of commonly-found RF transmitter chain components, which is facilitated by the wideband nonlinear measurement setup.
Figure 3.8: Measured IQ-mixer characteristics versus LO power and average
LF input power. The plotted values are averaged across the signal bandwidth.
(a) Conversion gain for the I-branch. (b) Conversion gain for the Q-branch.
(c) Input impedance for the I-branch. (d) Input impedance for the Q-branch.
(e) Even order nonlinearities, I-branch. (f) Even order nonlinearities, Q-
branch. (g) Odd order nonlinearities, I-branch. (h) Odd order nonlinearities, Q-
branch.
Chapter 4

Characterization of Dispersive Effects in GaN HEMTs

The performance of microwave systems is dependent on optimized circuit designs, but also on the selection of device technology. Gallium nitride is the semiconductor of choice in many RF applications, attributed to its high breakdown voltage and high electron saturation velocity. GaN-based electronics can offer superior output power [86], efficiency [23], and linearity [87] at high frequencies [88]. Despite being a relatively mature technology by now, epitaxial optimization of GaN HEMTs is ongoing research where focus lies on improving the thermal performance and minimizing electrical effects due to trapping. Although dispersion due to electrical and thermal effects is not exclusive to GaN, the high power densities and high density of defects in the material make these effects particularly pronounced in GaN HEMTs [89].

Dispersion causes the alternating current (AC) I-V characteristics to deviate from the DC I-V characteristics and introduces memory effects in the device, which leads to degraded RF performance [90], degraded reliability [91] and additional distortion [29,92]. Memory effects can also arise from the surrounding electrical networks of a transistor (biasing networks, matching networks, etc.), but those can be dealt with by proper circuit design. The operating currents and voltages in a transistor may vary drastically depending on the input signal, quiescent bias and load impedance. As exemplified in Fig. 4.1, the load-line for a class-B biased transistor passes through regions with high electric fields and varying power dissipation. This can charge or discharge traps in the material, and while the temperature might not vary in RF speed, the envelope of modulated signals can cause variations in device temperature and thereby distortion [92].

Proper understanding of electrical and thermal effects requires accurate characterization methods. In devices with both trapping phenomena and excessive heating, such as in GaN HEMTs, it can be hard to separate the different effects due to overlap in time constants [29]. Multiple characterization methods must be utilized to fully grasp the impact of trapping and heating,
40  CHAPTER 4. CHARACTERIZATION OF DISPERSIVE EFFECTS IN GAN HEMTS

Figure 4.1: Output characteristics of a transistor together with an ideal class-B load-line. The output signal will be influenced by both electrical and thermal effects, which depend on the instantaneous currents and voltages in the device.

Figure 4.2: I-V characteristics showing the effect of knee walkout on pulsed I-V and on the RF load line, kink effect on the DC I-V, and current collapse on the drain current after a high voltage pulse.

and to incorporate these effects in transistor models [93]. This chapter aims to give an overview of various electrical characterization methods for quantifying trapping phenomena and thermal effects. First, some sources of dispersion are discussed, and how they are affected by the GaN epitaxial design. The electrothermal characterization methods used in paper [C] and paper [F] are later presented. A new method for extracting thermal transfer functions between mutually coupled devices is derived at the end of the chapter.

4.1 Sources of Dispersion

Dispersion can generally be broken down in two categories depending on the origin; electrical effects and thermal effects. The underlying mechanisms are triggered by different processes, but the effect on the I-V characteristics can sometimes be similar. In this section, various sources of dispersion and approaches to mitigate them are presented. As shown in paper [C] and paper [F], the epitaxial design of GaN HEMTs influences thermal properties and trapping phenomena to a large degree.

4.1.1 Electrical Effects

Dispersion due to electrical effects can manifest itself on the I-V characteristics in the form of a DC kink effect [94], knee-walkout [95], and current collapse [96] among others. These effects are visualized in Fig. 4.2. As a consequence, the RF performance of GaN-circuits may suffer from e.g. a lower output power due to a limited voltage swing (knee walkout), or time-varying effects when
subject to high-power pulses (current collapse) as shown in paper [b]. Electrical effects in a transistor are generally attributed to the formation of trapping centers throughout the epitaxial material, due to impurities, crystal defects and surface states [97,98]. These trapping centers can either capture electrons/holes, leading to a charged state in the vicinity of the channel, or emit electrons/holes. Both of these processes lead to modulation of the current through the channel of the transistor, detrimentally impacting the I-V characteristics. The trapping centers are typically located either at the surface of the device or in the buffer in the GaN material [99].

As discussed in [100], surface states can become charged through tunneling injection from the gate. It has also been shown that surface traps can be activated in the presence of high electrical fields and atmospheric moisture, causing oxidization and a charged layer on top of the device [101]. This can be mitigated by means of surface passivation, and by controlling the ambient air composition [102].

Traps in the buffer arise mainly from defects in the GaN crystal structure and from impurities throughout the buffer. Since free-standing GaN substrates are not readily available, GaN is typically grown on silicon, sapphire or silicon carbide (SiC) substrates. The lattice mismatch between GaN and the substrate can introduce dislocations in the GaN buffer, although lattice matching can be improved by using a nucleation layer [103]. Impurities in the buffer can be introduced either intentionally or unintentionally. Depending on the growth conditions, impurities such as silicon and oxygen can be unintentionally incorporated into the GaN crystal, leading to n-type GaN [104, 105]. The n-type GaN leads to a leaky buffer with poor electron confinement in the channel, which is detrimental to the long-term reliability and high-frequency performance in a GaN HEMT [106,107]. Consequently, to increase the buffer resistivity, intentional impurities such as carbon (C) or iron (Fe) are often used, which form deep acceptors [105].

In paper [C], the buffer is rendered resistive using C as a compensation dopant. The doping profile in a C-buffer can be tailored in a different way compared to the industry-standard Fe-doping, and could potentially lead to a better electron confinement with limited influence of trapping. Three samples with different buffer designs were investigated: two with uniform doping profiles and varying carbon incorporation (High-C and Low-C), and one with a stepped doping profile (Stepped-C).

### 4.1.2 Thermal Effects

In simplified terms, the temperature ($T$) in a component will depend on the ambient temperature ($T_A$), the dissipated power in the component ($P_{diss}$) and the thermal impedance of the component ($Z_{TH}$) as follows

$$T = T_A + Z_{TH}P_{diss}. \quad (4.1)$$

Semiconductor properties such as electron velocity are temperature dependent, and therefore any change in temperature will alter the device performance [108]. The resulting change in temperature due to the power dissipation of a sinusoidal signal will depend on the frequency response of the thermal impedance, of which the cutoff frequency can be 100 MHz or more [109,110]. For CW signals
above the thermal impedance frequency response, the associated Joule heating will lead to a fixed temperature increase. For modulated signals and for signals below the cutoff frequency of the thermal response, however, the temperature will fluctuate in envelope speed, leading to a time-varying performance and distortion [92,111].

In integrated systems, self-heating is accompanied by mutual thermal coupling, which can be on the scale of thermal coupling between gate fingers on a transistor or between different functional blocks in e.g. a transceiver. The frequency response of the thermal transfer function between the thermally coupled devices will depend on the physical separation and the material properties, with cut-off frequencies ranging from kHz to MHz [92].

Thermal management in GaN-based circuits is challenging due to the high associated power densities, up to 40 W/mm [86]. Although GaN HEMTs can operate at high channel temperatures, the output power performance is to large extent restricted by self-heating [27,28]. Since the majority of the generated heat is conducted through the material stack, thermal optimization of the GaN HEMT epitaxial structure is highly relevant. This can include utilization of diamond substrates instead of Si or SiC [112], optimization of thermal boundary resistances [113], and optimization of material thicknesses [114]. As shown in paper [F], where three different materials (M1, M2, M3) were investigated, variations in the nucleation layer and the thickness of the buffer influences the device self-heating and mutual thermal coupling.

To quantify thermal effects and to determine the origin of trapping phenomena, relevant characterization methods are required. Although RF measurements can be used for electrothermal modeling [115] and for investing trapping-related performance degradation [90,116], many thermal and trapping extraction techniques are built on pulsed I-V and low-frequency AC measurements. An overview of common measurement techniques will be presented in the next sections.

4.2 Pulsed I-V

Pulsed I-V measurements typically involves pulsing the gate and drain of the transistor to cover various parts of the I-V characteristics, from different quiescent bias points. The pulse width and pulse period time can be varied to perform e.g. iso-thermal measurements.

4.2.1 Pulsed I-V for Trap Characterization

Dynamic I-V characterization using pulsed I-V can be used for quantitative analysis of how trapping affects the I-V characteristics, and thereby also for coarse estimation of the impact of trapping on large-signal RF performance. The gate and drain terminals are simultaneously pulsed with a narrow, low duty cycle pulse whilst measuring the in-pulse drain current. Different trapping centers can be probed through variations of the quiescent bias points, as exemplified in Fig. 4.3. In general, a negative quiescent gate bias will charge traps in the vicinity of the gate and parts of the buffer underneath the gate. This can lead to shifts in the threshold voltage of the device. In addition to a negative quiescent gate bias, a high quiescent drain bias can charge traps
4.2. PULSED I-V

Figure 4.3: Illustration of how pulsed I-V measurements can be used to probe different trapping effects. Iso-thermal trap characterization, assuming narrow pulse widths, from different quiescent bias points \((V_{GS,Q}, V_{DS,Q})\).

Figure 4.4: Pulsed I-V characteristics (paper [C]), \(T_{PW} = 1\) µs and \(T_P = 1\) ms, where the pulsing is done from different quiescent states, marked within parenthesis as \((V_{GS,Q}, V_{DS,Q})\). (a) High-C. (b) Stepped-C.

deeper down in the buffer, which can lead to higher access resistances due to back-gating effects [117,118].

Even though the effects of trapping on the I-V characteristics can be investigated through pulsed I-V, it provides limited information about the underlying trapping mechanism [117]. Furthermore, depending on the operating conditions, the charge and emission time constants can sometimes be asymmetrical for GaN HEMTs [96], causing pulse-to-pulse variation. Therefore, steady-state values of the drain current may or may not be reached within the measurement window. In addition to this, pulse widths in the order of nanoseconds are in some cases needed to achieve iso-thermal measurements [109], which might not be feasible in practice. A new pulsed I-V characterization technique, double-pulsed I-V, has been proposed to circumvent the asymmetrical charge and emission times in GaN HEMTs. A pre-pulse is utilized to set the states of the traps, followed by another pulse in which the current is monitored. In this way, different trap states and their influence on the measured I-V characteristics can be investigated in a more methodical way [119]. However, since double-pulsed I-V is not implemented in conventional measurement instruments, custom measurement hardware must be set up.

In paper [C], trap characterization using pulsed I-V was carried out with different quiescent bias states; one reference state where \((V_{GS,Q}, V_{DS,Q}) = (0, 0)\) V, one off-state with a reverse gate bias where \((V_{GS,Q}, V_{DS,Q}) = (-6, 0)\) V, and three states with reverse gate bias and different drain biases. The results in Fig. 4.4 indicate that the highly carbon-doped buffer suffers from severe trapping, which is emphasized when the drain bias is increased. Compared to the reference state, the ON-resistance increases nearly by a factor eight and the
saturation current drops with roughly 75% when pulsed from a (−6, 15)-bias. The buffer with a stepped carbon doping exhibits a slight drop when a negative gate voltage is quiescently biased, indicating some gate-related trapping.

4.2.2 Pulsed I-V for Thermal Characterization

The common approach for utilizing pulsed I-V measurements for thermal characterization is to investigate the influence of device self-heating on I-V characteristics using different pulse widths and/or pulse periods, illustrated in Fig. 4.5a. This can be used in comparison with DC I-V characteristics to extract simple models of thermal networks [120]. In the presence of trapping, however, it can be hard to distinguish self-heating from e.g. a reduction in current due to charging of traps, depending on the time constants.

Extraction of thermal resistances in GaN HEMTs can be carried out using a method presented in [121], which relies on the temperature dependence of the ON-resistance, $R_{ON}$. First, a look-up table of the $R_{ON}$ versus temperature is made by sweeping the ambient temperature using e.g. a thermal chuck. $R_{ON}$ is then extracted at a fixed ambient temperature whilst pulsing from different quiescent bias points, corresponding to different dissipated powers ($P_{diss}$), as shown in Fig. 4.5b. An equivalent thermal resistance, $R_{TH}$, can then be extracted by mapping the $R_{ON}$ versus $P_{diss}$ to the look-up table previously created. However, as discussed in [121], the method is highly dependent on sub-microsecond pulsing to limit cool-down of the device during the $R_{ON}$ versus $P_{diss}$ measurement. Furthermore, if the pulsing is made from high electric fields, additional uncertainties due to trapping may be introduced. In addition, as argued in [122], the thermal resistance in GaN HEMTs can be more accurately described using a polynomial model.

The method in [121] was used in paper [F] to extract the $R_{TH}$ for three different materials with varying epitaxial design. The measured $R_{ON}$ versus temperature and $R_{ON}$ versus $P_{diss}$ are shown in Fig. 4.6, and the extracted $R_{TH}$ is summarized in Table 4.1. The thermal resistance is lowest in the M2-material, which is grown by a different supplier compared to M1. M3, which is grown by the same supplier as M2 but has a thinner buffer, exhibits higher $R_{TH}$ for the smaller test structures. For the larger test structures, however, the difference between the materials are less pronounced.
4.3 Transient Methods

In this thesis, transient methods refer to characterization methods which rely on monitoring changes in e.g. current after or during a transient event such as a large voltage pulse. Trap discharging processes following a high power pulse can be studied to e.g. determine the origin of the trap. Lateral heat propagation in semiconductors can be studied by monitoring the change in resistance in a sensor due to adjacent heating elements driven by pulsed sources.

4.3.1 Drain Current Transient Measurements for Trap Characterization

Drain current transient (DCT) measurements focuses on characterizing the de-trapping processes following a trap-filling pulse. Depending on the quiescent bias point, different emission processes can be identified, and different ambient temperatures can be utilized to extract activation energies and capture cross sections of the corresponding traps through an Arrhenius diagram [97]. Furthermore, the filling pulse time can be varied to determine the origin of the trap [123]. DCT measurements can be made iso-thermal if the gate and drain are pulsed simultaneously to a pinched off state, where \( V_{DS} \) can be kept high without heating up the device. While mostly useful for extraction information about the underlying trap mechanism, DCT measurements can in combination with other measurements be used to predict pulsed-RF performance. For example, in combination with bias-dependent S-parameters, the RF gain collapse after a high-power pulse can be estimated, as shown in paper [b].

In paper [C], a DCT measurement was carried out on the carbon-doped GaN HEMTs with a filling time of 10 \( \mu \)s and a period time of 1 s. The gate quiescent bias was set to a constant value corresponding to 25% of the saturation current, whilst the drain side of the transistors was pulsed from \( V_{DS,Q} \) to 2\( V_{DS,Q} \). Since

\[ R_{ON} (\Omega) \]

\[ \text{Temperature (°C)} \]

\[ 0 \quad 25 \quad 50 \quad 75 \quad 100 \]

\[ 15 \quad 20 \quad 25 \quad 30 \]

\[ \text{M1} \quad \text{M2} \quad \text{M3} \]

\[ \text{P diss} \ (W) \]

\[ 15 \quad 16 \quad 17 \quad 18 \quad 19 \quad 20 \]

\[ \text{Material} \quad 2.5 \ \mu \text{m} \quad 4 \ \mu \text{m} \quad 32 \ \mu \text{m} \quad 45.5 \ \mu \text{m} \]

\[ \text{M1} \quad 51.2 \quad 39.6 \quad 21.6 \quad 21.6 \]

\[ \text{M2} \quad 42.9 \quad 35.5 \quad 19.2 \quad 17.2 \]

\[ \text{M3} \quad 60.9 \quad 47.2 \quad 23.3 \quad 17.8 \]

\[ \text{Table 4.1: Extracted thermal resistance (K/W) for TLM structures on three different GaN-on-SiC materials.} \]
only the drain was pulsed, the measurement was unfortunately not iso-thermal. The quiescent drain voltage was normalized to take into account differences in gate-source and gate-drain distances between the devices. The drop in drain current due to the filling pulse was monitored up to 10 ms after the pulse, with the result shown in Fig. 4.7. A polynomial model was fitted to the transient data in order to obtain a smooth derivative of the response [124]. Two emission peaks are visible for High-C, and one each for Low-C and Stepped-C. The emission peak for Stepped-C is very pronounced in amplitude, and could be a result of an unfair comparison between the devices. The electric field distribution in the Stepped-C devices is very different compared to the other devices due to differences in gate-source and gate-drain distances. The DCT measurement nevertheless displays the different time constants involved in each material.

### 4.3.2 Transient Lateral Thermal Coupling

Thermal coupling between closely placed semiconductor devices can be investigated using a transient measurement method presented in paper [j], where two resistors are used in a heater/sensor arrangement. A high-power voltage pulse is applied to the heating resistor, and the associated Joule heating will spread throughout the semiconductor material. The current flowing through the biased sensor will drop due to the increase in its operation temperature. The heating time constants and drop in current will depend on the physical separation between the heater and sensor, the heat propagation properties of the material, and the temperature dependence of the sensor.

This method was used in paper [F] for investigating the thermal coupling between test structures on different GaN-on-SiC materials. The current response of the sensor due to a 2.2 W heating pulse is shown in Fig. 4.8a. The resulting temperature change in the sensor, shown in Fig. 4.8b, can be estimated by using the temperature dependence of the sensor, obtained from pulsed I-V measurements at different ambient temperatures. The resulting temperature increase is lower for M2 (13.3°C) than for M1 (14.2°C), in spite of the larger current drop in M2. This is equivalent to a thermal coupling of 6.0 K/W and 6.5 K/W for M2 and M1, respectively. Trapping phenomena in M3, in combination with the increase in temperature, causes discharge of trapped electrons in the sensor, seen at around 0.2 ms in Fig. 4.8a. This would be
interpreted as a flattening of the temperature response, or even a drop in temperature depending on the sensor bias voltage. M3 is therefore omitted from the temperature estimation in Fig. 4.8b. A new $3\omega$-based method is later proposed for lateral thermal coupling measurements, which does not require any bias on the sensor, potentially minimizing the influence of trapping during measurements.

### 4.4 Small-signal AC measurements

Small-signal measurements are a powerful complement to the commonly used pulsed I-V and transient measurement methods. Dispersion does not exclusively affect the large-signal performance in transistors, but also the small-signal characteristics. This can be utilized to e.g. perform AC-sweeps over a large frequency range by means of low-frequency VNAs, enabling extraction of both fast and slow time constants related to electrothermal effects in a single measurement.

#### 4.4.1 Admittance Dispersion

Small-signal admittance dispersion can be intuitively understood from the illustration in Fig. 4.9. Dispersion can also arise in the transconductance [125], but is exemplified here for the output admittance. Nominally, the output admittance of a GaN HEMT at low frequencies will not produce any reactive currents since the drain-source capacitance, $C_{ds}$, is negligible. At certain frequencies however, time constants related to trapping or thermal processes can cause resonances and/or hysteresis which generate reactive currents, giving rise to peaks in the output susceptance (Fig. 4.9b). Trapping-related dispersion is typically temperature-dependent, and by sweeping the ambient temperature and tracking the frequency of peaks, the activation energy of the corresponding trap can be extracted through an Arrhenius-diagram [125]. For heating-related dispersion, the resonance frequency is static versus ambient temperature, assuming that the thermal properties of the materials do not change over the measured temperature range.

The operating conditions of the device are mainly set by the bias point, which...
can be varied to probe the trapping phenomena in different ways. For example, time constants related to trapping in admittance dispersion measurements have shown to be bias-dependent [126, 127], which can be utilized to investigate the effect of admittance dispersion on specific amplifier classes. For large-signal applications, admittance dispersion measurements can be combined with pulsed-RF measurements to provide better model accuracy [128]. Furthermore, in [125] some correlation between DCT measurements and admittance dispersion measurements is noted.

In [129] it is shown that, by modeling a single trap center in a GaN HEMT as an RC-network, the output admittance can be expressed as follows

\[
Y_{22}(\omega) = \left[ g_{ds} + \frac{\omega^2 C_T}{1 + (\omega/\omega_T)^2} \right] + j \left[ \omega C_{ds} + \frac{\omega C_T}{1 + (\omega/\omega_T)^2} \right], \tag{4.2}
\]

where \( g_{ds} \) is the output conductance, \( \omega \) is the angular frequency, \( C_T \) the modeled dispersion capacitance, \( \omega_T = \frac{2\pi}{\tau_T C_T} \), and \( C_{ds} \) is the output capacitance. The argument of (4.2) will exhibit a peak at the frequency of the emission rate of the trap, under the assumption that \( g_{ds} \) is large at low frequencies [129].

In paper [C], the output admittance of the GaN HEMTs were measured from 5 Hz to 3 GHz using a VNA, and the ambient temperature was swept from 20 °C to 160 °C. The results are shown in Fig. 4.10. The admittance dispersion in High-C is pronounced, with two emission peaks moving versus temperature, labeled \( E_2 \) and \( D_1 \). In Low-C only one emission peak is detected, labeled \( E_2 \). For Stepped-C, the observed dispersion peaks do not follow the typical temperature-dependence of trapping-related dispersion, and therefore is not considered to be an effect related to trapping. Activation energies of the traps causing the emission peaks have been extracted using an Arrhenius diagram, shown in Fig. 4.10d, where \( E_2 = 0.59 \) eV and \( D_1 = 0.15 \) eV. It is argued in [130] that the trap corresponding to the \( E_2 \) activation energy can be due impurities or intrinsic defects. Similarly, the \( D_1 \)-trap is shown in [131] to be linked with defects and dislocations, although in [132] it is correlated with surface-related trapping.
4.5. CHARACTERIZATION OF THERMAL COUPLING USING A MODIFIED $3\omega$-METHOD

4.5.1 Thermal Impedance Extraction from Z-parameters

As discussed in [128], thermal effects in GaN HEMTs can introduce shifts in the real part of the output conductance. Furthermore, it has been shown in [110] that the thermal impedance can be extracted from low-frequency impedance measurements. The frequency dependence of the impedance, assuming no trapping dispersion, is directly related to the time constants of the thermal response. As shown in Fig. 4.11, the real part of the input impedance of a GaN channel can vary dramatically depending on the applied electric field, stimulus frequency, and ambient temperature. For this particular device, the thermal response is visible up to roughly 50 MHz, indicated by the flatness of the impedance between 50 MHz and $\sim 2$ GHz. This implies that pulses of 20 ns or shorter are required for iso-thermal pulsed I-V characterization. The real part of the impedance drops above $\sim 2$ GHz due to device parasitics.

4.5. Characterization of Thermal Coupling using a Modified $3\omega$-method

The $3\omega$-method is well-established for characterization of thermal properties in materials. A heating element is driven with a sinusoidal current at an angular frequency of $\omega$, which will generate Joule heating at $2\omega$. The resulting temperature fluctuation due to the heating will depend on the thermal impedance of the surrounding materials. This thermal fluctuation will modulate the resistance of the heating element at $2\omega$, of which the amplitude of the modulation depends on the temperature sensitivity of the heater resistance. The mixing of the driving current at $\omega$ and the resistance modulation at $2\omega$ will generate
a voltage at $3\omega$. By measuring the in-phase and out-of-phase components of the $3\omega$ voltage, the thermal conductivity of the surrounding material can be calculated [133, 134].

For characterizing mutual thermal coupling between two resistive elements, a heater/sensor arrangement can be used, similar to the transient thermal coupling method described in Section 4.3.2. A heating signal is used to generate thermal fluctuation at the sensor, and a probe signal is used at the sensor in order to pick up on the thermal fluctuation. However, using the standard $3\omega$-method would result in self-modulation of the sensor, due to the injection of a low-frequency signal into the sensor resistor. Other solutions to this problem propose a DC-bias on the sensor side, and injection of a sinusoidal current on the heater side [135]. Since trapping phenomena is highly affected by the electric field [127], a zero DC-bias could potentially minimize the influence of trapping. In paper [F], a zero DC-bias $3\omega$-based method for thermal coupling characterization is introduced. In the proposed method, referred to as the modified $3\omega$-method, the heating element is driven with an LF signal, and the sensing element is probed with an RF signal, which has a frequency above the cutoff frequency of the thermal network. No DC-bias is therefore needed, and there is no resistance modulation in the sensor due to the RF signal, apart from a constant offset due to the average dissipated power in the RF signal.

The proposed method is illustrated in Fig. 4.12, where the heating element is driven with an LF signal at $\omega_{LF}$ and the sensor is driven with a small RF
4.5. CHARACTERIZATION OF THERMAL COUPLING USING A MODIFIED $3\omega$-METHOD

The probe signal at $\omega_{RF}$. The heat wave propagating throughout the material will modulate the resistance of sensor. The amplitude of the resistance modulation will depend on the heating power, temperature sensitivity of the sensor, and material properties. The RF signal will mix with the modulated resistance, generating intermodulation products at $\omega_{RF} - 2\omega_{LF}$ and $\omega_{RF} + 2\omega_{LF}$. A thermal transfer function between the heater and sensor can be extracted by measuring the intermodulation products and relating them to the dissipated power at LF and the temperature sensitivity of the sensor. A derivation of the method will follow.

4.5.1 Theory

The voltages and currents flowing in to the heating element are defined as

\[
v_{LF}(t) = A_{v,LF} \cos(\omega_{LF}t) \quad (4.3)
\]

\[
i_{LF}(t) = A_{i,LF} \cos(\omega_{LF}t), \quad (4.4)
\]

which will generate an apparent power at LF as

\[
S_{LF}(t) = \frac{1}{2} v_{LF}(t)i_{LF}(t) = \frac{1}{4} A_{v,LF} A_{i,LF} + \frac{1}{4} A_{v,LF} A_{i,LF} \cos(2\omega_{LF}t). \quad (4.5)
\]

The first term, at DC, will offset the nominal sensor resistance. The modulation of the sensor resistance due to the heating wave, at $2\omega$, is defined to be related to the power at LF as

\[
r_{\text{mod}}(t) = \alpha_r S_{LF}(t) = \frac{1}{4} \alpha_r A_{v,LF} A_{i,LF} \cos(2\omega_{LF}t), \quad (4.6)
\]

where $\alpha_r$ is a transfer function which relates the modulation of the resistance to the input power at LF. Due to the modulation of the sensor resistance, the incident RF probe current will create voltage intermodulation products around the fundamental RF signal as follows

\[
v_{\text{sens}}(t) = A_{i,RF} \cos(\omega_{RF}t) \cdot r_{\text{mod}}(t) = \frac{1}{8} \alpha_r A_{i,RF} A_{v,LF} A_{i,LF} \cos((\omega_{RF} - 2\omega_{LF})t) + \cos((\omega_{RF} + 2\omega_{LF})t)). \quad (4.7)
\]

As seen from the expression above, the magnitude of the generated intermodulation products is dependent on the RF probe current and LF power. By measuring the voltage at the $\omega_{RF} + 2\omega_{LF}$-component, referred from now on as the $3\omega$-frequency, the following relationship is obtained

\[
A_{v,3\omega} = \frac{1}{8} \alpha_r \gamma A_{i,RF} A_{v,LF} A_{i,LF}, \quad (4.8)
\]

where $\gamma$ is a current scaling factor, which takes into account the nominal (non-modulated) sensor resistance and the source impedance of the signal generator as follows

\[
\gamma = -\frac{A_{v,3\omega}/A_{i,3\omega}}{(A_{v,RF}/A_{i,RF}) - (A_{v,3\omega}/A_{i,3\omega})}. \quad (4.9)
\]

The transfer function $\alpha_r$ can now be calculated as

\[
\alpha_r = 8A_{v,3\omega}/(\gamma A_{i,RF} A_{v,LF} A_{i,LF}). \quad (4.10)
\]
This is an arbitrary transfer function which does not provide any absolute information of the thermal transfer between the heater and sensor. To extract the thermal transfer function between the heater and sensor, the temperature dependence of the sensor resistor needs to be known. This can be extracted by sweeping the ambient temperature and measuring the sensor resistance, as done in the $R_{ON}$ versus temperature measurement in Fig. 4.6a.

### 4.5.2 Method Verification

To verify the method, a harmonic balance simulation was set up in Keysight ADS, as shown in Fig. 4.13. Two two-port symbolically defined devices (SDDs) were defined as heating and sensing elements, with a thermal network connecting the temperature ports of the two SDDs together. The thermal network was modeled as a Π-network, with resistors and capacitors in each branch. The sensor resistance was modeled as a function of the temperature as

$$R_T(T) = R_{T0} + R_{T1}T, \quad (4.11)$$

where $R_{T0}$ is the nominal sensor resistance and $R_{T1}$ is the linear temperature coefficient. The frequency of the RF probe signal was chosen to a value beyond the response of the thermal network whilst the frequency of the LF signal was swept. The impedance of the signal generator was set to an arbitrary value of $25 + j25 \, \Omega$ to show that the current scaling factor, $\gamma$ in (4.9), compensates for the source impedance.

The response of the thermal coupling network is calculated as $V_{ch}/I_{ch}$ from the harmonic balance simulation (Fig. 4.13). It is shown versus the heating frequency in Fig. 4.14, together with the extracted values using the modified $3\omega$-method. The extracted values show excellent agreement with the simulated response of the thermal coupling network.

### 4.5.3 Thermal Coupling in GaN-on-SiC

The modified $3\omega$-method was used in paper [F] for characterization of mutual thermal coupling between two semiconductor resistors on different GaN-on-SiC materials. The wideband nonlinear measurement setup presented in Chapter 2...
4.5. CHARACTERIZATION OF THERMAL COUPLING USING A MODIFIED $3\omega$-METHOD

Figure 4.14: Impedance response of the thermal transfer function versus the heating frequency. The solid black line is the true thermal transfer response and the dotted red line is the extracted thermal transfer function using the modified $3\omega$-method. (a) Real part. (b) Imaginary part.

Figure 4.15: The wideband nonlinear measurement setup configured for measuring thermal coupling using the modified $3\omega$-method.

was used for calibrated measurements of both the LF and RF signals. A schematic of the measurement setup is shown in Fig. 4.15. A coaxial calibration was done at the LF and RF ports, and the on-wafer probes and device parasitics were de-embedded from the measurements in post-processing. The frequency of the LF signal was swept from 100 Hz to 1 MHz at a power level of 22 dBm, and the RF signal was set to a frequency of 1.2 GHz and a power level of -8.4 dBm.

The extracted magnitude and phase of the transfer function, $\alpha_r$, is shown in Fig. 4.16a and Fig. 4.16b, respectively. Some spread in the magnitude at lower frequencies is visible, which can be explained by different thermal coupling in the materials or by different temperature sensitivities in the sensors. The overall frequency response is similar, however, and no trapping-related frequency dispersion is observed in M3. This is in contrast to the previous thermal coupling measurements, which were highly influenced by trap discharging (Fig. 4.8).

To obtain the thermal transfer function between the heater and sensor, the different temperature sensitivities in the sensors can be factored out by using the temperature-dependence of $R_{ON}$ (Fig. 4.6a). The real and imaginary parts of the thermal transfer function are shown in Fig. 4.16c and Fig. 4.16d, respectively. The thermal coupling is slightly lower in M2 compared to M1, which is in line with the transient lateral thermal coupling measurements (Fig. 4.8). The real values of the thermal coupling, roughly 3.2 K/W for M1 and 3.0 K/W for M2 at 100 Hz, is also in line with the steady state values of roughly 6.5 K/W (M1) and 6.0 K/W (M2) obtained with the transient
Figure 4.16: Mutual thermal coupling between two semiconductor resistors on three different GaN-on-SiC materials (paper [F]). The frequency of the heating signal was swept from 100 Hz to 1 MHz. (a) Magnitude of the transfer function, $\alpha_r$. (b) Phase of the transfer function, $\angle \alpha_r$. (c) Real part of the thermal transfer function, $H_{TH}$. (d) Imaginary part of the thermal transfer function, $\mathfrak{I}(H_{TH})$.

measurement. The thermal coupling in M3 is lower than for the other materials, which could indicate that the thinner buffer in M3 causes more of the heat to diffuse through the substrate rather than laterally. No thermal coupling is seen above 80 kHz, which agrees fairly well to the 7 $\mu$s delay seen in the transient measurements (Fig. 4.8).

4.6 Chapter Summary

This chapter has given an overview of some common techniques for characterization of electrical and thermal effects. Quantification of these effects are challenging from a measurement perspective, as it puts high requirements on the measurement hardware, e.g. high-power nanosecond pulsing. Well-designed experiments are needed for separation of electrical and thermal effects. Characterization of trapping phenomena is challenging due to how it manifests itself on the DC I-V, pulsed I-V, RF, LF, small-signal and large-signal characteristics. There is no single figure-of-merit describing the general impact of trapping on transistor performance, and the type of measurement should be decided by the application need, in order to provide relevant information to the end user. With regards to thermal parameters, definitions are more established, but under the influence of trapping it can be complicated to distinguish the different effects from each other. Therefore, new methods which potentially minimize the influence of trapping during thermal characterization, such as the proposed modified $3\omega$-method, are of interest.
Chapter 5

Conclusions and Future Work

5.1 Conclusions

Designing cutting-edge microwave hardware for next generation wireless systems requires a multi-faceted approach, of which accurate measurements and device characterization play a crucial role. The utilization of wideband, high-PAPR communication signals in combination with strict energy efficiency requirements put high requirements on the microwave hardware. Furthermore, the utilization of III-V semiconductors such as GaN introduces additional challenges in the form of trapping phenomena. This, in combination with high power densities, degrades the achievable RF performance, and introduces dynamic effects which are highly dependent on e.g. the amplifier class, input signal, and voltage and current swing. This thesis aims to provide new tools for wideband, nonlinear characterization and to provide new insight into dispersive effects in GaN HEMTs.

A core part of this thesis has focused on the design of a wideband nonlinear measurement setup, which allows for simultaneous LF and RF acquisition from DC up to 4 GHz. It has been shown that a direct RF sampling architecture is feasible for nonlinear measurement setups. The fast-paced development of high-speed ADC/DAC technology could, in the future, enable such architectures to be used for mm-wave characterization setups, for decreased complexity and cost, but also for better measurement flexibility. This will open up new possibilities for characterization of mm-wave circuits and components using wideband stimuli.

In a transmitter, the PA is in many cases considered to be one of the most challenging blocks to design due to the trade-offs between linearity, efficiency, bandwidth, power and gain. As demonstrated in paper [A], the wideband nonlinear measurement setup can aid in multi-band PA design. It is shown that concurrent multi-band operation changes the fundamental frequency loading requirements, and load-pull on the intermodulation products can yield some efficiency gains. Improvements to the usable dynamic range of the measurement setup was carried out in paper [B], through the use of
a correction table for compensation of the internal VGAs of the oscilloscope. In paper [D], the wideband nonlinear measurement setup was extended with low-frequency measurement capabilities, allowing for analysis of the effect of LF termination impedances on in-band distortion in a transistor. The simultaneous LF and RF measurement capabilities of the measurement setup were later used in paper [E] for large-signal characterization of an IQ-mixer. Through the use of incommensurate frequency grids, it was shown that I- and Q-branch imbalances result in differences in generated distortion and conversion gain from each branch. Furthermore, the two branches exhibit different compression characteristics and significant in-band distortion due to even order nonlinearities is present in the upconverted spectrum.

GaN-based electronics are attractive for many RF applications. A long-standing problem in GaN has, however, been dispersion due to electrical and thermal effects. Relevant characterization is needed partly for implementation of dispersive effects in models, but also for improvements of the epitaxial design through material comparisons. Dispersion due to trapping is not easily captured as it is highly dependent on the operating conditions of the transistor. Furthermore, thermal effects can have similar time constants as trapping phenomena, which must be considered when measuring. The characterization should to some extent be driven by the application need, in order to provide relevant information to the end user. Methods such as the modified $3\omega$-method in paper [F], enabled by the wideband nonlinear measurement setup, are of interest since they can characterize thermal properties in semiconductor materials with potentially lower influence of trapping.

To summarize, this thesis has presented characterization tools and methods which facilitate the development of high-performance microwave circuits and systems for next generation wireless applications.

### 5.2 Future Work

The work presented in this thesis paves the way for continued research within the following proposed topics:

- **Measurement-aided design using wideband stimuli**
  The capabilities of the wideband nonlinear measurement setup opens up for further research within wideband circuit design. Although the multi-band load-pull characterization in paper [A] has demonstrated the wideband capabilities of the measurement setup to a certain extent, the measurement scenario can be expanded. For example, wideband modulated signals within each operating band is a more realistic scenario. A study in how to optimally design wideband or multi-band PAs using measurement data could be of interest.

- **Modeling of frequency-translating devices**
  The simultaneous LF and RF acquisition of the wideband nonlinear measurement setup can be useful for modeling of frequency-translating devices such as IQ-mixers. This can result in more accurate models which take into account port impedances, compression characteristics, sources of distortion, and imbalances between the I- and Q-branches.
• **Characterization of dispersive effects**
The flexibility of the proposed measurement setup can be utilized to study electrical and thermal effects in devices using an arbitrarily defined stimulus. This can be exploited to design better experiments for separation of the different effects, or for more methodical analysis. For example, further development of pulsed I-V techniques, such as the double pulsed I-V measurement presented in [119], are of interest.

• **Methods for iso-trap thermal characterization**
The modified 3ω-method presented in paper [F] is promising for characterization of thermal properties under zero-bias conditions. Further verification of the method is however needed. Although a zero DC bias could potentially limit the influence of trapping, this must be experimentally verified.

• **Further studies on epitaxial optimization of GaN HEMTs**
Further GaN material studies can provide more insight into how various epitaxial parameters influence trapping phenomena and thermal performance. For example, a study specifically aimed at variations of the buffer could be useful for e.g. determining the optimal thickness for minimizing the thermal resistance, similar to what is shown with simulations in [114].
Chapter 6

Summary of appended papers

This chapter summarizes the publications included in this thesis. A short summary of each publication is presented as well as my contributions to each paper.

Paper A


An active load-pull setup with wideband measurement capabilities is presented. The setup is used for carrying out multi-band load-pull on a gallium nitride (GaN) high-electron mobility transistor (HEMT).

**My contribution:** I assembled the measurement setup, did the calibrations and measurements, and wrote the paper.

Paper B


A correction algorithm for real-time oscilloscopes is presented, for correcting magnitude offsets and phase shifts introduced in the RF front end of the RTO.

**My contribution:** I formulated the algorithm, did the measurements, and wrote the paper.

Paper C

Trapping in carbon-doped GaN HEMTs is investigated using various characterization methods. It is shown that the buffer doping can be tailored to achieve a good trade-off between leakage and trapping effects. This paper was done in collaboration with Linköping University.

My contribution: I did the majority of the characterization and wrote the majority of the paper.

Paper D


The measurement setup introduced in Paper A is extended with low-frequency (LF) measurement capabilities. This is achieved through custom-designed LF sense boards based on operational amplifiers.

My contribution: I designed the sense boards, did the measurements, and wrote the paper.

Paper E


The LF-extension introduced in Paper D is utilized for four-port large signal characterization of an IQ-mixer. Through the use of an incommensurate frequency grid at the I- and Q-ports, nonlinearities can be separated into contributions from the I- and Q-branches, and even order and odd order nonlinearities.

My contribution: I introduced the incommensurate frequency grids, did the measurements, and wrote the paper.

Paper F


Electrothermal performance in three GaN-on-SiC wafers is investigated using various characterization methods. A new method for extraction of thermal transfer functions is presented. This paper was done in collaboration with UMS in Ulm, Germany.

My contribution: I did the majority of the measurements, formulated the new method, and wrote the majority of the paper.
Acknowledgments

This work would not have been possible without the support from many people, whom I therefore would like to express my gratitude towards.

First of all, I would like to thank my examiner Prof. Herbert Zirath for giving me the opportunity to conduct my research at the Microwave Electronics Laboratory, and Prof. Jan Grahn for creating the open research collaboration between industry and academia through the GigaHertz Centre.

This Ph.D. journey would never have been initiated without the trust and support from my supervisor Asst. Prof. Mattias Thorsell and my co-supervisor Prof. Christian Fager. You gave me the opportunity to work part-time during my M.Sc. studies as an amanuensis, which later led to a Master’s Thesis and eventually the start of a Ph.D. career. Your guidance, knowledge and inspiration has been instrumental in my growth as a researcher and I’m very grateful to have you as supervisors.

I also consider myself lucky to have worked with talented colleagues which have given me invaluable support, cooperation and feedback over the years. I have enjoyed every day at work thanks to you. Special thanks to Prof. Niklas Rorsman and Asst. Prof. Koen Buisman for fruitful discussions and inspiration. William and Ibrahim, I really enjoyed sharing the office with you, and thanks for the encouragement and cheering support. Congratulations to Johan Bergsten who recently passed his Ph.D. dissertation, and a big thanks for sharing this journey since 2013. Thanks to Dhecha, for enjoyable discussions and for fun times when we shared the teaching in the Wireless Link Project. Also, thanks to Johan Bremer for both the entertaining and more serious conversations we have had. Henric, thank you for everything IT-related and for sharing most of the lunches I’ve had at MEL. There are more people both inside and outside of Chalmers that I have met and that have supported me - a big thanks to all of you!

To the life outside of work; thanks to my family for your endless support and thanks to my friends for making life more colorful. To Elin, I am extremely grateful to have had you by my side during this period. You have given me the strength to finish this journey. I love you!

This research has been carried out in the GigaHertz Centre in joint projects financed by Swedish Governmental Agency of Innovation Systems (VINNOVA), Chalmers University of Technology, RISE (previously SP Technical Research Institute of Sweden), ComHeat Microwave AB, Ericsson AB, Saab AB, Infineon Technologies AG, Mitsubishi Electric Corporation, NXP Semiconductors, SweGaN AB, and United Monolithic Semiconductors.
Bibliography


