Characterization of Trapping Effects, Robustness and Linearity in GaN HEMTs for High Power and Low Noise Microwave Amplifiers

OLLE AXELSSON

Microwave Electronics Laboratory
Department of Microtechnology and Nanoscience – MC2
Chalmers University of Technology
Gothenburg, Sweden, March 2016
Characterization of Trapping Effects, Robustness and Linearity in GaN HEMTs for High Power and Low Noise Microwave Amplifiers

OLLE AXELSSON

© Olle Axelsson, 2016
ISBN 978-91-7597-335-7

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 4016
ISSN 0346-718X

Technical Report MC2-332
ISSN 1652-0769

Microwave Electronics Laboratory
Department of Microtechnology and Nanoscience – MC2
Chalmers University of Technology
SE-412 96 Goteborg, Sweden
Phone: 46 (0) 31 772 10 00

Printed by Kompendiet
Gothenburg, Sweden 2016
Abstract

The Gallium Nitride high electron mobility transistor (GaN HEMT) is being established as a commercial technology and used in increasingly large volumes. However, the first generation of GaN MMIC is optimized towards high output power, and hence challenges remain in order to realize the technology’s full potential and expand its use towards e.g. linear transmitters and robust receivers. The linearity of GaN HEMTs is severely limited by dispersive effects. To reduce these effects, further optimization of the technology is needed. Furthermore, GaN LNAs could offer better survivability and linearity than competing LNA technologies.

Issues connected to dispersion and degradation are not always captured by traditional characterization procedures, making development of new characterization methods important. This thesis serves to understand the consequences of dispersive effects and robustness, as well as suggesting new ways of characterizing such effects.

First a study of the dispersive effects for different Fe doping profile in the GaN buffer is presented. Traditional characterization procedures such as measurement of S-parameters and pulsed and DC I-V characteristics show no significant differences. It is shown that characterization of recovery after a high power pulse captures important differences in RF gain recovery transients for different Fe profiles. Furthermore it is demonstrated that this effect can approximately be evaluated by DC-pulsing the drain voltage, and thus simplifying the measurement procedure.

The potential of GaN for highly robust and linear LNAs is also explored. The impact of forward current through the mesa sidewalls in mesa-isolated HEMTs is examined and found to be a limiting factor for LNA survivability. Severe temporary gain degradation after high input power stress due to trapping effects is found in GaN LNAs fabricated in several different processes. This is an important issue since LNAs often need to be operational immediately after stress.

The linearity of GaN LNAs is explored by analysing the sources of nonlinearity and compared with a GaAs pHEMT. It is concluded that GaN has superior linearity due to being biased at higher drain voltages and which also leads to higher power consumption. Two LNAs have been designed focusing on maximizing the output third order intercept to DC power consumption ratio. The resulting ratios of 12 and 19 are the highest reported values for GaN amplifiers in the 6-8 GHz frequency band.

Keywords: GaN HEMT, low-noise amplifier, electron trapping, intermodulation distortion, robustness
List of publications

Appended papers

This thesis is based on the following papers.


Other publications

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


Thesis


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

Abbreviations

2DEG          Two-dimensional electron gas  
AlN          Aluminum nitride  
AESA          Active, electronically scanned array  
DLTS          Deep level transient spectroscopy  
DPD          Digital predistortion  
GaN          Gallium nitride  
GaAs          Gallium arsenide  
HBT          Heterojunction bipolar transistor  
HEMT          High electron mobility transistor  
ITRS          International technology roadmap for semiconductors  
LNA          Low noise amplifiers  
MOCVD          Metalorganic chemical vapor deposition  
MMIC          Monolithic microwave integrated circuits  
NF          Noise figure  
OIP3          Output third order intercept point  
PA          Power amplifier  
PAPR          Peak to average power ratio  
pHEMT          Pseudomorphic HEMT  
SiC          Silicon carbide  
SiNx          Silicon nitride

Notations

\( \sigma_n \)          Electron capture cross-section  
\( \tau_n \)          Time constant  
\( a_n \)          Amplitude corresponding to time constant \( \tau_n \)  
\( c_n \)          Capture coefficient of electron trap  
\( C_b \)          DC blocking capacitance  
\( C_d \)          RF decoupling capacitance  
\( C_{gd} \)          Gate-source capacitance  
\( C_{gd2} \)          Coefficient in the Taylor expansion of \( Q_{gd}(V_{gd}) \)
\( C_{gd3} \) : Coefficient in the Taylor expansion of \( Q_{gd}(V_{gd}) \)

\( C_{gs} \) : Gate-source capacitance

\( C_{gs2} \) : Coefficient in the Taylor expansion of \( Q_{gs}(V_{gs}) \)

\( C_{gs3} \) : Coefficient in the Taylor expansion of \( Q_{gs}(V_{gs}) \)

\( C_{in} \) : Input capacitance

\( e_n \) : Emission coefficient of electron trap

\( E_C \) : Conduction band energy level

\( E_T \) : Trap energy level

\( f_{max} \) : Maximum frequency of oscillation

\( f_T \) : Cutoff frequency

\( g_0 \) : Degeneracy of trap level state in empty state

\( g_1 \) : Degeneracy of trap level state in filled state

\( g_m \) : Transconductance

\( g_{ds} \) : Output conductance

\( g_{m2} \) : Coefficient in the Taylor expansion of \( I_{ds}(V_{gs},V_{ds}) \)

\( g_{m3} \) : Coefficient in the Taylor expansion of \( I_{ds}(V_{gs},V_{ds}) \)

\( g_{md} \) : Coefficient in the Taylor expansion of \( I_{ds}(V_{gs},V_{ds}) \)

\( g_{d2} \) : Coefficient in the Taylor expansion of \( I_{ds}(V_{gs},V_{ds}) \)

\( g_{d3} \) : Coefficient in the Taylor expansion of \( I_{ds}(V_{gs},V_{ds}) \)

\( g_{m2d} \) : Coefficient in the Taylor expansion of \( I_{ds}(V_{gs},V_{ds}) \)

\( g_{md2} \) : Coefficient in the Taylor expansion of \( I_{ds}(V_{gs},V_{ds}) \)

\( I_{ds} \) : Drain-source current

\( I_{gs} \) : Gate-source current

\( I_{max} \) : Maximum drain current

\( K \) : Boltzmann’s constant

\( n \) : Free electron concentration

\( N \) : Trap state concentration

\( N_t \) : Trapped electron concentration

\( P_{dc} \) : DC power consumption

\( P_{in} \) : Input power

\( P_{out} \) : Output power

\( R \) : Resistance

\( R_{bias} \) : Bias resistance

\( R_L \) : Load resistance

\( T \) : Temperature

\( <v> \) : Root mean square electron velocity

\( V_{ds} \) : Drain-source voltage

\( V_{gs} \) : Gate-source voltage

\( V_{gst} \) : Gate-drain voltage

\( V_{supply} \) : Supply voltage
$V_{\text{stress}}$  
Forward voltage during stress
Contents

Abstract ......................................................................................................................................................... iii
List of publications ......................................................................................................................................... v
Abbreviations and notations ...................................................................................................................... vii

1 Introduction .............................................................................................................................................. 1

2 Trapping effects in GaN HEMTs ........................................................................................................... 5
   2.1 Physical origins ............................................................................................................................ 6
      2.1.1 Capture and emission of electrons .................................................................................... 7
      2.1.2 Surface traps ....................................................................................................................... 9
      2.1.3 Impact of buffer .................................................................................................................. 9
   2.2 Characterization of trapping effects ............................................................................................... 11
      2.2.1 Current collapse .................................................................................................................. 11
      2.2.2 Distortion and transient effects .......................................................................................... 13
      2.2.3 Physical properties ............................................................................................................. 16
   2.3 Chapter summary ............................................................................................................................ 17

3 GaN HEMT based LNAs ......................................................................................................................... 19
   3.1 Robustness of GaN LNAs ............................................................................................................. 22
      3.1.1 Failure mechanisms in GaN LNAs .................................................................................... 26
   3.2 Recovery time ............................................................................................................................... 33
      3.2.1 Device level characterization .............................................................................................. 36
   3.3 Linearity and intermodulation distortion ......................................................................................... 39
      3.3.1 OIP3 analysis on device level – GaN vs GaAs ................................................................. 41
      3.3.2 OIP3 analysis on circuit level ............................................................................................. 44
      3.3.3 Design of highly linear GaN MMIC LNAs ................................................................. 47
3.4 Chapter summary ................................................................. 48

4 Noise Parameter measurements .............................................. 51
  4.1 Noise parameter measurements ......................................... 51
    4.1.1 Noise temperature of an electronic source impedance tuner 52

5 Conclusions and future outlook ............................................. 55
Acknowledgment ........................................................................ 57
References .................................................................................. 59
Chapter 1

Introduction

Established transistor technologies can often be evaluated and compared by a defined set of figures of merit and a designer can make an informed choice using only a data sheet. However, bringing an immature technology from promising performance to commercialization takes decades. Even when traditional performance indicators show excellent figures, new technologies often suffer from anomalous behaviour, unreliability and processing variations that need to be overcome before the technology is adopted in high-volume products. While many of the problems that occur on the journey to commercialization are similar between different technologies, each new type of transistor also has its own set of challenges. Overcoming these involves improvement of fabrication to produce more ideal behaviour, but also development of characterization methods to quantify non-idealities, models to predict them and methods to define acceptable ranges for specific applications.

The AlGaN/GaN HEMT is already being established as a mature technology, but development is still ongoing and challenges remain in order to realize its full potential. The first reports of the formation of a two-dimensional electron gas (2DEG) in the AlGaN/GaN interface [1] and the first high electron mobility transistors based on this heterostructure [2] were published in 1992 and 1993, respectively. The combination of high breakdown field and thermal conductivity of the materials in the heterostructure and the high electron mobility in the 2DEG makes it ideal for microwave and RF power applications, and continued research has resulted in steady performance improvements [3].

Today, GaN HEMTs are commercially available as discrete transistors as well as monolithic microwave integrated circuit (MMIC) components and foundry services. Power amplifiers (PA) and high power switches are the two components where GaN HEMTs have most successfully been utilized at microwave and RF frequencies. The combination of high saturated power, large bandwidth, small footprint, low losses and high efficiency offered by GaN PAs and switches cannot be rivalled by any other technology. For active, electronically scanned arrays (AESA) antennas in radar and electronic warfare transmitters, GaN is already on the way to being established as the technology-of-choice. During 2015, defense industries such
as Saab and Raytheon have started introducing GaN MMIC PAs and switches in their AESAs [4]. For space applications its insensitivity to ionizing radiation makes it an attractive choice, in addition to its performance advantages [5]. Wireless infrastructure offers a potentially larger-volume market for GaN. However, competition is tough from lower cost alternatives such as LDMOS in the lower frequency bands and GaAs and Si CMOS and BiCMOS technologies at higher frequencies. Future communication standards are moving towards higher frequencies and increasing demands on the hardware in terms of bandwidth and efficiency. This might provide an opportunity for GaN technology to reach high-volume markets, by offering performance advantages which enable architectures and solutions which are otherwise not feasible [5].

For low noise amplifiers (LNA), GaN HEMTs offer noise figures comparable to GaAs pHEMTs, while offering advantages in robustness and linearity [6] which may potentially result in better system performance in radar or communication systems by relaxing requirements on limiters and filters. However, although a GaN MMIC LNA is marketed by Qorvo (TGA2611), the adoption of GaN LNAs in real products is far behind that of GaN PAs and switches.

While maturing into a commercial technology, GaN technology is also being pushed to higher and higher frequencies, using transistors with shorter gate lengths and more advanced epistructures. The challenge is to, as stated in the 2013 International Technology Roadmap for Semiconductors [7], “increase their functionality in terms of operating frequency and modulation schemes while simultaneously meeting increasingly stringent linearity requirements at the same or lower cost”. This includes maintaining stable and reliable operation with new and down-scaled device structures and tailoring of transistor characteristics to obtain optimal efficiency and linearity in specific transmitter architectures.

This thesis aims to provide understanding of the impact on circuit and system performance from several properties of GaN HEMTs, and means to evaluate these effects on device levels. These properties are dispersive behavior, failure mechanisms and the nonlinear I-V characteristics. The focus is on PA and LNA applications, with special attention given to the latter, a previously less explored area.

Dispersive behavior in GaN HEMTs is one important challenge to overcome. GaN HEMTs suffer from discrepancies between their DC characteristics and their behavior at high frequencies [8]. Different manifestations of these effects affect performance in different ways. On slow time scales, changes in device performance over time leads to unreliable performance and discrepancy between models and actual devices. On faster time scales, GaN HEMTs suffer from decreased RF current and voltage swing under large signal conditions, compared to what the DC drain current-voltage characteristics would suggest, commonly referred to as current collapse [8]. This leads to reduction of efficiency and maximum output power.
Furthermore, dispersive behavior causes distortion of modulated signals which may be hard to correct for in systems with high linearity requirements. Development of SiN\(_x\) passivations have decreased these discrepancies and lead to improved RF performance [3, 9].

Device level evaluation of the effects of trapping on system performance requires new characterization procedures, since traditional characteristics such as DC I-V characteristics and S-parameters do not give information about these effects. Large signal current collapse is characterized using pulsed I-V and load-pull measurements, but effects on distortion and linearizability are difficult to evaluate without modulated measurements on circuit level. There is thus a need for new, application relevant characterization procedures and figures of merit that give information about dispersive effects and their influence on linearizability, both from a process optimization perspective and when evaluating technologies from a system perspective.

Chapter 2 of this thesis deals with trapping effects in GaN HEMTs, their consequences, physical causes and methods to characterize them. The main contributions of Paper A are summarized. The paper presents a comprehensive characterization of the effects of buffer Fe doping profile on transient effects as well as CW performance. Furthermore, it is shown how device level measurements can be used as an indicator of how dispersive behavior in the transistor affects distortion and linearizability of modulated signals.

Chapter 3 focuses on characterization of three important factors for GaN LNAs, namely robustness, temporary degradation due to trapping effects and linearity. Robustness is important since the GaN HEMT’s ability to survive adverse operating condition is one of its main competitive advantages. However, the limiting factors for the maximum input power to GaN LNAs are not well understood, since most studies on LNA level have not analyzed the causes of failure in detail while studies of physical degradation mechanisms have focused mainly on PA and switching operation [10]. This is an important impediment to widespread adoption since uncertainty about the reliability necessitates overly conservative ratings which hinders the technology from being used as its full potential. There is thus a need for application relevant testing procedures and specifications on device level, which can easily be translated to circuit level ratings using standard circuit design computer-aided design tools. In section 3.1, GaN HEMT robustness under LNA-like conditions is characterized both on circuit level and device level and different known degradation conditions are discussed in relation to robust LNA operation. An investigation of noise figure and gate leakage degradation due to forward current in GaN HEMTs using mesa isolation is presented. It is proposed that current through the mesa sidewalls may limit the survivability of LNAs using these devices. DC and
RF measurements on device level are used to estimate LNA robustness together with circuit simulations. The study is also reported in Paper B.

Temporary degradation in robust GaN LNAs during the first micro- and milliseconds after input power stress is another topic which has previously been studied only very little [11, 12]; most reports evaluating robustness of GaN LNAs have only tested for permanent degradation. However, some radar and electronic warfare applications require LNAs to be operational immediately after a high input power pulse. It thus important to evaluate whether high input powers into GaN LNAs lead to temporary performance degradation, how quickly they recover from such degradation and how the characteristics of the transistor and the LNA design influence these parameters. In this thesis it is demonstrated for the first time that trapping effects do indeed affect the performance after a high power pulse. Recovery times are the millisecond range, which is much too slow for many applications where LNAs may be disturbed by pulsed signals with pulse repetition frequencies in the kHz range or higher. Measurements on LNA as well as device level show that the gain drop after the pulse varies greatly between different GaN technologies and the interface between the AlGaN barrier and that both the GaN channel and the buffer design seem to be important.

GaN LNAs have also shown very high linearity in terms of output third order intercept point (OIP3), with reported values close to 50 dBm combined with sub-dB noise figures being reported [13, 14]. This is more than an order of magnitude higher than what is typically achieved by commercial GaAs pHEMT LNAs. Higher OIP3 could enable new receiver architectures e.g. in multiband communication links where a single LNA is used for several bands with sufficiently low intermodulation distortion (IMD). However, improvements in OIP3 have come at the cost of higher power consumption, since the best linearity performance is achieved at large drain bias voltages. To investigate whether this high linearity can be achieved with lower power consumption, Volterra series analysis is here used analyze the sources of intermodulation distortion in a GaN HEMT, and the results are compared to a GaAs pHEMT. Paper D presents design and characterization of to GaN MMIC LNAs focusing on high OIP3, low noise figure and low power consumption.

In chapter 4 and Paper E, the suitability of electronic tuners for measurement of noise parameters is evaluated, by measuring the excess noise produced by a tuner, compared to the thermal noise at room temperature. Noise parameter measurements are important for modeling noise performance during LNA design.

Finally, chapter 5 summarizes the most important conclusions from my work and gives recommendations for future research into these and related subjects.
Chapter 2

Trapping effects in GaN HEMTs

Issues with dispersive behavior in GaN HEMTs are some of the most important challenges facing the technology. These issues are caused by the existence of trap states within the band gap, most importantly in the GaN buffer and near the AlGaN surface. The interaction between the electrons in the conduction band and trap states alter the electrical characteristics of a transistor by introducing trapped charges that change the potential profile in the channel. The charging and discharging of these traps modulate the conductivity of the channel and their occupancy in turn modulated by the applied voltages as well as external factors such as temperature and radiation.

Depending on the location and physical properties of the trap level, as well as the transistor operating conditions, trap levels can affect performance in various ways. Early on, a major issue was a reduction of maximum current and increase in knee voltage in GaN HEMTs when operating as a power amplifier biased at a high drain voltage, compared to DC conditions [8]. This is illustrated in Fig. 2.1. In order to provide maximum amount of output power, a PA needs to maximize the amplitudes of both the voltage and current waveforms. These circuits are therefore biased at high drain voltages while the output matching is designed so that the transistor swings into the high current-low voltage (knee) region during large signal operation [15]. However, when the HEMT is biased at a high voltage, the population of trapped charges in the buffer and at the surface increases as highly energetic electrons are injected into these regions [16]. These trapped charges act as a virtual gate, reducing current and increasing the on resistance. Furthermore, the time constants of detrapping processes are typically much longer than the period of an RF signal, so the amount of electron trapping is determined mainly by the DC bias point, rather than the instantaneous voltage in the knee region [8]. The presence of traps thus reduces the swing into the knee region when operating from a high drain voltage bias point, and therefore the maximum output power delivered by the transistor.
Potentially more severe problems arise when trying to transmit a complex signal with high linearity requirements, such as in radio base stations of wireless telecommunications systems. In order to increase the spectral efficiency, modern communication standards such as LTE have moved to modulation formats resulting in signals with increasing peak-to-average power ratios (PAPR). In the time domain, these signals look like irregularly pulsed waveforms with power levels between the pulses much lower than the peak levels. If the trapping effects in a GaN PA are modulated by the output power level and the emission time constants are in the same range as the time between pulses, the nonlinear PA transfer function varies with time after a high power pulse, leading to distortion of the output signal. Aside from distorting the information contained in the signal, the signal also expands outside its assigned frequency band, disturbing communication in neighboring bands. Linearity requirements of transmitters in communications systems are therefore very strict.

The standard method to compensate for distortion due to nonideal hardware in radio base station transmitters is digital predistortion (DPD) [17]. These circuits use digital signal processing to create a pre-inverse of the transmitter’s transfer function and thus produce the desired output signal. DPDs can be used to compensate for memory effects as well as static nonlinearities. However, computational complexity and thereby power consumption increases as the DPD need to take into account memory effects with a wide range of time constants and their interactions with other nonidealities in the components [18].

2.1 Physical origins

Traps in GaN HEMTs originate from defects in the crystal structure in various parts of the transistor. Defects in a crystal are deviations from its periodic crystal
structure, most commonly for GaN the Wurzite structure. Many types of defects have been observed in GaN and these introduce energy levels that would normally be within the band gap of an ideal GaN crystal. Defects which involve only a single or a few atoms and do not extend in space in any dimension are called point defects. Some important point defects in GaN crystals are [19]:

- Nitrogen vacancies, acting as shallow donors.
- Contaminants such as oxygen, silicon or carbon replacing either a nitrogen or gallium atom in the lattice. Oxygen replaces a nitrogen atom and acts as a donor while carbon acts as an acceptor when replacing a nitrogen atom, which is the most common case. However, some carbon atoms will are also incorporated on gallium sites, in which case they act as donors.
- Intentional doping. Shallow donors can be used to increase the conductivity and electron concentrations whereas deep acceptors have the opposite effect.
- Antisites – A nitrogen atom occupies a place normally occupied by a Gallium atom (acting as an acceptor), or vice versa (donor).

Aside from point defects, there are also structural defects such as threading dislocations originating from the interface between the GaN buffer and the substrate, with a thin AlN nucleation layer between. This can be caused by, for example, mismatch in lattice constants or thermal expansion coefficients between the substrate, the nucleation layer and the buffer. These defects extend through the crystal, in some cases all the way to the surface. Electrons belonging to atoms close to the dislocation lines bond to the environment differently than electrons in a regular lattice, which changes their energy levels and may cause point defects to be gathered along the dislocations.

### 2.1.1 Capture and emission of electrons

Trap levels introduced by defects interact with the conduction band by capture and emission of electrons. In the case of a donor trap, the effect is an increase in free electrons and a positive trapped charge. An acceptor trap, on the other hand, reduces the number of electrons in the conduction band while introducing a trapped charge with a negative sign. If the traps are deep (acceptor energy levels higher than the Fermi level and donor levels below the Fermi level), only a fraction of them are ionized, whereas shallow traps are normally almost fully ionized.

Consider the case of electrons being injected into a region with traps, for example by a voltage pulse starting at \( t = 0 \). The capture and emission process can be
described by Shockley-Read-Hall statistics [20, 21]. The rate of capture is proportional to the amount of free electrons as well as to the number of unfilled trap states, while the emission rate is proportional to the number of trapped electrons and number of unfilled states in the conduction band. As long as only a small fraction of the states in the conduction band are occupied, this is almost independent of the carrier concentration. We assume the trap is in the upper half of the band gap and do not consider any interaction between the trap level and the valence band. The concentration of trapped electrons \( N_t \) is then given by the differential equation

\[
\frac{dN_t}{dt} = (N - N_t)c_n n - e_n N_t
\]  

(2.1)

Where \( N \) is the total concentration of trap states, \( n \) is the free electron concentration, \( e_n \) is the emission coefficient of the trap level and \( c_n = \sigma_n \langle v \rangle \) is the capture coefficient where \( \sigma_n \) is the trap’s capture cross section and \( \langle v \rangle \) is the mean velocity of electrons in the conduction band. The solution is, if we assume that \( n \gg N \) so that the free electron concentration is not much affected by the trapping process:

\[
N_t(t) = \frac{c_n n N}{c_n n + e_n} \left(1 - \exp\left[-(c_n n + e_n)t\right]\right) + N_t(0) \exp\left[-(c_n n + e_n)t\right]
\]  

(2.2)

With \( N_t(0) \) being the concentration of trapped electrons before the pulse. As \( t \to \infty \), \( N_t \) approaches the value:

\[
N_t(\infty) = \frac{c_n n N}{c_n n + e_n}
\]  

(2.3)

If the voltage returns to its original value and the injected electrons are swept away from the region, the electron concentration decreases close to zero and the trapped electrons are emitted. (2.1) now reduces to:

\[
\frac{dN_t}{dt} = -e_n N_t
\]  

(2.4)

With the solution:

\[
N_t(t) = N_t(0) \exp[-e_n t]
\]  

(2.5)

\( t = 0 \) is now at the end of the voltage pulse. The emission coefficient \( e_n \) is given by:

\[
e_n = \frac{g_0}{g_1} \sigma_n y T^2 \exp \left[ -\frac{E_C - E_T}{kT} \right]
\]  

(2.6)
Where \( g_0 \) and \( g_1 \) are the degeneracy of the trap level in its empty and occupied state, respectively, \( \gamma \) is a material constant and \( E_c - E_T \) is the energy difference between the trap level and the conduction band, referred to as the activation energy of the detrapping process. \( \tau = 1/e_n \) is referred to as the time constant of the detrapping process. Transient measurements of trap emission at different temperatures can thus be used to determine the energy level and capture cross section of traps. This will be described more in detail in section 2.2.

In GaN HEMTs, capture processes are typically very fast, often in the nanosecond range, whereas detrapping can be much slower. For modulated signals, the trapping can therefore often be considered instantaneous whereas the time constants of the detrapping process need to be studied in detail in order to accurately model the transistor behavior [22, 23]

2.1.2 Surface traps

Trapping of electrons from the gate at the AlGaN surface has been identified as a major contributor to the current collapse in HEMTs [8]. It has been found that deposition of a SiN\(_x\) layer on top of the AlGaN surface significantly reduces the current collapse, while also improving the breakdown voltage [9]. Optimization of passivation procedures [24] and development of field plates that decrease the peak electric field at the gate edges have led to large reductions of current collapse and improved in the output power capabilities of GaN HEMTs [3, 25]. While the origin of these effects and the mechanism by which the passivation works is still debated, it has recently been suggested that ambient water molecules bonding to the atoms at the surface and absorbing electrons from the gate is a major reason [26]. In that case, the SiN\(_x\) layer reduces trapping by shielding the gate edge and the surface from water molecules. Other mechanisms that have been proposed are that Si atoms are incorporated at the surface replacing the surface traps with a shallow donor whose occupancy is less influenced by the electric field [27], or that the passivation provides an escape path for trapped electrons [24, 28].

2.1.3 Impact of buffer

Because of the limited availability and high cost of native GaN substrates, GaN HEMTs are most commonly grown on Si, Sapphire or SiC substrates. SiC provides the best crystal quality while also taking advantage of its excellent thermal conductivity and high resistivity, while Si is also being explored due to its lower cost. Because of the lattice mismatch between SiC and GaN and to prevent diffusion of Si into the GaN layers, an AlN nucleation layer and a thick GaN buffer are needed between the substrate and the HEMT structure reducing the number of defects [29].
The properties of the GaN buffer have important implications for electrical performance as well as dispersive effects in GaN HEMTs. Doping of GaN with a deep acceptor was initially used as a method of compensating for shallow donors and maintaining high resistivity in GaN on sapphire, which otherwise show n-type behavior [30-32]. Although improvements in growth techniques of GaN on SiC have made it possible to obtain good quality, resistive crystals without deliberate compensation doping, deep acceptor doping of the buffer still has the advantage of creating a back-barrier underneath the channel which improves electron confinement close to the AlGaN/GaN interface. This improves the pinch-off, reduces short-channel effects and output conductance and increases the drain-source breakdown voltage, leading to improved high power performance at RF and microwave frequencies [33].

Incorporation of acceptors in the buffer can be controlled indirectly by varying the pressure during the MOCVD growth of the buffer, with lower pressure causing higher dislocation and impurity density, increasing the resistivity [32]. Impurities can also be incorporated by actively adding a dopant to the MOCVD reactor. Although there has been some interest in Carbon (C) [34] and Magnesium (Mg) [35] doped buffers, the most common dopant in commercial HEMTs is Iron (Fe) [30, 31].

However, doping of the buffer has been seen to correlate with dispersive effects such as current collapse [34, 36, 37]. To some extent, this behavior is unavoidable, since any dopant acts also as an electron trap whose occupancy is modulated by operating conditions. In the case of Fe doping, a trap with activation energy around 0.6 eV located in the buffer between the gate and the drain has been seen to increase in importance with increasing Fe concentrations [37-40]. However, similar trap levels have been seen also in GaN HEMTs with no Fe doping and it is believed that the trap could be an intrinsic defect which is indirectly influenced by the Fe presence rather than the Fe atoms themselves. In [41] an increased density of threading dislocations in Fe doped GaN grown on sapphire was reported, compared to n-type GaN. Several papers [37, 42-45] have also noted that the activation energy of this trap corresponds to the level of a nitrogen antisite, as calculated in [46].

Leaving the region closest to the channel undoped reduces the current collapse and may be a good trade-off in order to reduce dispersive effects while maintaining a resistive buffer [30, 31, 34, 47]. However, Fe doping suffers from memory effects during MOCVD growth leading to a slow roll-off of the Fe concentration after the Fe source is turned off, rather than a sharp transition. This is believed to be due to Fe accumulating at the GaN surface and being incorporated in the above layers upon further growth [30, 31]. Carbon doping, by contrast does not suffer from this effect, allowing more freedom in the buffer design. A few studies have investigated the trade-offs involved in the design of the buffer doping profile in Fe [37, 48] and C [34,
doped GaN HEMTs, but it is still not well understood how different performance metrics are affected by the profile.

Paper A presents a comprehensive characterization on the effects of different variations in the Fe doping profile in GaN HEMTs fabricated by UMS. Fig. 2.2 shows the tested variation in Fe concentration and thickness of the Fe-doped layer. The effects on trapping as well as on small signal and DC performance were measured. While the s-parameters as well as the pulsed and DC I-V characteristics showed only very small differences in the tested range of Fe profiles, more detailed characterization of the transient behavior after trapping is induced showed significant differences, which will be described later in this chapter.

![Variations in buffer doping profiles of the GaN HEMTs tested in Paper A. The devices are fabricated by UMS on Cree wafers.](image)

**2.2 Characterization of trapping effects**

**2.2.1 Current collapse**

From an applications perspective, the simplest method to characterize large signal current collapse is to measure the transient in drain current in the knee region after a step from either a lower gate voltage (gate lag measurement) or a higher drain voltage (drain lag measurement) [8]. These measurements give information not only on the current collapse but also on the emission times of the traps involved. If the gate and the drain voltages are simultaneously pulsed, the complete output I-V plane can be characterized when pulsed from different bias points [3, 8, 49]. If the pulses are short enough to not allow any emission or capture of electrons during the pulse, the trap occupancy is only a function of the quiescent bias point. Since capture processes in GaN are very fast, this is not easily achieved when pulsing from a bias point with less trapping to one with more trapping (pulsing to higher drain voltage or lower gate voltage). However, when pulsing to a state with less trapping, pulse lengths around one microseconds are often sufficient. These measurements can give
an indication of the performance deterioration due to trapping effects, by comparing the pulsed I-V from a typical power amplifier bias point to that from a reference point where both the drain and gate terminals are unbiased [23]. Fig. 2.3 shows pulsed I-V measurements of transistors with the three buffer doping profiles shown in Fig. 2.2. Pulsing from $V_{ds}=V_{gs}=0$ V represents a case with little trapping at the quiescent point, to which the other pulsed measurements are compared to evaluate the current collapse. Pulsing from $V_{ds}=0$ V and $V_{gs}=-4$ V (below pinch-off) activates trapping effects due to gate lag but not due to drain lag, and results in only a very small current slump, around 3 % compared to the (0,0) measurement. Pulsing from a class-C bias point ($V_{gs}=-4$ V, $V_{ds}=20$ V) approximates the current collapse during PA operation and activates both drain-related and gate-related trapping effects. This measurement shows a current slump around 15 %. The current slump is larger for the more highly doped standard material, but the differences are very small and not larger than differences between nominally identical devices.

![Fig. 2.3: Pulsed I-V measurements of transistors with the three buffer doping profiles shown in Fig. 2.2, from three different bias points.](image)

Direct characterization of the large signal current collapse can be done using load-pull measurements with a measurement system that can measure nonlinear waveforms in the transistor plane. The drain current-voltage waveform can then be studied directly under PA-like conditions for different load impedances, and the swing into the knee region can be directly compared to the DC or pulsed I-V characteristics [50].
2.2.2 Distortion and transient effects

Effects on distortion and linearizability of modulated signals due to trapping are more difficult to characterize on device level, because of the complex interplay between the characteristics of the transmitted signal, the DPD algorithm, the static nonlinearity of the PA and the memory effects. Testing a PA in an actual system with a realistic modulated signal gives information about the level of distortion, but it is hard to assess what can be done to improve it and how specific characteristics of the transistor, the circuit design and the DPD algorithms impact linearity.

However, some qualitative insights about distortion due to transient effects and which time scales these effects operate in can be gained from measuring the small signal gain transient after a high output power pulse. This was done in paper A on the tested UMS devices with different buffer doping. An IQ modulator was used to provide a pulsed RF signal to a matched transistor, with the power level backing off by 15 dB between the pulses. This allows measurement of the small signal gain after the pulse, using a signal analyzer to measure the RF output power transient. Fig. 2.4 shows the measured gain transients after a 300 μs pulse at 3 dB compression, plotted on a logarithmic time axis with the end of the pulse at $t=0$. The three materials show similar levels of gain drop, around 0.5 dB, after the pulse, but the recovery is faster for the more highly doped standard material.

The drain current transients after the pulse were also measured, shown in Fig. 2.5a. These measurements show that the current also experiences a drop after the pulse, with a recovery time in the same range as the gain recovery. The differences between the buffer profiles are more marked in the drain current, as the HEMT with the more highly doped buffer recovers close to its original value in around 10 ms whereas the two other devices take 100 ms to reach the same value. The correlation between the drain current transient and the gain transient indicates that measurement of drain current can be used to probe the effects on gain in cases where it cannot be measured directly.

To investigate how accurately the gain transients can be predicted from the drain current transient, a very simplified model of the effect of trapping on the s-parameters of the transistor is used. Firstly, it is assumed that only $S_{21}$ changes enough to significantly affect the gain. Secondly, it is assumed that $S_{21}(t)$ after the pulse is the same as the static $S_{21}$ at the same drain voltage but a lower gate voltage corresponding to $I_{ds}(t)$. This is equivalent to assuming that the trapped charge acts as a shift in the effective gate voltage, which is commonly used in modelling of trapping effects [23]. The dynamic $S_{21}(t)$ can then be estimated from static bias dependent s-parameter measurements. Lastly, it is assumed that the amplifier gain varies with $S_{21}$ as $|S_{21}|^2$, i.e. the effect on input and output matching is negligible. The black curves in Fig. 2.4 show the gain transients estimated from drain transient measurements in comparison
to the measured transients. The simplified model does not exactly predict the measured transients but is accurate enough to give a rough prediction of the gain drop as well as the recovery time.

It was also investigated whether DC pulsing the drain voltage produces the same effect on the drain current as an RF pulse, since this would facilitate characterization of these effects significantly. Fig. 2.5b shows the drain current transient after a pulse to $V_{ds} = 40$ V from the $V_{ds} = 20$ V and a quiescent drain current of 100 mA, the same bias point used in the pulsed RF measurements. The similarity between Fig. 2.5a and b shows that DC pulsing can be used to mimic the effects of an RF pulse, in order to give a rough estimation of the RF gain transients. This provides a simple way to get a first indicator of the impact of trapping effect on linearizability, both from a process optimization perspective and when evaluating a technology for a certain application.
The differences between the materials can be analyzed more in detail by looking at the time constants in the recovery process. As we saw in (2.5), the trapped charge decays exponentially after free electrons are removed from the trapping region, so we model the drain current transient as a sum of exponentials, in order to account for several different emission processes:

\[
I_{ds}(t) = I_{dsq} + \sum a_n \exp\left[-\frac{t}{\tau_n}\right]
\]  

(2.7)

A good fit to the measured transients for \( t > 100 \) μs was obtained using two terms in (2.7). The black lines in Fig. 2.5 show the fit of the model (2.7) to the measured transients. A third constant in the range of 10 μs was needed for some of the measurements but the corresponding amplitude was small and difficult to extract since the data is noisy for very fast processes. The analysis will therefore focus on the two dominant time constants \( \tau_1 \) and \( \tau_2 \), (where \( \tau_1 \) is the larger one). Fig. 2.6 shows the extracted time constants \( \tau_1 \) and \( \tau_2 \) and their corresponding amplitudes \( a_1 \) and \( a_2 \), as a function of the output power level in the RF pulsed measurements. Fig. 2.7 shows the parameters as a function of the voltage pulse step \( \Delta V_{ds} \) for DC pulsing. The extraction shows that the same time constants are present in all materials while the differences are due to different amplitudes of the two processes. Higher Fe concentration seems to enhance the faster emission process and decrease the slower one, leading to faster recovery. The total amount of trapped charge is largely unaffected by the Fe concentration, since the sum of the two amplitudes are similar for all tested devices.

Fig. 2.6: Parameters extracted by least square optimization of the model (2.1) against the measured transients after a 300 μs pulse to higher output powers at a 20 V, 100 mA bias point.
Fig. 2.7: Parameters extracted by least square optimization of the model (2.1) against the measured drain current transients after a 1 μs pulse to higher drain voltages ($\Delta V_{DS} = V_{p} \Delta V_{DS}$ from a 20 V, 100 mA bias point.

### 2.2.3 Physical properties

On a physical level, traps are characterized by their energy level and their capture cross-section. As mentioned in section 2.1.1, these can be determined from measuring the time constants of the emission process at different temperatures after stimulating trapping of electrons. This is done using an Arrhenius plot. Rewriting (2.6):

$$e_n = \frac{1}{\tau} = \frac{g_0}{g_1} \sigma_n T^2 \exp \left[ -\frac{E_C - E_T}{kT} \right]$$

$$\ln(\tau T^2) = \frac{E_C - E_T}{kT} - \ln \left( \frac{g_0}{g_1} \sigma_n \right)$$

The Arrhenius plot is constructed by plotting $\ln(\tau T^2)$ on the y-axis and $1/kT$ on the x-axis. The activation energy $E_C - E_T$ is then obtained from the slope of the line whereas the capture cross-section can be determined approximately from the constant term, if the other constants are known.
Different methods can be used to probe the trapped charge transient and determine the time constants. Deep level transient spectroscopy (DLTS) [51], using measurements of capacitance over a Schottky diode after a voltage pulse, is perhaps the most common. Drain current transients [52], low frequency s-parameters [53] and low frequency noise measurements [54] are also used.

Fig. 2.8 shows Arrhenius plots for the two time constants extracted from drain current transient measurements in the transistors studied in paper A. The plots reveal that the two time constants have very similar activation energies, and the energy level is also close to the trap level identified in many other studies as being correlated to Fe doping of the GaN buffer [37-40]. The time constants are also very similar between the different doping profiles. Studies by Uren et. al. on carbon doped GaN buffers [55] have showed several time constants with the same activation energy can be due to different vertical and lateral transport mechanisms for electrons and holes in the buffer. Fe doped buffers differ from Carbon doped buffers in that the Fe acceptor level is in the upper part of the band gap leading to smaller hole concentrations in the buffer, but it is possible that similar effects are present in Fe doped buffers.

![Fig. 2.8: Extraction of the activation energies E1 and E2 of the detrapping processes behind the two time constants $\tau_1$ and $\tau_2$, for a drain voltage pulse from 20 V to (a) 30 V and (b) 40 V at a quiescent drain current of 100 mA. The extraction is made from drain current transient measurements at different temperatures. The different colors represent transistors with the different buffer doping profiles shown in Fig. 2.2.](image)

### 2.3 Chapter summary

This chapter has described some issues with trapping effects in GaN HEMTs their physical origins and methods to characterize them. It was proposed that pulsing of the DC drain voltage to a higher value from its quiescent point can be used to evaluate transient effects on the gain due to an output power pulse, which in turn is an indicator of distortion of modulated signals. This method was used to investigate
the effect of variations in the buffer Fe doping concentration in GaN HEMTs. It was found that the static characteristics as well as the current collapse was affected very little by the doping concentration in the tested range, but the recovery is faster in the more highly doped samples. This was explained by the presence of two different emission processes with the same activation energy, but different time constants. Higher doping concentration enhances the faster process, while it suppresses the slower one, making the recovery faster.
Chapter 3
GaN HEMT based LNAs

Although the GaN HEMT has primarily been considered as a high power technology for PAs and switches, its ability to operate at very high voltages combined with low noise figure makes it an interesting candidate also for LNAs in certain applications. Compared to other LNA technologies such as GaAs pHEMT, robustness and linearity are the main advantages of GaN LNAs. A very robust LNA relaxes the requirements of protecting circuitry between the receiving antenna and the LNA. A very linear LNA may enable simpler and more sensitive receiver architectures, for example in multiband communication links. Using a single LNA and mixer for multiple bands eliminates the need of narrowband filters before the LNA and the filtering can be done at baseband instead. This improves performance and decreases cost, since these microwave filters are both expensive and lossy. However, a multiband receiver needs to be very linear in order to not cause intermodulation distortion (IMD) between signals in different bands.

The performance in terms of noise figures, maximum safe input power and output IP3 of published GaN based LNAs is listed in Table 3.1. The robustness is typically tested by subjecting the LNA to increasing input power levels until degradation or failure occurs. As seen in Table 3.1, several reports have shown GaN LNAs surviving input powers higher than 40 dBm without permanent damage. By contrast, commercial GaAs LNAs are commonly rated at maximum input powers in the range of 20 dBm. However, LNAs in GaAs technology with integrated limiters [56] provide survivability and noise figures in the same range as state-of-the-art GaN LNAs. The additional advantage of GaN LNAs over an integrated GaAs limiter/LNA is the possibility to integrate it on the same chip as GaN switches and PAs [57] which enables highly integrated transceiver front-ends with superior performance. These are useful in wideband active, electronically scanned array (AESA) antennas where a high number of transceiver modules are needed in a limited amount of space.

The values of OIP3 and noise figure from Table 3.1 are plotted against frequency in Fig. 3.1 (a) and (b), and compared with commercial GaAs HBTs and pHEMT amplifiers from Analog Devices and Qorvo. In terms of noise figure, the best results on GaN LNAs are comparable to commercial GaAs pHEMT LNAs. GaN LNAs
combine low noise figures with OIP3 larger than 40 dBm, with [14] being perhaps the most impressive example reporting sub-dB noise figure between 1-8 GHz combined with an OIP3 of 46 dBm. By contrast, GaAs LNAs typically have OIP3 figures around 30 dBm in the same frequency range.

Table 3.1
Survey of published GaN amplifiers

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[58]</td>
<td>11</td>
<td>1-4</td>
<td>41</td>
<td>400/15</td>
<td>2.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[58]</td>
<td>15</td>
<td>1-4</td>
<td>41-44</td>
<td>400/15</td>
<td>4.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[59]</td>
<td>20</td>
<td>42-47 (45)</td>
<td>4</td>
<td>28</td>
<td>35^2/5</td>
<td>3.6^2</td>
<td></td>
</tr>
<tr>
<td>[60]</td>
<td>13</td>
<td>1-25 (22)</td>
<td>4.6</td>
<td>34</td>
<td>75/12</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>[14]</td>
<td>20-10</td>
<td>1-8</td>
<td>1</td>
<td>46</td>
<td>400/15</td>
<td>6.6</td>
<td></td>
</tr>
<tr>
<td>[13]</td>
<td>20</td>
<td>0.25-3</td>
<td>3-4</td>
<td>52</td>
<td>750/40</td>
<td>5.3</td>
<td></td>
</tr>
<tr>
<td>[61]</td>
<td>20</td>
<td>0.5-3.5</td>
<td>2.8</td>
<td>50</td>
<td>500/40</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>[62]</td>
<td>15</td>
<td>0-24</td>
<td>4</td>
<td>41</td>
<td>300/30</td>
<td>1.4</td>
<td></td>
</tr>
<tr>
<td>[62]</td>
<td>12</td>
<td>0-20</td>
<td>6</td>
<td>43</td>
<td>400/30</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>[63]</td>
<td>13-7</td>
<td>4-8 (6)</td>
<td>1.9</td>
<td>24</td>
<td>30/4</td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td>[64]</td>
<td>18</td>
<td>3</td>
<td>3</td>
<td>39</td>
<td>105/20</td>
<td>3.8</td>
<td></td>
</tr>
<tr>
<td>[65]</td>
<td>9</td>
<td>0.5-6</td>
<td></td>
<td>49</td>
<td>650/28</td>
<td>4.4</td>
<td></td>
</tr>
<tr>
<td>[66]</td>
<td>13</td>
<td>0-3.5</td>
<td></td>
<td>49</td>
<td>450/28</td>
<td>5.6</td>
<td></td>
</tr>
<tr>
<td>[67]</td>
<td>15</td>
<td>1.6-2.3</td>
<td></td>
<td>49</td>
<td>1100/15</td>
<td>4.8</td>
<td></td>
</tr>
<tr>
<td>[68]</td>
<td>14</td>
<td>2.14</td>
<td></td>
<td>48</td>
<td>240/48</td>
<td>5.5</td>
<td></td>
</tr>
<tr>
<td>[69]</td>
<td>11.8</td>
<td>2.5</td>
<td>2.7</td>
<td>33</td>
<td>50/14</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>[70]</td>
<td>18</td>
<td>2</td>
<td>2</td>
<td>54</td>
<td>864/28</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>[71]</td>
<td>11</td>
<td>6</td>
<td>2.7</td>
<td>20</td>
<td>10/6</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>[72]</td>
<td>10</td>
<td>0.75-13.5</td>
<td>3</td>
<td>23/1.5</td>
<td>36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[73]</td>
<td>17</td>
<td>1-8</td>
<td>2</td>
<td>29-32</td>
<td>80/10</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>[74]</td>
<td>10</td>
<td>4-16 (7)</td>
<td>3.5</td>
<td>24</td>
<td>50/3</td>
<td>1.7</td>
<td>37</td>
</tr>
<tr>
<td>[75]</td>
<td>21</td>
<td>6</td>
<td>1.2</td>
<td></td>
<td>42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[75]</td>
<td>20</td>
<td>14</td>
<td>1.9</td>
<td></td>
<td>42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[76]</td>
<td>18</td>
<td>28</td>
<td>4.0</td>
<td></td>
<td>28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[77]</td>
<td>18</td>
<td>8-19</td>
<td>2.8</td>
<td>38</td>
<td>300/15</td>
<td>1.4</td>
<td>40</td>
</tr>
<tr>
<td>[78]</td>
<td>23</td>
<td>8-11</td>
<td>2</td>
<td></td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[79]</td>
<td>23</td>
<td>2-20 (4.5)</td>
<td>3.3-4.7</td>
<td>29</td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[80]</td>
<td>12</td>
<td>27-31</td>
<td>4</td>
<td></td>
<td>33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[6]</td>
<td>25</td>
<td>8-11</td>
<td>2.2</td>
<td></td>
<td>41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paper D</td>
<td>10</td>
<td>7</td>
<td>1.7</td>
<td>39</td>
<td>63/10</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Paper D</td>
<td>7</td>
<td>7</td>
<td>2.4</td>
<td>41</td>
<td>63/10</td>
<td>19</td>
<td></td>
</tr>
</tbody>
</table>

1The figure within parentheses represents the frequency where OIP3 is measured.  
2For multistage LNAs the power consumption only of the final stage is taken into account when calculating OIP3/Pdc, in order to obtain a more fair comparison to single-stage designs.
Fig. 3.1: Scatterplots of (a) OIP3, (b) noise figure and (c) OIP3/P_{dc} versus frequency of published GaN LNAs and PAs together with a sample of commercially available amplifiers in GaN HEMT, GaAs pHEMT and GaAs/InGaP HBT technologies.

The superior linearity of GaN comes with a higher power consumption, due to the larger drain biasing voltages used in these amplifiers. The ratio of IP3 to DC power consumption can be used to evaluate the inherent linearity of a technology, since both, to a first approximation, scale linearly with gate periphery. This figure of merit is also listed in Table 3.1 and plotted in Fig. 3.1 (c). When it comes to OIP3/P_{dc}, GaN does not seem to have any significant advantage over GaAs. GaAs/InGaP HBTs provide excellent linearity for moderate power consumption up to around 5 GHz and are used for power amplifiers in mobile handsets, but their high noise figure makes them less suitable for low noise amplifiers. There are numerous circuit topologies that can be used to enhance linearity, including negative feedback [13, 81], and various techniques relying on several nonlinear current components from different transistors cancelling at the output [82].

Despite the very good results on LNA level using GaN HEMTs in terms of both linearity and robustness, the fundamental limitations of the technology and their origins on device level are not much studied. In order to identify possible areas of improvement on device and circuit level, it is important to evaluate different
degradation mechanisms and sources of nonlinearity relevant for LNAs. Furthermore, studies on robustness have focused on permanent degradation, but it is also important to consider what happens during the first micro- and milliseconds after a high input power pulse. In sections 3.1-3 the robustness, recovery time and linearity of GaN LNAs are discussed in more detail for specific GaN technologies.

3.1 Robustness of GaN LNAs

This section will describe characterization of LNA robustness on circuit and device level, before moving on to investigations of some of the failure mechanisms that limit robustness in GaN LNAs in section 3.1.1. The robustness of an LNA is here defined as its ability to survive high input power levels without permanent performance degradation. The most straightforward way to characterize robustness is to subject the LNA to increasing input power levels and measure the performance in terms of gain and noise figure between each power level. The stress can be done using either a pulsed or a CW signal. While CW stress is useful as it constitutes a worst case scenario, pulsed signal is a more realistic scenario for LNAs in radar or electronic warfare receivers. Fig. 3.2 shows the gain and noise figure after stress as functions of the input power for four different samples of a 2-6 GHz GaN MMIC LNA fabricated in UMS GH25-10 GaN foundry service. Two of the LNAs were stressed with a CW signal and two with a pulsed signal with 1% duty cycle and 1 µs pulse width. All the tested devices experienced degradation in both gain and noise figure from input power levels higher than 41 dBm, and catastrophic failure occurred for power levels between 42 dBm and 44 dBm. These values are in line with other reported values in Table 3.1. Notably, the LNAs subjected to pulsed stress did not survive higher input power than the CW stressed LNAs, showing that the failure mechanism is fairly fast.

![Fig. 3.2: (a): Noise figure and (b): gain degradation after stress, as a function of input power.](image-url)
In order to understand the limits of GaN technology in terms of LNA robustness and the causes for failure, we need to take a look at the operating conditions of the transistor in an LNA during stress. A series resistor in the kΩ range is commonly used in the gate bias network of LNAs in order to limit the gate current [83], as shown in Fig. 3.3. When the input power gets large enough, the gate diode is turned into forward conduction and a rectified DC current starts to flow through the bias network. The voltage drop over the resistor then causes the transistor’s operating point to fall below the supply voltage, limiting the increase in gate current. As the input power increases further, the DC value of the gate voltage continues to decrease until the device reaches the reverse breakdown in the negative peaks of the waveform. When reverse breakdown is reached, the peak currents in both directions will increase rapidly with input power since the current can no longer be limited by adjusting the DC operating point.

The measurement setup shown in Fig. 3.4 have been used to measure the operating conditions of the transistor during stress under LNA-like conditions and investigate how they are connected to permanent degradation or failure. The use of a large signal network analyser calibrated at the device plane enables measurement of the current and voltage waveforms in the transistor during stress. The gate bias resistor is connected externally, facilitating an LNA-like characterization of the transistor, and allowing its value to be easily varied. This facilitates the study of how variations in technology influence robustness when used in an LNA. Pulsed signals can be used as well as CW signals; the DC components of the drain and gate waveforms for pulsed signals are measured with an oscilloscope. The same setup will also be used to characterize recovery time in section 3.2.1.
Fig. 3.4 Measurement setup for characterizing robustness and recovery time on device level under LNA-like conditions.

This method has been used to characterize robustness in GaN transistors with different bilayer passivation. The robustness characterization was done in a similar way as in when characterizing LNAs, where the input power was increased in 1 dB steps and the gate and measuring the drain and gate I-V characteristics after each step. However, no RF or noise figure characterization was done between the stress sessions in this test, since changes in the I-V characteristics are usually a good qualitative indicator of gain and noise figure degradation.

Four different SiNₓ passivations are compared, all with a total thickness around 50 nm but with different thickness of the two layers. The thicknesses of the lower, low-resistivity layer thickness are 0 nm, 3 nm, 10 nm and 50 nm, respectively, whereas the remaining 50 nm, 47 nm, 40 nm and 0 nm consist of high resistivity SiNₓ. The devices are fabricated similarly to the devices in [24], and have a 100 μm gate periphery. The transistors were subjected to 10 μs pulses of increasing power level at 2 GHz.

Fig. 3.5 shows the waveforms during stress for the four passivations. The waveforms suggest that the transistors with bilayer passivations experience reverse breakdown at lower negative voltages than the ones with monolayer passivations, regardless of the resistivity.
Fig. 3.5: Waveforms during stress for the four different passivations, at different power levels. (a): 50 nm High resistivity, (b): 3 nm low resistivity, 47 nm high resistivity; (c): 10 nm low resistivity, 40 nm high resistivity; (d) 50 nm low resistivity.

$I_d-V_{gs}$ and $I_g-V_{gs}$ sweeps were done after each power level, and the transconductance and gate current at a typical low noise bias point (10 V, 30 mA) are shown as a function of input power in Fig. 3.6. The measurements reveal that the gate current experiences degradation in all devices at input powers between 36 dBm and 40 dBm. Both bilayer passivations degrade 3-4 dB earlier than the monolayer passivations, regardless of the resistivity, and the degradation seems to occur around the same input powers as the first signs of reverse breakdown in the negative peaks of the voltage waveforms (Fig. 3.5). It is thus essential for LNA robustness to design transistors for large gate-drain and gate-source breakdown voltages. There was also a small degradation in transconductance from around 35 dBm in the transistor with highly resistive passivation and (to a lesser extent) in the one with 3 nm low-resistivity layer. Fabrication and measurements of more devices are required to determine whether the differences between monolayer and bilayer passivations are due to inherent properties of the passivations or whether it is due to the specific processing of this particular batch of transistors.
Having studied the operating conditions of transistors in an LNA under input overdrive and their effects on performance, in the next section we will take a look at potential causes of these effects.

3.1.1 Failure mechanisms in GaN LNAs

Several physical mechanisms have been identified as causes of failure or degradation in GaN HEMTs, but these studies have mostly been done with a focus on PA or switching applications. Few have considered the failure mechanisms from the perspective of LNA robustness. It is therefore of value to discuss different known failure and degradation mechanisms and consider whether the failure conditions may occur in an LNA under input overdrive.

Joh et al. [84-87] showed that DC gate-drain or gate-source voltages above a critical value lead to degradation in access resistances, maximum drain current, gate leakage and current collapse. This was attributed to dislocations near the gate edges forming due to inverse piezoelectric stress from high vertical electric fields and adding to the inherent strain due to lattice mismatch in the AlGaN/GaN interface. The critical voltage can be increased by reducing the AlN content in the barrier and thereby the lattice mismatch [85]. Short-gate transistors show lower critical voltages than longer gates, because of increased interaction between the fields on the source and drain side of the gate [88]. These effects could be important for LNAs since the DC value of the gate voltage falls to very low values at high input powers and a dramatic increase in gate current as reported in [87] would cause an increase in noise figure.

Highly energetic electrons colliding with the lattice causing defects to be formed, referred to as hot electron stress, have also been suggested to have an effect on degradation in GaN HEMTs [10]. This effect is strongest when there are simultaneous high current and electric fields in the channel, for example during power amplifier operation. In LNAs, it is unlikely that simultaneous very high...
electric fields and current would occur in the channel in the lateral dimension, but it is conceivable that high vertical fields in the barrier or across the gate Schottky diode during forward or reverse breakdown could cause hot electron effects.

High forward gate current is known to cause a permanent increase in gate leakage [87, 89-91], which has a detrimental effect on noise figure. This was also found in Paper B, where the effects of gate current in mesa isolated GaN HEMTs were investigated. Fig. 3.7 shows the effect on noise parameters, S-parameters and DC characteristics of five minutes with the gate forward biased. A large increase in noise figure and shift in the optimum source reflection coefficient, accompanied by increased gate current is seen after forward gate voltages larger than 3.4 V. The S-parameters show no notable differences after stress and a 0.5 V shift of the threshold voltage is the only difference in the output $I_d-V_g$ characteristics. It is thus clear that the degradation of noise performance is due to the increased gate leakage.

![Fig. 3.7: Measured (a) minimum noise figure and (b) optimum source impedance after different stages of the stress test. The changes in noise parameters were accompanied by an increase in the reverse gate leakage, shown in the inset in (a). (c) and (d) shows the measured s-parameters at a low noise bias point and the $I_d-V_g$ characteristics before and after the stress test, indicating only a small shift in the gate threshold voltage.](image-url)
In [87, 89-91], it is hypothesized that this degradation is caused by damage to the gate Schottky contact as the temperature increases from the power dissipation across the Schottky barrier. In [91], it was also found that using a thermally stable tungsten nitride gate metallization greatly improves the stability of GaN HEMTs when subjected to forward gate currents. As we have seen in section 3.1, high input powers in LNAs cause forward conduction in the positive voltage peaks, and this has been identified as a prominent cause of failure in GaN LNAs [83]. It is therefore meaningful to take a look at heating as a cause of failure in semiconductors and how this can be applied to the case of forward or reverse gate conduction during high input power levels in an LNA.

*Heating as a failure mechanism*

Heating due to high power dissipation in a small area is a common cause of failure in many semiconductors. Modeling the survivability of a device can thus often be reduced to modeling the temperature rise in the most critical parts and determining the critical temperature at which damage occurs [92].

When considering heating as a failure mechanism in an LNA, it is important to note that the temperature rise in the critical parts is determined mainly by the local power dissipation in that area, which may have a rather nonlinear relationship to the total available input power into the LNA. Whereas a power amplifier dissipates most of its power in the channel, an LNA under input overdrive will experience less heat generation there, since the gate voltage is below pinch-off during most of the waveform. However, as the power increases, the forward conduction through the Schottky diode generates heat in the Schottky barrier. When reverse breakdown is reached, heat is generated also in the negative peaks.

Due to the high voltages involved (reverse breakdown in AlGaN/GaN diodes typically happens around -100 V), the power dissipation during reverse breakdown is significantly higher than during forward conduction. On the other hand, during forward conduction the heat generation is concentrated in a very thin region close to the gate metal where most of the voltage drop occurs, whereas the voltage drop at reverse bias is more evenly distributed across the AlGaN barrier layer.

Fig. 3.8 shows the simulated dissipated power in an AlGaN/GaN diode from forward and reverse current as a function of input power, for different values of the series resistance in the bias network. The simulation was performed using the schematic shown in Fig. 3.3, connected to a 50 Ω one-tone power source at the input. The diode was modelled with a forward voltage of 1.5 V, a reverse voltage of -100 V and a 10 Ω series resistance in both directions. Fig. 3.8 shows the trade-off in choosing the bias resistance, since a higher value leads to a larger reverse breakdown current whereas a smaller value gives more forward conduction.
Fig. 3.8: Power dissipation in the gate diode from forward and reverse conduction as a function of input power, for different bias resistances in the model in Fig. 3.1b. The transistor is modeled a diode with a forward voltage of 1.5 V, a reverse breakdown voltage of ~100 V and a series resistance of 10 Ω.

Fig. 3.9: Top view photograph of one of the tested transistors and a schematic cross-section view of the device along one of the gate fingers. The unit gate width is 50 µm.

**Impact of mesa isolation**

Paper B specifically addresses the case when devices are isolated electrically by placing them on mesas, around which the top part of the wafer down to the semi-isolating buffer is etched away. The gate fingers extend outside the mesa in such devices, providing an additional conduction path from the gate directly to the GaN channel at the edges of the mesa, parallel to the AlGaN/GaN Schottky diode (Fig. 3.9). A way to avoid this problem is to isolate devices using ion implantation instead,
which is used in many MMIC processes [93]. Using some kind of passivation to insulate the side walls is another possible solution.

From a degradation perspective, we have to treat the two diodes separately, since excessive current in one of the diodes could cause localized degradation without affecting the other diode. Fig. 3.10a shows the measured forward I-V characteristics up to $V_g = 3.5$ V, for several mesa isolated AlGaN/GaN HEMTs with different gate peripheries but all of them with two fingers. The fact that the forward gate leakage does not scale with the unit gate width and the high series resistance (around 500 Ω) indicate that it is dominated by leakage though the mesa sidewalls rather than through the AlGaN barrier. Although the devices cannot survive forward DC voltages larger than around +4 V, RF measurements reveal a second diode with a significantly lower series resistance, turning on at around $V_g = +6$ V (Fig. 3.10b). Fig. 3.10c shows a DC model of the gate with two parallel diodes, and the I-V characteristics of the model is plotted with the measured characteristics in Fig. 3.10b. It is likely that the diode with the higher threshold voltage is represents the current through the AlGaN barrier.

![Graphs and diagrams explaining the diode behavior and measurements](image)

Fig. 3.10 (a): DC forward I-V characteristics of mesa isolated AlGaN/GaN HEMTs with two gate fingers of different widths. (b): Deembedded RF waveform measurement of the gate current and voltage of a 2x50 μm HEMT under input overdrive, compared to a model with two parallel diodes, shown in (c).
Detailed characterization of the $I_g$ characteristics after increasing levels of stress was performed, in order to find the limits of safe operation and evaluate the effects of forward gate biasing of these devices. DC forward bias as well as RF input signals were used to stress different devices, so that the degradation conditions could be compared between the two cases. Fig. 3.11 shows the gate I-V characteristics after different levels of DC and RF stress. The devices degrade in a very similar manner, indicating that the forward biasing of the gate causes the degradation also in the RF stress case.

By comparing the levels of stress required to damage the devices during DC and RF stress in terms of parameters of the input signal we can find out which parameters define the limits of safe operation. Fig. 3.12 shows the gate leakage at a typical LNA bias ($V_{ds}=10$ V, $V_{gs}=-4$) as a function of peak gate voltage, average gate current and power dissipation. Since there was significant variation between nominally identical devices in the series resistance of the gate diode under forward bias between nominally identical devices, two DC stressed devices are shown in Fig. 3.12, representing a low resistance and a high resistance sample (note that the devices measured in Fig. 3.10a are from another batch with smaller variation but otherwise comparable characteristics). Revisiting the hypothesis that the gate current is dominated by sidewall leakage at low forward bias, the different series resistances could be due to variations in the not-so-well-defined area of the sidewall diode.

When taking the total gate current into account, none of the chosen parameters seems to be able to predict degradation across both DC and RF stress. The voltage predicts the degradation more uniformly for the DC stressed devices which at first glance could indicate an electric field induced degradation mechanism. However, the devices survived far higher forward voltages during the RF signal peaks, which shows that the degradation mechanism is significantly slower than the RF frequency
of 600 MHz. Furthermore, if the different resistances are due to different diode areas, the global current does not necessarily reflect the current densities at the degradation spots and a current or power induced degradation mechanism cannot be excluded since the degradation may be highly localized. Indeed if we calculate the DC current and dissipated current during RF stress only in diode 1 using the model in Fig. 3.12c, the degradation vs power and current (the green line in Fig. 3.12b and c) look much more similar to the DC stress case. In conclusion, based on these result it is not possible to isolate one main failure mechanism, but degradation due to high local power dissipation in the side wall diode seems like the most plausible mechanism.

Fig. 3.12: Comparison of degradation after DC stress and RF stress. The degradation of gate leakage at a gate bias of -4 V is plotted as a function of forward voltage (a), forward gate current (b) and dissipated power in the gate (c). Two otherwise similar DC-stressed devices are shown, one with low series resistance below 5.5 V and one with high resistance. For RF stress, the degradation is shown as a function of total gate current as well as only the $I_1$ component in Fig. 3.10c.
3.2 Recovery time

The study of recovery time in LNAs brings together the topics of trapping effects in GaN, covered in Chapter 2, with the particular operating conditions of highly robust GaN LNAs, discussed previously in this chapter. Many applications, such as RADAR and electronic warfare systems, need receivers that can not only survive high input powers, but also be operational within microseconds after a high input power pulse. In [11], recovery time due to the discharging of the large negative gate voltage through the large resistor in the bias network was analyzed. It was found that the matching network time constants in a typical GaN LNA are in the nanosecond range. Experimentally, recovery was measured after a 38 dBm pulse and a small gain drop was found after the pulse, with a recovery time in the microsecond range, attributed to heating. However, as was seen in Chapter 2, many GaN HEMT technologies have transient effects due to trapping in the buffer and near the surface of the HEMT structure. Considering the very high voltages at the gate during high input power overdrive, it needs to be investigated whether these trapping effects could lead to a transient in the small signal performance after a high power pulse in GaN LNAs.

In paper C, gain transients after a high power pulse were measured for MMIC LNAs fabricated in three different GaN processes. LNA #1 and #3 are designed by Saab Electronic Defense Systems in UMS GH25-10 foundry and a Chalmers in-house MMIC process, respectively. LNA #2 is commercially available from Qorvo. The in-house process differs from commercial processes in that the buffer is not actively doped while the commercial process use Fe doped buffers. Furthermore the Chalmers process utilizes an AlN exclusion layer between the AlGaN barrier and the GaN channel. As we saw in Chapter 2, Fe doping in the buffer of GaN HEMTs is commonly used for increased buffer resistivity and electron confinement in the channel but has an effect on trapping. The wide-bandgap AlN layer increases the electron confinement in the channel and reduces the electron penetration into the barrier but makes it more challenging to obtain good uniformity and low contact resistances.

The measured gain transients for the three amplifiers at different power levels are seen in Fig. 3.13. The two LNAs fabricated in commercial processes suffer a substantial drop in gain after the pulse at power levels as low as 30 dB, far below the damage levels of these circuits. This could constitute a vulnerability for receivers using these LNAs, allowing an intentional or unintentional jammer to temporarily disable essential systems with a pulsed signal at significantly lower power level than is required to permanently damage them. The recovery times are in the range of 20 ms, which could be much too slow since disturbing signals often have pulse repetition frequencies in the kHz range. The drain current transients also saw a
decrease after the pulse and a recovery time similar to that of the gain. The Chalmers LNA, on the other hand, did not experience any decrease in gain after the pulse, and a smaller drain current drop. This indicates that the mechanism causing gain drop in the commercial LNAs is present to a smaller extent also in this technology.

Fig. 3.13: Small signal gain transients in (a) LNA #1, (b) LNA #2 and (c) LNA #3, after a 10 μs pulse to different levels of input power, on a logarithmic time axis with the end of the pulse at t=0.
As mentioned above, the discharging of the gate through the biasing resistance can be one of the factors determining the recovery time of an LNA. Analysing the response time of the matching network using circuit analysis can therefore be useful to confirm or exclude it as a possible explanation to the measured behaviour. All the tested amplifiers have a similar input network topology, depicted in Fig. 3.14, with a series and a shunt inductor $L_1$ and $L_2$ for in-band matching, a DC blocking capacitor $C_b$ at the RF input, an RF decoupling shunt capacitor $C_d$ and a series resistor $R_{bias}$ in the DC biasing network. Typical element values are also listed in Fig. 3.14. The input capacitance to the transistor, $C_{in}$, is also included.

To simulate the recovery of the gate voltage after a high $P_{in}$ pulse due to the matching network, the transistor can be considered as a pulsed current source, representing the forward current through the gate during input overdrive, in parallel with $C_{in}$. Fig. 3.15 shows a transient simulation of the gate voltage during and after a 1 μs long 1 mA current pulse. The RF port was terminated in 50 Ω in the transient simulation and the DC supply voltage was -2 V. Obviously, the recovery due to the matching networks is much faster than the measured transients seen in Fig. 3.13, and can be excluded as an explanation. Trapping in the transistor is thus likely to be the cause.
The large difference between LNA #2 and the other two LNAs indicate that the transistor technology greatly impacts the gain collapse due to input overdrive. More insights about how to minimize these effects can be gained by studying variations in HEMTs on device level.

3.2.1 Device level characterization

The measurement setup in Fig. 3.4 can also be used to characterize recovery time on device level. In these measurements, a vector signal generator is used to send a low input power signal between the pulses like when characterizing LNA recovery in Paper C, while an oscilloscope measures the DC voltages and currents in the device plane during and after the pulse. Like with robustness characterization, device level measurements of recovery time facilitates study on effects of technology and bias resistance variations. Furthermore, nonlinear waveform measurements in the device plane makes it possible to correlate effects on recovery time to device operating conditions during the pulse. This setup has been used to evaluate the effects on recovery time of variations in the AlGaN/GaN interface.

The interface between the AlGaN barrier and the GaN channel is essential in order to obtain good high frequency performance in GaN HEMTs. The extension of the two dimensional gas into the AlGaN layer must be minimized in order to avoid impurity scattering which reduces the electron mobility. This requires a sharp increase of the Al content at the interface, and failure to obtain this due to residual gallium in the MOCVD reactor has been cited as a detrimental factor limiting the high-frequency performance of GaN HEMTs [94].

Growing a few nanometers wide-bandgap AlN exclusion layer at the interface increases the mobility by decreasing the 2DEG extension into the barrier and is therefore often used for high frequency GaN HEMTs. However, difficulties in controlling the thickness and uniformity of the exclusion layer and of forming low-resistivity ohmic contacts through the AlN makes fabrication of AlGaN/AlN/GaN HEMTs more complicated than standard AlGaN/GaN HEMTs [94].

Efforts to obtain a sharp AlGaN/GaN interface without the presence of an exclusion layer by altering the growth conditions of the GaN layers close to the interface [94, 95] have resulted in mobility numbers and high frequency performance in the same range as AlGaN/AlN/GaN devices, while avoiding the problems associated with it. However, the AlN layer could also have an influence on trapping both in the AlGaN barrier, at the surface and the buffer, since electrons in the channel are less likely to tunnel through the higher AlN barrier and occupy trap states at the surface, or in the opposite direction from the gate and occupy trap states in the buffer. It is therefore interesting to evaluate how recovery time in LNAs is affected by a standard AlGaN/GaN interface, an optimized, sharp interface and an AlGaN/AlN/GaN interface. Otherwise identical devices with standard, sharp and
AlN interfaces have been measured using the setup in Fig. 3.4. These devices use an intentionally Carbon doped buffer. The tested devices have a gate width of 2x50 μm. The pulse length was set to 10 μs and the transistors were biased at 10 V and 30 mA current.

![Fig. 3.16: (a) gain and (b) drain current transients after 10 μs pulses of different power levels for a transistor with standard AlGaN/GaN interface.](image1)

![Fig. 3.17: (a) gain and (b) drain current transients after 10 μs pulses of different power levels for a transistor with sharp AlGaN/GaN interface.](image2)
Fig. 3.18: (a) gain and (b) drain current transients after 10 μs pulses of different power levels for a transistor with AlGaN/AlN/GaN interface.

Fig. 3.19: Gate current-voltage waveforms during stress for (a) standard, (b) sharp and (c) AlN interface.

$S_{21}$ and drain current transients after input power pulses between 24 dBm and 30 dBm are shown in Fig. 3.16-18, and the waveforms during stress are shown in Fig. 3.19. Both the standard and the sharp AlGaN/GaN interface lead to around 5 dB decrease in gain and 10 mA decrease in drain current after a 30 dBm pulse, whereas the transistors with an AlN exclusion layer do not experience significant drop. The AlN layer thus seems to be important in order to enable short recovery times. This
could be due to the presence AlN layer growth altering the trap concentration in the barrier or, more likely, the transport of electrons or holes across the barrier. The severe gain collapse for the devices without exclusion layer also indicates that devices with Carbon buffer doping have the same problems with gain collapse as Fe doped devices.

The recovery was also measured during the robustness test of GaN HEMTs with different bilayer passivations described in section 3.1 and the transistors show very little drop in both gain and drain current up to input power levels as high as 40 dBm, despite gate voltage swings down to around -100 V. This is an interesting result since these devices use an unintentionally doped buffer like the Chalmers LNA tested in the previous section, but does not have an AlN exclusion layer. Unintentionally doped buffers thus seems to recover faster than both Carbon and Fe doped buffers.

In summary, the recovery measurements on transistors from different GaN technologies on LNA and device level indicate that both Carbon and Fe doping of the buffer leads to larger gain collapses, while an AlN exclusion layer decreases the gain collapse. Direct comparison of devices with different buffer doping is a suitable topic for future studies. The physical properties and location of the traps leading to these effects and the mechanism by which they cause the gain collapse in LNAs should also be investigated.

3.3 Linearity and intermodulation distortion

For receivers required to linearly detect a low power signal in the presence of another, much stronger signal, low levels of intermodulation distortion (IMD) are required. IMD is the distortion of a signal because of mixing with signals in other frequency bands. Small signal amplifiers such as LNAs can in many cases be considered as weakly nonlinear when it comes to analyzing IMD. In this case, the signal stays in the region around the DC operating point where the bias-dependent, nonlinear elements in the equivalent circuit can be modeled accurately by a third order Taylor expansion around the bias point. The Volterra model is commonly used to analyze IMD in weakly nonlinear systems [96]. It relies on Taylor expansions around the DC operating point of the I-V and Q-V characteristics. If the coefficients are accurately determined, this model predicts device behaviour close to the bias point very well, and can also be used to analyze the importance of different nonlinearities in the device. In principle, Volterra series analysis is not limited to weakly nonlinear systems, but the model fails outside the region where the Taylor polynomial accurately follows the curvature of the nonlinearities. The number of Taylor coefficients necessary thus increases for highly nonlinear components such as mixers or power amplifiers in output power compression making the models less useful.
In a Volterra model capable of predicting third order IMD, the nonlinear elements in the device equivalent circuit are replaced with polynomials up to the third order. The current source $I_{ds}(V_{gs},V_{ds})$ is thus replaced by a two-dimensional polynomial.

$$I_{ds}(V_{gs},V_{ds}) = g_m V_{gs} + g_{ds} V_{ds} + g_{m2} V_{gs}^2$$
$$+ g_{md} V_{gs} V_{ds} + g_{d2} V_{ds}^2 + g_{m3} V_{gs}^3$$
$$+ g_{m2d} V_{gs}^2 V_{ds} + g_{md2} V_{gs} V_{ds}^2$$
$$+ g_{d3} V_{ds}^3$$

(3.1)

And the nonlinear gate charge and drain charge functions are approximated by:

$$Q_{gs}(V_{gs}) = C_{gs} V_{gs} + C_{gs2} V_{gs}^2 + C_{gs3} V_{gs}^3$$

(3.2)

$$Q_{gd}(V_{gd}) = C_{gd} V_{gd} + C_{gd2} V_{gd}^2 + C_{gd3} V_{gd}^3$$

(3.3)

This approach makes it possible to derive analytical expressions for the IMD components and the OIP3 when a circuit is excited by a two-tone signal, based on the Taylor coefficients of the active device and the impedance matrix of the linear network embedding it. The contribution of different parameters can be analyzed separately.

However, a complete Volterra series analysis, taking into account all nonlinear effects of a transistor based amplifier tends to produce very large expressions which are very difficult to gain any understanding from. There is thus a trade-off between accuracy and completeness on one side, and simplicity on the other side in choosing which effects and parameters to include when doing an analysis of nonlinear distortion. Moreover, some of the Taylor coefficients are very difficult to determine from only S-parameter and DC measurements. When working with wide band gap technologies, this is even more difficult as these transistors often suffer from low frequency dispersion due to self-heating and trapping. As the IMD products in an amplifier are vector sums of dozens of partially cancelling terms, the IMD level is very sensitive to small errors in the Taylor coefficients and other parameters.

Nonetheless, analysis of the amplifier linearity performance based on a simplified model with the nonlinear elements represented by polynomials can be useful despite not predicting the measured IP3 accurately. If the model is simple enough, the analytical expressions for the linearity figure of merits are fairly simple and understandable and can be used to provide insight into the sources of nonlinearity in the amplifier and the different trade-offs involved in the design.
3.3.1 OIP3 analysis on device level – GaN vs GaAs

To analyze IMD in an amplifier, not only the nonlinear elements but also the linear parts, such as parasitics and matching networks, must be taken into consideration. In order to analyze the impact of the transistor characteristics and how different contributions affect IP3, a very simple amplifier model is chosen, consisting of a nonlinear current source $I_{ds}(V_{gs}, V_{ds})$ and a resistive output termination (Fig. 3.20). The same analysis will also be performed on a GaAs pHEMT to explore whether there are differences in device characteristics which affect linearity. The analysis will later be extended to a more general case with a general passive network embedding the active device and a nonlinear gate capacitance, to evaluate how these factors affect linearity.

The influence of second order nonlinearities will not be included in the analysis. Second order nonlinearities produce components at baseband and at the second harmonic. When these components again mix with a fundamental component in a second order nonlinearity, they create third order products, some of which are in-band. These components produce a large number of terms with different dependencies on bias and on the matching networks at the fundamental frequency as well as at the second harmonic or at baseband. As they in most cases have limited influence on OIP3, the analysis is simplified by omitting these products. Short circuiting the device nonlinearities at the second harmonic and baseband frequencies is one way to minimize the influence of second order products.

The circuit model in the simplified IMD analysis is shown in Fig. 3.20. The nonlinear current includes the linear transconductance and the third order terms.

\[
I_{ds}(V_{gs}, V_{ds}) = g_m V_{gs} + g_m^3 V_{gs}^3 + g_{m2d} V_{gs}^2 V_{ds} + g_m V_{ds}^2 + g_{d3d} V_{ds}^3
\]  

The linear output conductance is considered as part of the output termination. Using the nonlinear currents method [96], OIP3 can be calculated by successively calculating the first order voltages, the third order currents and finally the third order voltage at the output. OIP3 becomes:
\[ OIP_3 = \frac{2R_L}{3 \left[ g_{m3}/g_m^3 - g_{m2d} R_L / g_m^2 + g_{md2} R_L^2 / g_m - g_{d3} R_L^3 \right]} \]  

Depending on the device Taylor coefficients, the terms in the denominator may add either in phase or out of phase. Also, they depend in different ways on \( g_m \) and on the output termination, so the optimum termination may vary strongly with device characteristics and with bias. For some sets of Taylor coefficients, there is a load termination which causes the different terms to cancel out, resulting in IMD sweet-spots.

The \( g_m \) and \( g_{ds} \) characteristic extracted from S-parameter measurements of a 4x25 \( \mu \)m transistor from the Triquint 3MI 0.25 \( \mu \)m GaN on SiC MMIC process are shown in Fig. 3.21. The values for a GaAs transistor are shown for comparison. The Taylor coefficients in (3.5) are extracted from the derivatives of \( g_m \) and \( g_{ds} \) with respect to \( V_{ds} \) and \( V_{gs} \). In order to assess the relative importance of the different terms in the denominator of they are calculated separately assuming a load resistance of 30 \( \Omega \cdot \)mm. Fig. 3.22 and Fig. 3.23 show the contribution from each term at different bias, as well as the sum of all contributions, for GaN and GaAs respectively.

For both technologies, the terms depending on \( g_{m3} \), \( g_{m2d} \), and \( g_{md2} \) are all in the same order of magnitude and decrease with increasing current. The \( g_{m3} \) contribution depends almost only on the current while \( g_{m2d} \) and \( g_{md2} \) contributions increase with higher load resistance and lower drain voltage. There is partial cancellation between these terms. Consequently, the total IMD is smaller than each of the individual terms for current levels below 20 \% of \( I_{max} \). At high currents, the \( g_{d3} \) contribution becomes dominant and sets the OIP3. This term decreases for high drain voltages and small load resistances, which is therefore favourable for OIP3 at high current levels.

Fig. 3.21 Transconductance and output conductance of a GaN and a GaAs transistor, extracted from measured S-parameters. Different traces correspond to different drain voltages. For GaN, the drain voltages are 5 V, 10 V and 15 V, and for GaAs 1.5 V, 2 V, and 3 V.
\[ \Sigma = \frac{g_m^3}{g_m^3} - \frac{g_m^2 R_L}{g_m^2} + g_m d R_L^2 / g_m - g_d R_L^3 \]

Fig. 3.22 Contributions to GaN amplifier intermodulation distortion from different terms in the drain current Taylor expansion at (a) 5 V, (b) 10 V and (c) 15 V bias. The blue line is the sum of all contributions in the denominator of (3.6).

\[ \Sigma = \frac{g_m^3}{g_m^3} - \frac{g_m^2 R_L}{g_m^2} + g_m d R_L^2 / g_m - g_d R_L^3 \]

Fig. 3.23 Contributions to GaAs amplifier intermodulation distortion from different terms in the drain current Taylor expansion at (a) 1.5 V, (b) 2 V and (c) 3 V bias. The blue line is the sum of all contributions in the denominator of (3.6).

Fig. 3.24 Simulated OIP\textsubscript{3} of (a) GaN and (b) GaAs amplifier, calculated using (3.6), a load resistance of 30 \(\Omega\) mm and Taylor coefficients extracted from bias dependent S-parameters.

In the GaAs device, the nonlinearity of the output conductance is much more important, resulting in much smaller load resistances for the optimum OIP\textsubscript{3}. In general, however, the characteristics and the roles of the different IMD terms are similar in the two technologies. GaN transistors, with their higher breakdown
voltages, offer more possibilities to improve linearity by increasing the drain bias, thereby suppressing the nonlinearity of transistor output characteristics. For the same drain voltage, however, GaAs seems to offer slightly better linearity than GaN (Fig. 3.24), mainly because of higher $g_m$. The linearity advantage of GaN thus comes at the cost of higher power consumption.

The IMD can be expected to be very sensitive to variations in the Taylor coefficients and accurate simulation of OIP$_3$ requires very accurate characterization of the device. This is true particularly in the most linear regions of the bias and impedance space, where cancellation between different contributions to the total IMD current is strong.

### 3.3.2 OIP3 analysis on circuit level

For an extended analysis, a very general representation is chosen, where the linear network embedding the intrinsic device is represented by a four port and its $Z$ matrix (Fig. 3.25). Device parasitics, matching networks as well as the input and output terminations (usually 50 Ω) are considered as part of this four port. The amplifier is assumed to be narrowband, so that the $Z$-parameters are the same at both tones in a two tone test. By representing the passive network by a general four port, fairly simple expressions for the OIP$_3$ can be obtained without making unrealistic assumptions about device and amplifier characteristics or assuming a certain topology. This sheds light on how the OIP3 can be impacted by passive network design in an amplifiers. The extended analysis also serves to check to which extent the general conclusions from the simplified analysis hold true also in a more general case. On the other hand, the influence of particular elements in the passive networks cannot be determined without calculating how they influence the $Z$-parameters. A nonlinear gate capacitance is also included in the analysis, where the linear part is included in the passive network and the nonlinear part is represented by a nonlinear charge connected to port 3 (Fig. 3.25).

\[
Q_{gsn} = C_{gs3} V_{gs}^3 \tag{3.6}
\]

![Fig. 3.25: Nonlinear model of a general one stage amplifier.](image-url)
In this case, OIP3 becomes:

\[
OIP_3 = \frac{2 \left| g_m - \frac{Z_{21}(1 + Z_{34}g_m)}{Z_{24}Z_{31}} \right|^3 |1 + Z_{34}g_m| |Z_{24}|^2}{3|Y_3|R_L}
\]  
(3.7)

Where \( R_L \) is the impedance of the output termination, usually 50 \( \Omega \), and \( Y_3 \) is given by:

\[
Y_3 = g_{m3} + \frac{1}{3} g_{m2d} \left[ 2 \left( \frac{Z_{41}(1 + Z_{34}g_m)}{Z_{31}} - g_mZ_{44} \right) \right.
\]

\[
+ \left( \frac{Z_{41}(1 + Z_{34}g_m)}{Z_{31}} - g_mZ_{44} \right)^2
\]

\[
+ \frac{1}{3} g_{md2} \left( \frac{Z_{41}(1 + Z_{34}g_m)}{Z_{31}} - g_mZ_{44} \right)^2
\]

\[
+ 2 \left[ \frac{Z_{41}(1 + Z_{34}g_m)}{Z_{31}} - g_mZ_{44} \right]^2
\]

\[
+ g_{d3} \left( \frac{Z_{41}(1 + Z_{34}g_m)}{Z_{31}} - g_mZ_{44} \right)
\]

\[
\cdot \left[ \frac{Z_{41}(1 + Z_{34}g_m)}{Z_{31}} - g_mZ_{44} \right]^2
\]

\[-j\omega C_{gs3} \left( g_mZ_{33} + \frac{Z_{23}(1 + Z_{34}g_m)}{Z_{24}} \right) \]
(3.8)

In common source amplifiers with significant gain, the drain voltage component which is amplified by the active device is much larger than the component transmitted passively from the input, so we can assume that:

\[
|g_mZ_{44}| \gg \left| \frac{Z_{41}(1 + Z_{34}g_m)}{Z_{31}} \right|
\]  
(3.9)

\[
|g_mZ_{33}| \gg \left| \frac{Z_{23}(1 + Z_{34}g_m)}{Z_{24}} \right|
\]  
(3.10)
\(|g_m| \gg \left| \frac{Z_{21}(1 + Z_{34}g_m)}{Z_{24}Z_{31}} \right| \quad (3.11)\)

These assumptions simplify the expressions for \(\text{OIP}_3\) and \(Y_3\), giving:

\[
\text{OIP}_3 = \frac{2}{3} \cdot \left| \frac{g_m^3}{Y_3} \right| \cdot \frac{|Z_{24}|^2}{R_L} \cdot |1 + Z_{34}g_m| \quad (3.12)
\]

\[
Y_3 = g_{m3} - \frac{1}{3} g_{m2d} g_m (2Z_{44} + Z_{44}^*) + \frac{1}{3} g_{md2} g_m^2 (Z_{44}^2 + 2|Z_{44}|^2)
+ \frac{1}{3} g_{md2} g_m^2 |Z_{44}|^2 - j\omega C_{gs3} g_m Z_{33} \quad (3.13)
\]

If \(Z_{41}=Z_{23}=Z_{34}=Z_{21}=0\), \(Z_{44}=Z_{24}=R_L\) and \(C_{gs3}=0\), this is identical to (3.6). If the
device is matched to a resonant load at the output, the interplay between the different
terms in \(Y_3\) is thus identical to the simplified case described in the previous section.
The nonlinear capacitance produces a term 90° out of phase with the drain current
nonlinearity, giving rise to AM-PM distortion. One way to mitigate this effect is to
tune the output load, so that \(Z_{44}\) is no longer purely resistive. This load can then be
tuned so that the imaginary part of the terms originating from the drain current
nonlinearity cancels the capacitance nonlinearity. This approach, however, is very
sensitive to model inaccuracy. Another way to minimize the influence of the
nonlinear gate capacitance is to match to low impedance on the input, so that \(Z_{33}\)
becomes small. This may, however, lead to worse gain and noise figure.

One conclusion that can be drawn from this analysis is that a circuit designer can
in theory, regardless of technology, always improve linearity by increasing the
feedback, represented by the factor \(|1 + Z_{34}g_m|\). For maximizing this factor, \(Z_{34}\) should
be large, real and positive. This corresponds to negative feedback, since the drain
current going out of port 4 in Fig. 3.25 gives a negative voltage at port 3, producing
a drain current in the opposite direction. In a common source amplifier, the feedback
factor can be increased by adding networks to the source or between the input and
the output. Resistive feedback networks are broadband but introduce noise and losses
while reactive networks can result in negative feedback in a limited frequency band.
3.3.3 Design of highly linear GaN MMIC LNAs

In paper D, two GaN MMIC LNAs in Triquint’s 3MI GaN on SIC MMIC process were designed and characterized. Fig. 3.26 shows microphotographs of the two amplifiers. The focus was on high linearity with reasonably low power consumption and noise figure. This was achieved using reactive negative feedback. In LNA 1, inductive source feedback was used, as is common in LNAs. The feedback as well as the matching networks were designed as a compromise between return loss, gain, linearity and noise figure. In LNA 2, additional feedback is introduced by adding extra gate-drain capacitance externally. Fig. 3.27 shows the feedback factor $|1 + Z_{34}g_{m}|$ for the two circuits.

Fig. 3.26: Microphotographs of LNA 1 (left) and LNA 2 (right)

Fig. 3.28 shows the measured OIP3 and linearity figure of merit (OIP3/PDC) at 7 GHz for the two circuits. The optimum figure of merit is achieved at 10 V and 63 mA bias for both amplifiers, where LNA achieves a value of 12 and LNA a value of 19. As was seen in Fig. 3.1, this is higher than previously published GaN amplifiers as well as typical commercial amplifiers in GaAs HEMT and HBT technology at similar frequencies.
Chapter summary

This chapter covers several aspects of GaN LNAs, namely robustness, temporary gain degradation after high input power operation and linearity. Published GaN LNAs where robustness and/or linearity values are reported were reviewed. GaN LNAs typically show at least an order of magnitude higher values of maximum safe input power as well as output IP3, compared to GaAs pHEMT LNAs. Measurements of GaN LNA robustness on circuit level as well as on device level were presented, in both cases showing degradation for input power levels around 40 dBm. In order to better understand the limits of GaN technology for robust LNAs, degradation mechanisms and their impact on LNAs were discussed. The special case where the HEMTs are placed on a mesa for device isolation was studied, and it was found that excessive forward current through the mesa sidewalls could be a limiting factor for
the survivability of these devices, which degrade after being subjected to forward gate currents as low as 3-6 mA.

Temporary degradation after input power pulses below the damage levels were evaluated on LNA and transistor level. It was found that some technologies suffer from severe gain degradation after pulses lower than 10 dB below the damage levels, whereas other technologies show little or low gain drop. The recovery times of these effects are in the millisecond range, suggesting that this is due to trapping effects in the transistors. The design of the buffer and the AlGaN/GaN barrier seems to be important in order to reduce these effects.

Sources of nonlinearity in a GaN HEMT were also analyzed in detail and compared to a GaAs pHEMT, in order to examine the potential of GaN LNAs in terms of OIP3. The analysis showed that the superior linearity of GaN is due to being biased at higher drain voltages which reduces the output conductance nonlinearity. The advantage of GaN HEMTs thus comes with the cost of higher DC power consumption. Two 6-8 GHz GaN MMIC LNAs were designed with the goal of maximizing the OIP3 to $P_{dc}$ ratio, combined with low noise figure. This was achieved by increasing the negative feedback from the output to the input using inductive source feedback as well as parallel capacitive feedback. The resulting OIP3/$P_{dc}$ values of 12 and 19 are the highest reported for GaN amplifiers in this frequency band.
Chapter 4
Noise Parameter measurements

Successful LNA design requires a device model capable of predicting the noise figure of a transistor when connected to different embedding networks. In order to describe a general linear noisy two port and calculate the noise currents and voltages at its terminals, a set of four noise parameters, along with its linear network parameters, is needed. The parameters used in the standard IEEE representation [97] are the minimum noise temperature $T_{\text{min}}$, the noise resistance $R_n$, and the complex optimum source admittance $Y_{\text{opt}}$. With this representation, the noise figure of a component is:

$$T_n = T_{\text{min}} + T_0 \frac{R_n}{G_s} |Y_s - Y_{\text{opt}}|^2$$  (4.1)

where $Y_s$ is the admittance of the source termination and $G_s$ its real part, and $T_0$ is defined as 290 K. When measuring noise parameters of a device, the noise temperature is measured for a number of different $Y_s$ and the noise parameters are determined by fitting against (4.1). However, to calculate the noise temperature of the device under test, the equivalent noise temperature of the input termination must be known. In commercial systems, this is often assumed to be equal to the ambient room temperature [98].

In this chapter, the principles behind noise parameter measurements will be described, in order to provide background to Paper E, where a characterization of the noise temperature of an electronic impedance tuner used for noise parameter measurements is presented.

4.1 Noise parameter measurements

The noise temperature of a device for a certain input termination can be determined from a single measurement of output noise power $P_{\text{out}}$, if the equivalent noise temperature $T_{\text{in}}$ of the input termination, the S-parameters of the device, the
reflection coefficients looking into the source termination $\Gamma_s$ and into the measurement receiver $\Gamma_R$, are known. The device noise temperature is then:

$$T_n = \frac{P_{out}}{k G_{DUT} \Delta f} - T_{in}$$  \hspace{1cm} (4.2)

where $G_{DUT}$ is the transducer power gain of the device under test, calculated from the device S-parameters, $\Gamma_s$ and $\Gamma_R$. Measurement of the output noise power requires calibration of the measurement receiver, which involves determining the operating power gain as well as the noise temperature of the receiver, for the impedance presented to the receiver looking into the device. If an isolator is present on the receiver input, the noise parameters and gain of the receiver can be determined from a single Y-factor measurement [99], but the measurement will then be limited by the bandwidth of the isolator. Another way to determine the receiver noise parameters and gain is to perform a Y-factor measurement with a 50 $\Omega$ noise source and direct noise measurements of several different room temperature impedances [100].

If the noise figure can be measured in several terminations, the noise parameters may be determined without any assumptions about the device under test. In commercial noise parameter test sets, a tuner is used at the input to present different source terminations at room temperature, and a VNA is used to measure S-parameters and noise power in the same setup [98]. The measured noise figures are then fitted against (4.1) to determine the noise parameters.

4.1.1 Noise temperature of an electronic source impedance tuner

Measurement of noise figure from direct noise power measurement requires accurate knowledge of the equivalent noise temperature of the input termination, $T_{in}$ in (4.2). In passive, mechanical tuners, which dissipate little DC power, this is very close to room temperature. If electronic tuners are used, however, there may be excess noise, as the current through the p-i-n diodes produces noise. Furthermore, the diodes dissipate DC power, so the physical temperature inside the tuner is higher than the ambient DC power. The most accurate, modern noise parameter systems therefore use mechanical tuners [101]. However, because of their speed and ease-of-use, electronic tuners, for example using p-i-n diodes, are used in many older commercial systems that are still in use [98]. It is therefore of interest to quantify the excess noise produced by these tuners.
In paper C the equivalent noise temperature at 10 GHz for all impedance states of an electronic tuner used in the commercial NP5 noise parameter system is characterized. As seen in Fig. 4.2, the noise temperature is significantly higher than room temperature and varies between different impedance states. The general trend is towards higher noise temperatures for more reflective loads, but for the outermost points (marked in red in Fig. 4.1 and with rings in Fig. 4.2), the temperature is instead similar to the low Γ points. If not accounted for when doing measurements, the excess noise in the tuner leads to large errors in measured noise parameters.

Fig. 4.1 The 353 impedance states of the tuner at 10 GHz in the Smith Chart.

Fig. 4.2 Measured noise temperature of electronic tuner for all 353 impedance states. The different colors correspond to different measurement series of the same states. The points marked with rings correspond to the red points in Fig. 4.1.
Chapter 5
Conclusions and future outlook

The work presented in these thesis serves to improve understanding of the properties of GaN HEMTs and the consequences of these properties on circuit and system performance for both PAs and LNAs. It also gives several examples of how these effects can be evaluated using device level measurements.

It is shown that the Fe doping profile in the tested ranges impacts the recovery time of trapping effects, while traditional characterization procedures show no or insignificant differences between the tested materials. Firstly, this means that the buffer profile can to some extent be optimized only for recovery time. Secondly, non-standard device characterization methods such as transient recovery measurements are needed to evaluate the effects of the buffer on device performance. It is also shown that DC-pulsing the drain voltage gives similar effects on the drain current transient as RF power pulsing, and that the drain current transient is in turn correlated to the RF gain transient. DC-pulsing can thus be used as a tool to qualitatively characterize memory effects of the RF gain due to trapping effects.

The work on robustness of GaN LNAs shows that forward current through the mesa sidewalls in mesa-isolated HEMTs could be a limiting factor. Device isolation using ion implantation might thus be a better solution and a direct comparison would be an interesting future study. In order to evaluate transistors from an LNA robustness perspective taking all potential degradation mechanisms in account, large signal measurements of the gate I-V waveforms during input power stress with an externally connected resistor is a good method. The effects of the stress on the device can then be directly measured and degradation can be correlated to the operation conditions during stress. By testing with different resistor values, the limits of the device during forward as well as reverse bias stress can be evaluated.

Other important parameters to consider are pulse length and duty cycle of the stress signal, since different degradation mechanisms work on different time scales. The LNA robustness measurements in section 3.1 indicate that failure mechanisms in GaN HEMTs can be fairly fast, no significant increase in survivability was seen for a 1% duty cycle and 1 µs pulse length signal, compared to CW input power.
Testing survivability for different pulsing conditions, especially pulse lengths down in the sub-microsecond range, is an important priority for future research on GaN LNA robustness.

The work on recovery time in GaN LNAs shows that temporary gain degradation after a high input power pulse is an important issue that could be a showstopper for GaN LNAs in electronic warfare applications, since the possibility to temporarily disable a receiver with power levels far smaller than the damage rating constitutes a serious vulnerability. However, the problem seems to be present only in some GaN technologies, indicating that the problem can be solved. Although it seems very likely that trapping causes the degradation, the exact trapping mechanisms and the location and physical properties of the traps causing the degradation remains to be investigated. Further studies on device level to optimize HEMTs towards short recovery times is another important priority. Solutions need to be compatible with good performance for LNAs as well as PAs and switches, since one of the largest advantages with GaN LNAs is the possibility to integrate it on the same chip as these other functions.

GaN HEMT LNAs generally offers better linearity than other LNA technologies. This is due to being biased at a higher drain voltage, which reduces the nonlinearity of the output conductance. This means that the benefit of higher linearity comes at the cost of higher power consumption. The methods to improve linearity without increasing power consumption are the same in GaN as in other technologies: negative feedback and cancellation techniques.

Two MMIC GaN LNAs were designed, achieving higher OIP3/P_{dc} figures than all previously published GaN LNAs and PAs in the same frequency range, with values of 12 and 19. This was achieved by increasing the negative feedback from the output to the input using series inductance on the source for the LNA with the lower OIP3/P_{dc} value and both source inductance and parallel drain-gate capacitive feedback for the circuit with the higher value.
Acknowledgment

I would like to express my gratitude to the people who made this work possible. Thanks to my supervisor Niklas Rorsman for giving me the technical feedback, guidance and encouragement necessary for me to complete this work. Also thanks to co-supervisors Mattias Thorsell and Kristoffer Andersson for all the valuable guidance and rewarding discussions during my time as a PhD student.

I owe many other colleagues my thanks for their assistance and collaboration. Special thanks to Tongde Huang and Johan Bergsten for providing devices and helping me to characterize them, and for Sebastian Gustafsson for all the assistance in the lab.

My project partners from industry and academia have given me useful feedback and a broader perspective on my research. Special thanks to Niklas Billström and Johan Carlert at Saab, Thomas Emanuelsson and Tobias Wegeland at Ericsson and Hervé Blanck at UMS for their feedback from an industrial point of view.

To all the colleagues and friends during my time at MC2, thank you for making it fun to go to work and for being great friends inside and outside the work place.

My deepest gratitude to my family, whose loving support is a constant source of comfort and helps me put struggles and failures at work into perspective.

Most of all, thanks to my love 革丽 for all the happiness you give me and for making my work with this thesis feel more meaningful. I dedicate this thesis to you.

This work has been performed as part of several projects:
"HPM-skydd av AESA baserat på GaN" financed by the Swedish Governmental Agency of Innovation Systems (VINNOVA) and Saab AB within the National Aviation Engineering Research Programme
"Reconfigurable Linear Amplifier –RECLAM” project financed by VINNOVA and Ericsson
"Advanced III-Nitrides-based electronics for future microwave communication and sensing systems” financed by the Swedish Foundation for Strategic Research (SSF).
“Swedish MANGA” financed by the Swedish Defence Materiel Administration
“GHz center” supported by VINNOVA, Ericsson AB, SAAB AB, UMS, Mitsubishi, Infineon AG, and Comheat AB
References


References


References


Paper A

Application Relevant Evaluation of Trapping Effects in AlGaN/GaN HEMTs With Fe-Doped Buffer


Application Relevant Evaluation of Trapping Effects in AlGaN/GaN HEMTs With Fe-Doped Buffer

Olle Axelsson, Sebastian Gustafsson, Student Member, IEEE, Hans Hjelmgren, Senior Member, IEEE, Niklas Rorsman, Member, IEEE, Hervé Blanck, Jörg Splettstoesser, Jim Thorpe, Thomas Roedle, and Mattias Thorsell, Member, IEEE

Abstract—This paper investigates the impact of different iron (Fe) buffer doping profiles on trapping effects in microwave AlGaN/gallium nitride (GaN) high electron mobility transistors (HEMTs). We characterize not only the current collapse due to trapping in the buffer, but also the recovery process, which is important in the analysis of suitable linearization schemes for amplitude modulated signals. It is shown that the simple pulsed dc measurements of current transients can be used to investigate transient effects in the RF power. Specifically, it is revealed that the design of the Fe-doping profile in the buffer greatly influences the recovery time, with the samples with lower Fe concentration showing slower recovery. In contrast, traditional indicators, such as S-parameters and dc as well as pulsed I–V characteristics, show very small differences. An analysis of the recovery shows that this effect is due to the presence of two different detrapping processes with the same activation energy (0.6 eV) but different time constants. For highly doped buffers, the faster process dominates, whereas the slower process is enhanced for less doped buffers.

Index Terms—Dispersion, gallium nitride (GaN), high electron mobility transistors (HEMTs), semiconductor device doping, trap levels.

I. INTRODUCTION

T

HE HIGH output power density and efficiency offered by power amplifiers (PAs) based on gallium nitride (GaN) high electron mobility transistors (HEMTs) make them a strong candidate for high-frequency power applications. However, the dispersive behavior of these transistors inhibits both Continuous Wave performance and linearity, and has delayed their adoption in many microwave systems. The dispersive effects are mainly caused by deep electron traps in the buffer and at the surface. These dispersive effects are manifested as a decrease in maximum current (current slump) and ON-state conductance (knee walkout) during large signal operation, leading to lower maximum output power and efficiency [1], [2]. Also of great concern from a system perspective but less studied on device level is the distortion of amplitude modulated signals due to these memory effects. Linearization schemes using digital predistortion are impractical for long time constants, which require a large computational capacity and memory [3], [4].

Research on traps in GaN has primarily been focused on understanding their physical mechanisms and minimizing the current collapse. Earlier, much efforts were dedicated to minimize the effects of surface traps by utilizing different passivation schemes [1], [2], [5]. More recently, there has been a growing interest in the influence of the characteristics of the GaN-buffer on the large-signal performance of GaN HEMTs. The GaN-buffer is commonly doped with a deep acceptor such as iron (Fe) to reduce leakage currents and increase breakdown voltage [6], [7]. However, the design of the buffer, in terms of compensation doping concentration and profile is known to have an effect on trapping phenomena. In particular, a trap with activation energy \( \sim 0.6 \text{ eV} \) is thought to be localized in the buffer, and it has been seen that its concentration increases with the Fe-dopant concentration [8]–[11]. In [11], it was predicted from simulations that Fe traps would cause a small current collapse that is almost independent of the Fe concentration above a certain threshold level. However, it has also been suggested that the 0.6 eV traps seen in measurements on the Fe-doped buffers are not due to the Fe atoms but another trap indirectly affected by the doping [8].

This paper presents an application-focused investigation of the effect of different buffer Fe-doping profiles on GaN HEMT PA performance, focusing in particular on transient effects due to trapping after high-power operation. This is important in the analysis of suitable linearization schemes but often not considered in other studies. It is shown that the simple pulsed dc measurements of current transients can be used to investigate transient effects in the RF power. While standard device characterization methods (e.g., dc and pulsed I–V, S-parameters, and so on) reveal only small differences between

samples with different buffer Fe profiles, we show that the recovery process after a high-power pulse is significantly affected by the Fe doping. Furthermore, these effects are analyzed by determining the time constants and activation energies of the traps involved.

This paper is organized as follows. In Section II, the devices under test (DUT) and the differences in buffer doping are explained. Section III compares the dc and small signal performance of the devices. Section IV describes the measurements used to characterize trapping effects in the devices and presents the results of these measurements. Finally, the conclusions are drawn in Section V.

II. TECHNOLOGY

The DUT is 0.5-μm gate length GaN HEMTs with 6-μm × 400-μm gate width, fabricated by United Monolithic Semiconductors using Cree GaN wafers grown on SiC. The gate-source and gate-drain distances are 1.5 and 3 μm, respectively. A gate-connected field plate extends 0.5 μm toward the drain side and a source-connected field plate 2.3 μm from the source and above the gate. The lower part of the GaN buffer is Fe doped in order to obtain high resistivity and increase the breakdown voltage. Since Fe-doping in the channel has been shown to cause severe current collapse [7], the Fe-doping profiles are designed to have an exponential decrease with a distance toward the AlGaN/GaN interface. This is inherently achieved during the metal–organic chemical vapor deposition growth by memory effects after the closing of the Fe source. Fig. 1(a) shows the three tested buffer Fe profiles. The total buffer thickness is 1.8 μm for all devices. The standard doping profile has an approximately 1-μm-thick plateau with an Fe concentration in the 10¹⁸ cm⁻³ range, decreasing exponentially to around 10¹⁶ close to the AlGaN/GaN interface. Besides the standard buffer design, the Fe-doping close to the channel was varied by reducing the thickness of the Fe plateau by 20% while keeping the total buffer thickness the same (reduced thickness) or by reducing the Fe concentration in the plateau by 50% (reduced concentration). All measurements have been carried out on two devices of each type to assure the reproducibility of observed variations. However, only the data for one device of each type are presented, unless there are specific reasons to comment on differences between nominally identical devices. The transistors were mounted in a package, and bond wires were used to connect the gate, drain, and source terminals to the package [Fig. 1(b)].
TABLE II

<table>
<thead>
<tr>
<th>Buffer</th>
<th>f_{max} [GHz]</th>
<th>f_c [GHz]</th>
<th>g_m [mS]</th>
<th>g_R [mS]</th>
<th>C_{out} [pF]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>28</td>
<td>8.2</td>
<td>340</td>
<td>3.2</td>
<td>2.3</td>
</tr>
<tr>
<td>Reduced thickness</td>
<td>28</td>
<td>8.9</td>
<td>344</td>
<td>3.8</td>
<td>2.3</td>
</tr>
<tr>
<td>Reduced concentration</td>
<td>28</td>
<td>8.9</td>
<td>352</td>
<td>3.8</td>
<td>2.3</td>
</tr>
</tbody>
</table>

defined as the difference in gate threshold voltage in order to obtain 1-mA current between 0.5 and 50 V drain bias, and SS is the slope of log_{10}(I_{ds}) with respect to the gate voltage in the subthreshold region, measured in decade/V. As shown in Table I, DIBL is slightly higher in the devices with decreased Fe doping. Although small, the differences are consistent across samples and expected, since a higher doping level reduces the short-channel effects by improving the electron confinement in the channel [13]. The differences in SS were small and inconsistent across different samples. The OFF-state drain current is ∼0.1 mA and dominated by gate leakage for all devices at V_{ds} = 50 V.

Furthermore, two-tone measurements have been performed on conjugate matched HEMTs at 2.65 GHz at several bias points and tone spacings and showed no significant variations in the third-order intercept point (IP3) between the three device types up to output powers of 29 dBm per tone. The HEMTs achieved an output IP3 of 43 dBm at 28 V and 200 mA bias.

In conclusion, the different Fe-doping profiles did not demonstrate any significant effects on dc, small signal performance and linearity at low power levels and none of the tested devices showed leakage currents or breakdown voltages that are problematic for PA operation. This implies that the Fe buffer can be optimized to a large degree for other characteristics, such as reduced dispersive effects.

IV. CHARACTERIZATION OF TRAPPING EFFECTS

To characterize the different consequences of dispersive effects caused by electron trapping in the buffer, several characterization methods are needed. The traditional method of characterizing trapping effects is pulsed I–V measurements, which are carried out in Section IV-A. This gives information on the current collapse and knee-walkout that affect the output power and efficiency. However, the pulsed I–V characteristics do not reveal transient effects on the device RF performance after a change in operating conditions, such as moving from a high RF power level to a lower power level.

Direct characterization of the transient behavior is performed in Section IV-B. We use a vector modulator to supply a high-power RF input signal to a matched HEMT at 2.65 GHz for 300 μs before backing off the input power level by 15 dB. The output power is monitored in the time domain during and after the pulse using a vector signal analyzer.

However, this method requires expensive instrumentation and is impractical on device level, since it requires engineering of the output current and voltage waveforms (e.g., using tuners, matching networks, or active injection) in order to create realistic operating conditions for the device. Therefore, a more practical method to obtain information of these effects on device level is desired. In Section IV-C, we propose that the effects of high-power microwave operation can be mimicked by dc-pulsing to a high drain voltage, and the transient effects can be monitored by measuring the drain current after the pulse at an application relevant bias.

A. Pulsed I–V Characterization

To obtain information about current collapse due to gate and drain lag, both the drain and gate voltages were pulsed for 1 μs from three different quiescent points: \( V_{gs} (V), V_{ds} (V) = \{0, 0\}, \{-4, 0\}, \{-4, 20\} \). Fig. 4 shows the measured drain current–voltage characteristics at \( V_{gs} = +1 \) V, for all quiescent points along with the dc data.
Fig. 5. Transient measurement of (a) gain and (b) drain current after the three different transistors were subjected to a 300-μs pulse into 3-dB compression. The drain bias was 20 V, and the quiescent current was 100 mA. Black lines in (a): estimation based on drain current transients combines with the static bias-dependent S-parameters. Black lines in (b): prediction of model (1) with two exponential terms. (c) Gain collapse 0.5 ms after the pulse, measured and estimated from $I_{ds}$ transients.

Table III summarizes the drop in current at the knee voltage ($V_{gs} = +1$ V, $V_{ds} = 4$ V) for the quiescent points, relative to the $\{0, 0\}$ point. Pulsing the gate from below the pinchoff voltage ($−4$ V) results in a small drop in current, by 2%–3%, whereas a large dc voltage (20 V) on the drain cause a larger degradation in the maximum current, more than 10% for all buffers. This indicates that traps excited by high fields could cause degradation in the RF performance. The devices with the standard buffer show slightly larger current slump compared with HEMTs with reduced doping.

### B. Pulsed RF Transient Characterization

Fig. 5(a) and (b) shows the gain and dc drain current as a function of time after the RF-pulse for the three devices at a quiescent drain voltage of 20 V and a drain current of 100 mA. The output power is 35 dBm, leading to 3-dB gain compression. Immediately after the pulse, the gain drops by $−0.5$ dB, accompanied by a drop in the drain current down to 60 mA. The gain drop is larger for quiescent bias points closer to the pinchoff.

Whereas all the different devices experience a similar drop in gain and current, the devices with the more highly doped standard buffer recovers very close to the quiescent levels within ~10 ms, whereas the initial recovery is slower in the two variations with a lower Fe content.

The correlation between the drop in gain and drain current makes it possible to roughly evaluate transient effects measuring only the latter, allowing a much simpler setup. This is shown in Fig. 5(a), where the gain transient is estimated using the measured $I_{ds}(t)$ and the static S-parameter measurements at different values of $I_{ds}$. It is assumed that the matching is only slightly affected by the current transient, so that the amplifier gain varies with drain current as $|S_{21}|^2$. Furthermore, it is also assumed that the dynamic $|S_{21}|$ during the transient is equal to the static $|S_{21}|$ measured at the same drain current and voltage. Fig. 5(c) shows the measured and predicted gain collapse versus output power 0.5 ms after the pulse, compared with the quiescent level, for the three doping profiles. While this method does not exactly predict the transient, it gives a rough indication of the gain collapse as well as the recovery time.

Analysis of time constants can give insights on the mechanisms behind the current collapse and subsequent recovery [14], and on the reasons to the differences between different buffers. By fitting the measured data to a sum of exponentials

$$I_{ds} = I_{dsq} + \sum a_n \exp \left(\frac{-t}{\tau_n}\right)$$  \hspace{1cm} (1)

we can extract the different time constants $\tau_n$ (in the following numbered in order of magnitude with $\tau_1$ being the largest one) and corresponding amplitudes $a_n$ using optimization.

A good fit could be obtained for different conditions with two exponential terms in (1). Fig. 5(b) shows the fit of the model to the measured transients. The time constants $\tau_n$ and the corresponding amplitudes $a_n$ at different output power levels are shown in Fig. 6. Two distinct time constants in the millisecond range, separated by 1–1.5 orders of magnitude, can be seen in all devices. The time constants are very similar across the different devices, but their magnitudes differ. The highly doped standard buffer has a larger magnitude of the faster time constant but a smaller magnitude of the slower one, explaining the faster recovery in these devices. The time constants increase with power level, as well as the amplitudes.

### C. Pulsed DC Transient Characterization

In order to mimic PA operation, we have pulsed the drain voltage from a quiescent point of 20 V to higher values for 1 μs and measured the current response after the pulse. The gate voltage was kept constant at a value giving 100 mA quiescent current. The current was chosen as low as possible.
to minimize effects of self-heating but high enough to keep the transconductance in a fairly linear region during the experiment. Fig. 7 shows the drain current transients after pulses to $V_{ds} = 40$ V from 20 V for a quiescent current of 100 mA, compared with model (1) with three time constants extracted using optimization. The time constants $\tau_n$ (with $\tau_1$ being the largest one) and the corresponding amplitudes $a_n$ are plotted in Fig. 8 versus the magnitude of the voltage pulse ($\Delta V_{DS} = V_{pulse} - V_{quiescent}$).

The drain current shows a similar drop in current and subsequent recovery as during the pulsed RF measurement (Figs. 5 and 6), indicating that pulsing the drain voltage can indeed be used to characterize transient effects in these transistors and find their associated time constants. The initial current drop is a few milliamperes larger in the standard devices, which is consistent with the results of the pulsed $I$–$V$ characterization in Section IV-A, where the more highly doped standard devices showed slightly stronger current collapse. On the other hand, the recovery is faster for the more highly doped buffer, similar to what the RF transient measurements showed.

The two slowest time constants $\tau_1$ and $\tau_2$, around 300 and 10 ms, respectively, after a 40 V pulse, have the largest amplitudes and, thus, dominate the current collapse and subsequent recovery. Similar to the case of RF pulsing, the difference between the samples with varied buffer doping profile is that reduced doping seems to lead to increased $|a_1|$ but reduced $|a_2|$, thus making the recovery slower by shifting it to the slower time constant. The time constants also decrease with smaller voltage pulses, which is consistent with RF pulsing leading to shorter time constants at lower power levels.

Drain current transient measurements were also carried out at elevated ambient temperatures in order to extract the activation energies of the traps behind the different time constants. Fig. 9(a) shows the Arrhenius plots and the extracted energies. The two larger time constants have similar activation energies in the range of 0.5–0.6 eV for all samples, whereas the activation energy of the fastest time constant is difficult to extract without faster measurement systems. The activation energies were extracted using temperatures measured on the fixture, without considering any temperature gradient inside the device. Assuming a 25 K/W · mm thermal resistance between the fixture and the buffer would result in a 21 K
temperature rise at the quiescent bias and an 80 meV higher activation energy. The elevated temperature also affected the temperature rise at the quiescent bias and an 80 meV higher amplitude of the slowest time constant [Fig. 9(b)]. The amplitudes associated with the two time constants [Fig. 9(b)]. The amplitude of the slowest time constant $|a_1|$ increased for higher temperatures, whereas $|a_2|$ decreased, resulting in the total current collapse staying constant with temperature.

V. CONCLUSION AND DISCUSSION

In this paper, the trapping effects in differently Fe-doped GaN buffers in high-power microwave AlGaN/GaN HEMTs were evaluated. Traditional performance indicators, such as dc and pulsed $I–V$ characteristics, $S$-parameters, and IP3, are very similar between the different buffers and the small differences cannot conclusively be attributed to the buffer doping rather than to unintended variation in epitaxy or processing.

However, a large impact of the buffer doping level is seen in the recovery after a large dc or RF pulse. We detect two detrapping processes with the same activation energy but different time constants. In the highly doped buffers, the initial current collapse is slightly larger, but a larger proportion of the trapped electrons is emitted through the faster process, resulting in a faster recovery. In contrast, lower Fe concentration in the buffer and channel leads to a slower recovery, since a larger proportion of the electrons is emitted through the slower process.

The activation energy of the two time constants (0.6 eV) is in the same range as what have been found in the previous studies of GaN buffers with as well as without Fe doping [8]–[10], [15]–[18]. The faster time constant increased in significance with increasing Fe doping similar to what was reported in [8]–[10]. However, the second time constant associated with the same energy with the opposite Fe concentration dependence was not found in other studies. In [8], a slower time constant was found at 0.82 eV, but it differed from our observations not only in the activation energy, but also in the fact that it did not show notable dependence on Fe doping.

The similarity of the drain current transients after pulsing the drain voltage from its operating point to the transient after a high-power RF pulse shows that the former can be used as a simple method to characterize memory effects in the RF domain and can be used as a complement to pulsed $I–V$ characterization.

The almost identical dc and S-parameter performances of the different Fe-doping profiles imply that the buffer doping profile largely can be optimized for minimizing the negative effects of trapping in the buffer. However, it should be noted that the buffer doping differences between the tested samples are relatively small. Reducing the compensation doping further would eventually lead to short-channel effects and decreased breakdown voltage [13], as well as larger effects on trapping [11].

REFERENCES


Olle Axelsson received the M.Sc. degree in engineering physics from the Chalmers University of Technology, Gothenburg, Sweden, in 2010, where he is currently pursuing the Ph.D. degree with the Microwave Electronics Laboratory. His current research interests include the characterization of robustness, linearity, and trapping effects in GaN HEMTs.

Sebastian Gustafsson (S’13) received the M.Sc. degree in electrical engineering from the Chalmers University of Technology, Gothenburg, Sweden, in 2013, where he is currently pursuing the Ph.D. degree with the Microwave Electronics Laboratory Group. His current research interests include RF metrology and GaN HEMT characterization.

Hans Hjelmren (M’91–SM’11) received the Ph.D. degree in electrical engineering from the Chalmers University of Technology, Gothenburg, Sweden, in 1991. His thesis work dealt with the numerical simulation of hot electrons in GaAs devices. He is currently an Associate Professor with the Chalmers University of Technology. His current research interests include the different aspects of TCAD.

Niklas Rorsman (M’10) received the M.Sc. degree in engineering physics and the Ph.D. degree in electrical engineering from the Chalmers University of Technology, Gothenburg, Sweden, in 1988 and 1995, respectively. He returned to the Chalmers University of Technology in 1998, where he is currently leading the microwave wide bandgap technology activities and investigating the application of graphene in microwave electronics.

Hervé Blanck received the M.S. and Ph.D. degrees from the University of Strasbourg, Strasbourg, France, in 1986 and 1989, respectively. He then joined the Central Research Laboratory, Thomson-CSF, Orsay, France, where he was involved in the GaAs HBT development from 1990 to 1996. In 1996, he joined United Monolithic Semiconductors Gesellschaft mit beschränkter Haftung, Ulm, Germany, where he has been the Manager of the Technology Development Department since 2003.

Jörg Splettstoesser, photograph and biography not available at the time of publication.

Jim Thorpe, photograph and biography not available at the time of publication.

Thomas Roedle, photograph and biography not available at the time of publication.

Mattias Thorsell (S’08–M’11) received the M.Sc. and Ph.D. degrees in electrical engineering from the Chalmers University of Technology, Gothenburg, Sweden, in 2007 and 2011, respectively. He is currently an Assistant Professor with the Chalmers University of Technology. His current research interests include the characterization and modeling of nonlinear microwave semiconductor devices.
Paper B

The Effect of Forward Gate Bias Stress on the Noise Performance of Mesa Isolated GaN HEMTs

O. Axelsson, M. Thorsell, K. Andersson, and N. Rorsman

The Effect of Forward Gate Bias Stress on the Noise Performance of Mesa Isolated GaN HEMTs

Olle Axellsson, Mattias Thorsell, Member, IEEE, Kristoffer Andersson, and Niklas Rorsman

Abstract—This study investigates degradation of gallium nitride (GaN) high-electron mobility transistor (HEMT) noise performance after both dc and RF stress with forward gate current. The results are used to facilitate optimization of the robustness of GaN low-noise amplifiers (LNAs). It is shown that forward biasing the gate of a GaN HEMT results in permanent degradation of noise performance and gate current leakage, without affecting S-parameters and drain current characteristics. The limit of safe operation of the 2 × 50 μm devices in this study is found to be between 10 and 20 mW dissipated in the gate diode for both dc and RF stress. We propose that degradation could be caused by excessive leakage through the mesa sidewalls at the edges of each gate finger. Circuit simulations may be used together with device robustness rating to optimize LNAs for maximum input power tolerance. Using a resistance in the gate biasing network of 10 Ω, it is estimated that an LNA utilizing a 2 × 50 μm device could withstand input power levels up to 33 dBm without degradation in noise performance.

Index Terms—Semiconductor device noise, semiconductor device reliability, robustness, MODFETs, MODFET amplifiers.

I. INTRODUCTION

DUE to its combination of high power handling capability and low noise figure, the Gallium Nitride (GaN) High Electron Mobility Transistor (HEMT) has received much attention for its potential in highly rugged low noise amplifiers (LNAs). Destructive tests have shown that LNAs with GaN transistors survive input power levels more than one order of magnitude higher than a typical Gallium Arsenide (GaAs) LNA [1]. The use of GaN technology thus alleviates or eliminates the need of protection circuitry in receivers, potentially improving the need for system robustness and performance. We specifically address the case where the devices are isolated mesa and directly contacting the GaN channel through the mesa sidewall, giving rise to additional conduction path for the gate current. The effect of this leakage path on the reverse leakage has been studied in [9], but the effects on device robustness have previously not been tested. Moreover, it is shown how degradation tests on device level can be used to analyze or optimize robustness of LNAs using circuit level simulations.

In Section II, it is experimentally demonstrated that forward gate bias stress results in degradation of noise performance. The effects on device noise parameters after stress are analyzed using noise modeling. Degradation in the HEMTs is investigated in more detail in Section III. The purpose is to establish a region of safe operation for these devices. DC stress is compared to RF input power stress, where the device is forward biased during the positive peaks of the gate voltage waveform, leading to high instantaneous gate currents. The consequences of the results on amplifier robustness are investigated and discussed using circuit
simsulations in Section IV. The results are discussed in relation to other studies in Section V and some conclusions are drawn in Section VI.

II. EFFECTS OF GATE BIAS STRESS ON NOISE PERFORMANCE

S- and noise parameters of dc-stressed devices were measured to evaluate the effect of forward bias stress on noise performance. Measurements were performed on 0.25 μm gate length AlGaN/GaN HEMTs fabricated in a Chalmers in-house MMIC process [9]. For the noise measurements, devices with 4 × 75 μm gate periphery were used. The HEMTs were dc-stressed for 5 minutes by applying a constant positive dc voltage on the gate while keeping the drain and source terminals grounded. After each stress session the noise parameters between 2–18 GHz at a low noise bias (10 V and 50 mA, corresponding to 20% of IDSS). This process was repeated for stress voltages stepped from 0 V to 4 V in 0.2 V steps.

The stress resulted in changes in the transistor noise parameters, accompanied by an increase of reverse gate leakage (Fig. 1). Specifically, the minimum noise figure increased from around 1.5 dB to 4 dB at 10 GHz and the optimum source impedance shifted towards the center of the Smith chart, particularly for low frequencies. The drain current characteristics were largely unchanged apart from a 0.5 V shift in the pinch-off voltage towards less negative values. Similar burn-in effects on the threshold voltage have previously been seen in other GaN HEMT processes [11] and do not explain the dramatic degradation in the noise performance, and the effects are not necessarily related. There are no significant differences between the S-parameters measured before and after stress for the same drain current and voltage (Fig. 2). The explanation to the dramatic effects on $F_{\text{min}}$ and $\Gamma_{\text{opt}}$ despite the S-parameters being unchanged is instead the increase in gate leakage, which is known to result in degradation of noise performance due to the increased shot noise through the gate [12]. The reason for the increase in gate leakage is further studied in the following sections.

III. DC AND RF ROBUSTNESS CHARACTERIZATION

For successful design of highly robust LNAs, it is important to know which operating conditions the transistor can sustain without permanent damage. Therefore, more detailed stress tests were conducted on several GaN HEMTs from the same wafer, in order to find the limiting factors of safe operation in terms of forward gate bias and RF input power rating.
Fig. 3. Intrinsic forward gate $I-V$ characteristics of $2 \times 50 \, \mu m$ HEMT, during dc and RF stress. Two dc-stressed HEMTs and one RF-stressed with +27 dBm available input power are shown. The intrinsic RF gate voltage is deembedded from the measured extrinsic waveforms shown in the inset for different RF power levels. The available power levels for the different stress sessions were chosen so as to have a constant 1 V increase in voltage amplitude between each session ($P_{in, dBm} = 16.7, 18.6, 20.1, 21.4, 22.6, 24.4, 25.3, 26.0, 26.7$).

**A. Setup and Procedure**

DC stress was performed on $2 \times 50 \, \mu m$ gate periphery devices in the same way as described in Section II and the dc $I-V$ characteristics were measured after each stress session. The current was measured continuously during the stress.

RF input power stress was also performed on a $2 \times 50 \, \mu m$ transistor in order to compare the degradation seen during constant dc forward bias to a more realistic case where the device is forward biased by the RF swing. The device was subjected to RF input power stress for five minutes at a typical LNA operation bias ($V_{gs} = -4 \, V$ and $V_{ds} = 10 \, V$). A large signal network analyzer (LSNA) was used to measure the current and voltage waveforms at the input during stress. The RF input power was increased in steps from +17 dBm to +27 dBm. After each RF power stress level, the dc characteristics were measured. The RF frequency was chosen as low as possible with the available equipment (600 MHz), to facilitate deembedding of the pads, the series resistance and the nonlinear gate capacitance.

**B. Gate Diode Characterization**

To understand what happens during stress, it is important to study the forward gate $I-V$ characteristics of the devices under test. Fig. 3 shows the measured dc characteristics during dc stress and the intrinsic waveform during RF stress. The measured extrinsic waveforms from which the intrinsic curves were extracted are shown in the inset. Under dc conditions, forward conduction starts from around 0.8 V forward bias and initially has a large series resistance, varying between 350 $\Omega$ and 700 $\Omega$ for different devices. The gate current (and the diode series resistance) does not scale with the gate periphery of the device. One dc-stressed device with a low series resistance and one with a high resistance are shown in Fig. 3. The RF measurement is not directly comparable since there is also a voltage on the drain terminal, but the characteristics for low forward voltages is similar to the dc case with a high series resistance. However, the RF measurement also reveals that another diode turns on at $\sim 5.5 \, V$, where the series resistance decreases to around 60 $\Omega$. This behavior cannot be seen under dc conditions since the devices suffer catastrophic failure before the second turn-on region is reached. Based on these observations, an equivalent circuit is extracted (Fig. 4), which describes the dc and RF forward characteristics of the gate (Fig. 5).

Leakage through the sidewalls of the mesa used for device isolation may be a physical explanation to this behavior, since it provides an additional Schottky diode directly to the GaN channel at the edges of each gate finger where the gate metal extends outside the mesa [9] (Fig. 6). This diode is expected to turn on at lower forward voltages than the AlGaN Schottky diode, because of GaN’s lower band gap. This, together with the observation that diode 1 (Fig. 4) has a high series resistance which does not scale with gate periphery per finger, makes it probable that diode 1 is associated with the sidewall leakage, concentrated in a small area at the edges of each finger. The second turn-on of diode 2 starting around 5.5 V is then associated with the common gate diode. If the two diodes in Fig. 4 are physically separated, damage due to high gate current could be localized to one of them and both components should be considered separately when evaluating the robustness of the devices.
The dc stress resulted in permanent degradation of the gate current–voltage characteristics, with large increases of gate leakage both in the forward and reverse direction. Fig. 7(a) shows the $I_G-V_G$ characteristics for one device after different stages of the stress test. Degradation started at dc forward voltages higher than 2.8 V, and another dramatic increase of gate leakage occurred at 3.8 V. At dc forward bias, the devices were thus severely damaged before reaching the turn-on of diode 2 (Fig. 4). It is thus likely that the main contribution of stress effect is caused by degradation in diode 1.

For RF stress, devices experienced similar degradation (Fig. 7(b)). The first noticeable effects on the $I_G-V_G$ characteristics occurred for an input power of 25.3 dBm, corresponding to an intrinsic peak gate voltage and current of 10 mA and 6 V respectively. The dc component was 2 mA and the gate power dissipation 10 mW at this input power. Very severe degradation was seen after an input power of 26.7 dBm corresponding to of 35 mA peak current, 7 V peak $V_{GS}$, 5.5 mA dc current and 36 mW power dissipation.

Fig. 8 shows a comparison between degradation after dc and RF stress, displaying gate leakage for $V_{GS}=-4$ V as a function of stress level, in terms of gate voltage, current and dissipated power in the gate. The two dc stressed devices, exhibiting different series resistance, start to degrade at a similar forward voltage, even though the corresponding gate current is significantly different. At a first glance this indicates an electric field induced degradation mechanism. In this case, the fact that the RF stressed device survived peak voltages up to 6 V without notable degradation, whereas the dc stressed devices degraded below 4 V points to a degradation mechanism slower than the length of the RF gate current pulses, which are in the nanosecond scale.

However, the complete picture is more complicated and a current or dissipated power induced degradation mechanism cannot be excluded. The damage may be localized and the global current level does not necessarily reflect local current densities. For example, the variation in critical current and power dissipation between different devices may be due to variations in the not so well defined area of the mesa sidewall diode (diode 1).

When comparing the RF and dc case, it is important to note that local current density in diode 1 or 2 may be the cause of failure during RF stress whereas only diode 1 is conducting during dc forward bias, making the current highly localized to the mesa sidewalls. This could explain the large differences in power dissipation and dc current required to damage the devices between the dc and the RF case. In Fig. 8(b) and (c) the contribution of diode 1 as extracted from the model is displayed separately. When only considering the current through diode 1, the large discrepancy between dc and RF stress disappears, and the diode degrades at similar current and power dissipation.

Fig. 7. Gate current–voltage characteristics at $V_{DS}=0$ V after different stages of the (a) dc and (b) RF stress test. The dc forward bias is swept from 0 V to +4 V and the available RF input power is swept from +17 dBm to +27 dBm. The colors in (b) correspond to measurements after different stress sessions in Fig. 3.

Fig. 6. Top view photograph of one of the tested transistors and a schematic cross-section view of the device along one of the gate fingers. The unit gate width is 50 μm.
Fig. 8. Comparison of degradation after dc stress and RF stress. The degradation of gate leakage at a gate bias of $-4\, V$ is plotted as a function of forward voltage (a), forward gate current (b) and dissipated power in the gate (c). Two otherwise similar dc-stressed devices are shown, one with low series resistance below 5.5 V and one with high resistance. For RF stress, the degradation is shown as a function of total gate current as well as only the $I_1$ component in Fig. 4.

In conclusion, based on these results it is not possible to isolate one main failure mechanism. Furthermore, a dual gate diode model is needed to enable accurate prediction of the device characteristics in a circuit simulator.

IV. AMPLIFIER ROBUSTNESS ANALYSIS

In an LNA, it is common to protect the transistor from excessive forward conduction by using a large series resistance (typically in the $k\Omega$ range) in the gate bias network [13]. When dc current starts to flow, the voltage drop over the gate bias resistance causes the dc gate voltage to decrease, limiting the forward gate current. However, other studies have shown that large negative gate voltages are also harmful to devices and cause degradation [3]–[5]. The amplifier robustness is thus determined by the transistor’s survivability to both forward and reverse bias stress, in interplay with its gate $I-V$ characteristics, the bias resistance and the matching network. Evidently, there is a trade-off in choosing the series resistance to avoid both forward and reverse bias.

Based on the current, voltage and power dissipation levels presented above, we may define a region of safe operation. If the transistor’s $I-V$ characteristics and region of safe operation are known, harmonic balance simulations can be used to simulate the sustainable input power of an amplifier and optimize the bias resistance for maximum robustness. To demonstrate this, we use a simple model of an LNA input, using the model in Fig. 9 with a $2 \times 50 \mu m$ HEMT. The transistor is assumed to be matched to an impedance close to 50 $\Omega$ at the gate. In real LNAs, this value depends on frequency, device size and the matching networks, but the size is often chosen as to provide simple matching to 50 $\Omega$ at the design frequency. Losses in the input matching network and parasitic resistances are not included in this model, so the robustness will likely be higher in real amplifiers.

Assuming for example that power dissipation in the mesa sidewall diode is the main cause of failure during forward bias stress, we plot contours of the power dissipation in both diodes levels as during dc stress. If the degradation is power or current induced, this could be an indication that the current associated with mesa leakage, alone is responsible for degradation also during RF stress.

In conclusion, based on these results it is not possible to isolate one main failure mechanism. Furthermore, a dual gate diode model is needed to enable accurate prediction of the device characteristics in a circuit simulator.
and gate voltage as functions of available input power and bias resistance (Fig. 10). According to the stress tests in this paper, the power dissipation in diode 1 should not exceed 7.5 mW. A typical value of the critical drain–gate voltage for degradation is 30 V [5], which means that if the drain is biased at 10 V, the gate voltage should not go below −20 V. The maximum input power that can be sustained without running into any of these limits would then be approximately 33 dBm, which is achieved with a resistance of 10 kΩ. The power dissipation in diode 2 is then around 1 mW, corresponding to 10 mA/mm. Choosing the resistance too small is more detrimental to robustness than choosing it too large, since the current increases rapidly for decreasing resistance whereas the dc voltage drop increases more slowly when the resistance increases (Fig. 10).

V. DISCUSSION

The devices under test in this study have very different forward gate I–V characteristics compared to previous studies of degradation due to forward gate current, which have used devices with a more traditional diode behavior, turning into full conduction between 1 V and 2 V [5]–[8]. Whereas the effects on the devices from forward bias–gate stress are similar in these studies and the dc voltages required to damage the devices in the same range, the devices in this study therefore suffered degradation at much lower gate current densities than in [5]–[8]. This is in line with the proposed leakage path through the mesa edge, since the current is concentrated to a very small area which could lead to localized damage.

This has important effects on LNA robustness. Doing the analysis in Section IV using the results from [7], where the devices withstood a gate current of 900 mA/mm and voltage of 3.5 V without severe degradation, would result in an optimal bias resistance of 30 Ω • mm, giving a maximum input power of 43 dBm for a 2 × 50 μm device. Minimization of mesa leakage or use of other methods such as ion implantation, which can provide device isolation without the formation of a mesa where the GaN channel is in contact with the gate metal [14], could thus be essential for robust LNA design.

VI. CONCLUSION

DC and RF stress tests have shown that devices subjected to forward gate bias stress degrade permanently, resulting in increased reverse gate leakage and impaired noise performance.

More detailed analysis of the devices under test showed that the devices behave like two parallel gate diodes under forward bias. The high resistance diode that dominates for low forward bias voltages is proposed to be associated with mesa sidewall leakage and seems to cause degradation during both dc and RF stress in this study.

Circuit simulations have been used to estimate that a 2 × 50 μm transistor used in an LNA survives 33 dBm input power to the gate if a 10 kΩ resistor is used in the gate bias network, given the damage levels indicated by stress tests. Device design for optimizing LNA robustness should focus on gates with minimized mesa leakage.

The increase in noise figure resulting from forward bias stress could significantly degrade the sensitivity of a GaN receiver. The stress effects can most easily be followed by monitoring the gate current or the noise figure in the system, as the increase in noise figure is noticeable at stress levels far below the levels required to affect S-parameters and output characteristics.

REFERENCES

Olle Axelsson was born in Göteborg, Sweden, in 1986. He received the M.Sc. degree in engineering physics from Chalmers University of Technology, Göteborg, in 2010, where he is currently working toward the Ph.D. degree. His research interests are design and characterization of robust and linear gallium-nitride-based amplifiers.

Mattias Thorsell (S’08–M’11) received the M.Sc. and Ph.D. degrees in electrical engineering from Chalmers University of Technology, Göteborg, Sweden, in 2007 and 2011, respectively. He is currently an Assistant Professor with Chalmers University of Technology. His research interests are characterization and modeling of nonlinear microwave semiconductor devices.

Kristoffer Andersson received the M.Sc. and Ph.D. degrees in electrical engineering from Chalmers University of Technology, Göteborg, Sweden, in 2001 and 2006, respectively. Until 2013, he was an Assistant Professor with the Microwave Electronics Laboratory, Department of Microtechnology and Nanoscience (MC2), Chalmers University of Technology, where he worked on modeling and characterization of gallium nitride high-frequency transistors. Since 2013, he has been a Senior Researcher at Ericsson Research, Göteborg, working with gallium nitride MMICs.

Niklas Rorsman received the M.Sc. degree in engineering physics and the Ph.D. degree in electrical engineering from Chalmers University of Technology, Göteborg, Sweden, in 1988 and 1995, respectively. His thesis work dealt with the development of InP-based HEMT and MMIC materials and processes, as well as modeling of HEMTs. From 1996 to 1998, he was with Ericsson Microwave Systems, Mölndal, Sweden, where he was involved in the modeling of III–V devices and MMIC design. In 1998, he returned to Chalmers University, where he is currently leading the microwave wide bandgap technology activities and investigating the application of graphene in microwave electronics.


Paper C

Impact of Trapping Effects on the Recovery Time of GaN Based Low Noise Amplifiers

O. Axelsson, N. Billström, N. Rorsman, and M. Thorsell

IEEE Microwave and Wireless Components Letters vol.26, no.1, pp.31-33, December 2015
Impact of Trapping Effects on the Recovery Time of GaN Based Low Noise Amplifiers
Olle Axelsson, Niklas Billström, Niklas Rorsman, Member, IEEE, and Mattias Thorsell, Member, IEEE

Abstract—This study investigates recovery time of the gain of AlGaN/GaN HEMT based low noise amplifiers (LNA) after an input overdrive pulse. Three LNAs, fabricated in two commercial MMIC processes and a Chalmers in-house process, are evaluated. The Chalmers process has an unintentionally doped buffer instead of the intentional Fe doping of the buffer which is standard in commercial GaN HEMT technologies. It is shown that the LNAs from the two commercial processes experience a severe drop in gain after input overdrive pulses higher than 28 dBm, recovering over a duration of around 20 ms. In contrast the LNA fabricated in-house at Chalmers experienced no visible effects up to an input power of 33 dBm. These results have impact for radar and electronic warfare receivers, which need to be operational immediately after an overdrive pulse. The long time constants suggest that these effects are due to trapping in the transistors with the Fe doped buffer playing an important role.

Index Terms—Gallium nitride, MMICs, low-noise amplifiers, robustness.

I. INTRODUCTION

The high survivability of low noise amplifiers in GaN technology has made them an interesting candidate for radar and electronic warfare receivers, which may be subjected to high input powers from electronic warfare weapons or close-range reflections. Several studies have shown that GaN LNAs survive input powers up to around 40 dBm [1], [2], whereas GaAs LNAs only can withstand input powers up to around 20 dBm. Although GaAs LNAs with integrated limiters offer survivability up to 40 dBm with good noise performance [3], the use of robust GaN LNAs reduces the need of limiting and enables integration with GaN switches or other functions.

High survivability is commonly achieved by use of a series resistance in the kΩ range in the gate bias network to protect the transistor from high forward gate currents during input overdrive, which has been identified as a major failure mechanism in GaN LNAs [4]. The voltage drop across the bias resistance decreases the dc gate voltage below the supply voltage as a gate current starts to flow. However, the LNA often need not only to survive without permanent damage but also to be operational immediately after the pulse. In [5], recovery back to normal operation was analyzed. It was found that the recovery time is in the nanosecond range dominated by the RC constant between the bias resistance and the capacitances in the input network and the transistor. However, many studies have also shown slow transients due to electron trapping in the buffer and at the surface in GaN transistors during power amplifier operation with high drain voltages [6]–[8], giving rise to DC-to-RF dispersion and impaired performance at high output power levels. The impact of these phenomena on recovery times in LNAs has previously not been studied.

In this paper, we study gain transients and recovery times after a high power pulse in LNAs fabricated in three different GaN-MMIC processes. Trapping is found to have a significant impact on the recovery time, causing a dramatic decrease in gain for several milliseconds after a pulse at power levels as low as 10 dB below the input power ratings. Optimizing the GaN HEMT technology to reduce trapping is thus essential in order to obtain highly robust GaN LNAs with short recovery times.

II. MEASUREMENTS

A. Devices Under Test

Three two-stage, 2–6 GHz GaN MMIC LNAs were tested. These circuits are suitable for example in wideband radar and electronic warfare systems. Specifications, figures of merit, and recommended bias for the three amplifiers are listed in Table I. LNA #1 was designed using UMS GH25-10 foundry service and LNA #2 [2] using a Chalmers GaN process [9], whereas LNA #3 (TGA2611) is commercially available from Qorvo. All the amplifiers use the same nominal gate length, 0.25 μm. It is however possible that the gate length of the Chalmers MMIC deviates more from the nominal value. The Chalmers process differs in that it does not use Fe doping of the GaN buffer, the standard method to decrease leakage currents and increase the breakdown voltage. Fe doping has been identified as a cause of electron trapping in GaN HEMTs [7].

LNA #1 uses both parallel and inductive series source feedback to obtain flat gain, good match and low noise figure. LNA #2 and #3 use only inductive series source feedback, which give excellent noise performance but a non-flat gain over the 2–6 GHz band. The input matching network of all the tested amplifiers have the topology shown in Fig. 1 (b). The transient response of the matching network after a gate current step [5] is dominated by the RC constant between the gate bias resistance $R_{bias}$ (in the kΩ range) and the capacitors $C_g$ and $C_d$ (typically...
TABLE I
SPECIFICATIONS OF AMPLIFIERS UNDER TEST

<table>
<thead>
<tr>
<th>LNA</th>
<th>Manufacturer</th>
<th>Frequency range [GHz]</th>
<th>Gain @ 3 GHz [dB]</th>
<th>NF @ 3 GHz [dB]</th>
<th>$I_{ds}$ [mA]</th>
<th>$V_{dd}$ [V]</th>
<th>Max $P_{out}$ [dBm]</th>
<th>Gate bias resistance [Ω]</th>
<th>Gate width [stage 1/ stage 2] [μm]</th>
<th>$f_{max}$ [GHz]</th>
<th>$P_{MAX}$ [dBm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>LNA #1</td>
<td>UMS</td>
<td>2-6</td>
<td>21</td>
<td>2.4</td>
<td>80</td>
<td>10</td>
<td>41</td>
<td>4000</td>
<td>4x100/4x100</td>
<td>31.82</td>
<td>25</td>
</tr>
<tr>
<td>LNA #2</td>
<td>Chalmers</td>
<td>2-6</td>
<td>15</td>
<td>1.5</td>
<td>80</td>
<td>10</td>
<td>39</td>
<td>1000</td>
<td>8x50/8x50</td>
<td>23.82</td>
<td>20</td>
</tr>
<tr>
<td>LNA #3</td>
<td>Qorvo</td>
<td>2-6</td>
<td>28</td>
<td>1.0</td>
<td>110</td>
<td>10</td>
<td>33</td>
<td>N/A</td>
<td>N/A</td>
<td>351590</td>
<td>29</td>
</tr>
</tbody>
</table>

1 Recommended low noise bias settings. The measurements are done at the same bias points ($I_{ds}$=110 mA, 165 mA and $V_{dd}$=10 V) for all LNAs for a more fair comparison of recovery times.

2 Max $P_{out}$ is defined as the highest safe input power in order to avoid permanent degradation. The value for LNA #1 and #2 is found by a step stress test where gain and noise figure are measured after 10 seconds of stress with increasing input power levels [2]. The value for LNA #3 is taken from the datasheet, and the lower value could thus be due to differences in definition or methodology rather than in the devices themselves.

B. Setup and Procedure

Fig. 1 shows a schematic of the setup used to characterize the recovery of the LNAs after a high power pulse. A vector signal generator is used to provide an input signal at 3 GHz which is pulsed to high RF power levels before backing off to −20 dBm. The output power and drain current transients during and after the pulse are measured using a vector signal analyzer and an oscilloscope with a current probe, respectively. For LNA #1 and #2, we measure the first stage drain current, whereas LNA #3 has a common bias pad for both drain terminals, allowing only measurement of the total current. The period of the input signal was 100 ms to allow for full recovery after the pulse. The peak input power was swept from 20 to 33 dBm in 1 dB steps. The measurements were carried out at 50%, 100% and 150% of the recommended quiescent drain current of TGA2611 (110 mA). The impact of the pulse length was investigated by a parametric experiment (1 μs, 10 μs and 100 μs).

III. RESULTS

Fig. 2 shows the gain transients after a 10 μs pulse for the three tested amplifiers, with the $I_{ds}$ transients shown in insets. The gain of both LNA #1 and #3 drops significantly after input power levels around 28 dBm or higher, and recovers within around 20 ms. The gain transients of the two LNAs look very similar. However, LNA #3’s gain exhibits a drop at lower input power (2 dB). The drop in gain is accompanied by a decrease in the drain current in both amplifiers. The drain current plots are not directly comparable since the currents of the two stages in LNA #3 cannot be measured separately. However, most of the current drop should occur in the first stage which is subjected to the larger input power.

The LNA fabricated in the Chalmers process experienced almost no gain collapse, in contrast to the two commercial processes. The drain current did decrease immediately after the pulse, but the drop was markedly smaller compared to the other two LNAs at similar input power levels.

A. Bias and Pulse Length Dependence

Fig. 3 shows the gain of the three LNAs 0.1 ms after the end of the pulse, versus peak input power, at different quiescent currents. The gain collapse is more severe for lower quiescent cur-
and forth the commercial processes compared to the other two LNAs. An increased gain immediately after a 100 μs pulse, whereas after longer pulses. The Chalmers LNA experienced a gain collapse after short pulses are less severe than after longer pulses. The Chalmers LNA experienced an increased gain immediately after a 100 μs pulse, in sharp contrast to the other two LNAs.

![Fig. 3. Gain 0.1 ms after a 10 μs pulse, at different bias points for the three amplifiers: LNA #1 (blue), LNA #2 (red) and LNA #3 (yellow).](image1)

Fig. 4 shows the pulse length dependence of the gain 0.1 ms after the pulse. The gain collapse after short pulses are less severe than after longer pulses. The Chalmers LNA experienced an increased gain immediately after a 100 μs pulse, in sharp contrast to the other two LNAs.

![Fig. 4. Pulse length dependence of the gain 0.1 ms after the pulse for the three amplifiers: LNA #1 (blue), LNA #2 (red) and LNA #3 (yellow). The circuits were biased at 110 mA and a drain voltage of 10 V.](image2)

At higher drain current levels, the gain collapse is reduced and occurs at higher input power levels. Fig. 4 shows the pulse length dependence of the gain 0.1 ms after the pulse. The gain collapse after short pulses are less severe than after longer pulses. The Chalmers LNA experienced an increased gain immediately after a 100 μs pulse, in sharp contrast to the other two LNAs.

IV. CONCLUSION

This paper presents transient measurements after input overdrive pulses in LNAs fabricated in three different GaN HEMT MMIC processes. These measurements reveal a significant gain and drain current collapse after input power pulses higher than 28 dBm for the two commercial processes, whereas the LNA fabricated in the in-house Chalmers process did not experience any gain collapse up to 33 dBm and a much smaller current drop. The results for the commercial processes are also in contrast to the results of [5], which showed recovery times in the order of microseconds for a GaN LNA subjected to 38 dBm input power. The cause of the gain and drain current collapse is probably related to the transistors, since time constants in the bias networks are typically in the nanosecond range whereas the measured transients have time constants in the range of milliseconds. The large differences between the Chalmers process as well as the MMIC used in [5], compared to the two commercial processes also indicate that optimizing the transistor technology to reduce trapping effects is essential in reducing the recovery time of robust LNAs. In particular, the buffer doping seems to play an important role and the unintentionally doped buffer of the Chalmers process leads to smaller gain and drain current drop compared to the intentionally Fe-doped buffers which are common in commercial GaN transistors. The pulse length dependence of the gain collapse indicates that there are also transient effects during the high power pulses, with time constants in the range of 1–100 μs. These transients are probably due to either heating or trapping effects in the transistor.

For radar and electronic warfare applications, these effects could mean that the input power ratings of the receivers must be lowered under conditions where fast recovery is required, for example in the presence of High Power Microwave (HPM) weapons designed to disable receivers. To a large degree, the gain collapse after high power pulses can be mitigated, at the cost of increased noise figure and power consumption, by biasing the LNAs at a higher quiescent current. It is likely that amplifiers operating at higher frequencies suffer from similar effects for a similar voltage swing on the gate during input overdrive, but this remains to be investigated. The correlation between the gain and drain current transients suggests that measurement of the latter could be sufficient to characterize the recovery time, allowing a simpler measurement procedure.

While causing recovery time issues in LNAs, the Fe doped buffer is favored in power amplifiers and switch circuits since it increases the resistivity (lower leakage) and breakdown voltage. The higher resistivity which increases the output conductance (but also to other differences like AlGaN barrier thickness and aluminum content, and gate length). Trade-offs or new solutions are thus necessary in order to integrate different functions on a single chip with good performance.

REFERENCES

Paper D

Highly Linear Gallium Nitride MMIC LNAs

O. Axelsson, and K. Andersson

*IEEE Compound Semiconductor Integrated Circuit Symposium*, La Jolla, USA, October 2012
Highly Linear Gallium Nitride MMIC LNAs

Olle Axelsson, Student Member, IEEE and Kristoffer Andersson, Member, IEEE
Microwave Electronics Laboratory, Chalmers University of Technology, 412 96, Göteborg, Sweden

Abstract — In this paper, two Low Noise Amplifiers designed in Gallium Nitride HEMT MMIC technology are presented. The focus of the designs is to achieve good linearity at low power consumption and acceptable noise figure. The first design achieves an OIP3/PDC of 12 using traditional LNA design techniques. In a second design, the OIP3 is improved by 2 dB, raising OIP3/PDC to 19, among the highest figures reported for GaN LNAs. This is achieved by using both inductive source feedback and drain-gate RC feedback.

Index terms — MMICs, Low-noise amplifiers, HEMTs, Gallium Nitride, Linearity

I. INTRODUCTION

Gallium Nitride HEMTs have, besides their obvious potential for high power microwave and RF applications, also received attention as a competitor to GaAs pHEMTs in LNAs for receiver applications. While so far being comparable or slightly higher in noise figure, Gallium Nitride HEMTs are believed to offer more in terms of robustness, integration and linearity, which could potentially give them an advantage in certain applications [1]. However, these prospects have so far been realized, investigated and analyzed only to a very small extent. While some studies have been investigating robustness issues and failure modes of GaN LNAs [2] and steps have been taken towards MMIC GaN transceivers [1], not much is known about the suitability of GaN LNAs for applications requiring extremely low intermodulation distortion. Very high figures of third order output intercept points (OIP3) have been reported but are accompanied by a similar increase of DC power consumption [3-8]. When comparing linearity for a specific amount of DC power consumption, GaN LNAs have yet to significantly outperform commercial GaAs pHEMTs, which can offer OIP3/PDC as high as 17 [9]. Also, little has been reported about specific design issues and considerations when designing GaN LNAs towards high linearity.

In this paper, two 6-8 GHz GaN LNAs are designed with the aim of providing as good linearity as possible while keeping the DC power consumption low, and some of the design considerations and trade-offs are analyzed and discussed. In a first design, traditional design techniques are used but the design is optimized towards high OIP3. When biased with 20 V and 75 mA, the amplifier achieves an output third order intercept of

Fig. 1. Microphotographs of LNA 1 (a) and LNA 2 (b). The two chips measure 1.6 mm × 1.2 mm and 1.5 mm × 1.1 mm respectively.

43 dBm at 7 GHz. The optimum OIP3 to DC power ratio of 12 is achieved at 10 V and 63 mA bias, with an OIP3 of 39 dBm.

In a second design, the linearity is enhanced by adding capacitive feedback between the input and output and by mismatching the input. This improves OIP3 by 1.5-2 dB at 7 GHz for all bias points compared to the first LNA, leading to a maximum OIP3 of 44 dBm at 75 mA and 20 V bias and 41 dBm at 63 mA and 10 V bias. The optimum OIP3 to DC power consumption ratio of 19 for the latter bias point is among the highest reported for GaN LNAs.

II. DESIGN APPROACH

The circuits were designed in the TriQuint 3MI GaN on SiC MMIC process with a gate length of 0.25 μm. Both amplifiers consist of a single stage based on four finger transistors with 100 μm unit gate widths.

A. Linearity figures of merit

We have chosen to design for maximizing the third order intercept point at the output, OIP3, rather than at the input, IIP3. This choice has important impact on the design approach. For example, reflections or losses at the input directly improve IIP3 but leave OIP3 unchanged while losses at the output degrade OIP3 but not IIP3. Also, when using negative feedback, IIP3 is improved three times more than OIP3 (in dB) because of the closed-loop gain reduction.

Which of these figures of merit is more relevant depends on the application. For the first LNA stage in a receiver, the IIP3 to noise figure ratio is an appropriate choice.
measure of the dynamic range. On the other hand, when designing a driver amplifier or the last stage in a multistage LNA which is required to deliver a certain amount of gain or output power, optimizing towards OIP\textsubscript{3} is probably a good choice. To obtain a figure of merit independent of scaling the transistor size and to enable a fair comparison between different technologies, the ratio of OIP\textsubscript{3} to DC power consumption is used as a figure of merit and as a design goal in this paper.

B. Large signal model

To simulate IMD, a large signal Angelov model is used [10]. For weakly nonlinear systems, models based on Volterra series up to the third order should in theory be better suited to predict IMD and can also be used to analyze the importance of different nonlinearities in the device. However, some of the coefficients in the Taylor series expansions of the device nonlinearities are very difficult to determine from only DC and S-parameter data. This is even more problematic when using Gallium Nitride devices which suffer from significant self heating and low frequency dispersion. We have found that a large signal model which is verified and optimized using large signal measurements is more accurate in predicting OIP\textsubscript{3} than a Volterra model determined from small signal and DC measurements. More advanced characterization methods using multitone excitations to find the Taylor coefficients [11] could improve the model and facilitate amplifier design.

C. General design considerations

A first amplifier (LNA 1) was designed using traditional LNA design techniques, but focusing on high linearity rather than low noise. Inductive source degeneration was employed to ensure in-band unconditional stability, enhanced linearity and noise matching. In order to boost OIP\textsubscript{3} by increasing negative feedback in the design frequency band, more inductance was added than would be the case in a pure low noise design. This leads to potential instability at frequencies slightly higher than the design frequency band. To stabilize, a series LC resonant circuit in parallel with a resistance was added at the transistor input. The series resonance is designed to short circuit the resistance in the design frequency band. The load impedance was then chosen for maximum OIP\textsubscript{3} while the chosen $\Gamma_S$ is a trade-off between input match, gain and linearity.

A second LNA (LNA 2) was designed using capacitive feedback between the output and the input of the transistor, increasing the OIP\textsubscript{3}. To avoid instability at high frequencies, a resistance was added in series with the feedback capacitance. The RC feedback also helps stabilize the transistor, eliminating the need for the stabilization circuit necessary in LNA 1. To further boost the OIP\textsubscript{3} of LNA 2, the input matching network was also chosen so as to increase the negative feedback, improving OIP\textsubscript{3} at the expense of gain and input return loss. Photographs of the two amplifiers are shown in Fig. 1.

D. Input matching network

The choice of input matching at the fundamental frequency may affect OIP\textsubscript{3} either because of the nonlinearity of the gate capacitance or through modifying the feedback mechanisms in the amplifier. For this

![Fig. 2. Contours of OIP\textsubscript{3} and gain in the $\Gamma_S$ plane at 7 GHz, for transistor biased with 10 V and 63 mA, with inductive degeneration and the output matched for maximum simulated OIP\textsubscript{3}. $\Gamma_S$ of the designed input matching network between 6 and 8 GHz is also shown in the figure. (a) shows contours with no added RC feedback and (b) is with RC feedback.](image1)

![Fig. 3. Contours of OIP\textsubscript{3} and gain in the $\Gamma_S$ plane at 7 GHz, for transistor biased with 10 V and 63 mA, with inductive degeneration and the input matched as indicated in Fig. 1. $\Gamma_L$ of the designed output matching network between 6 and 8 GHz is also shown in the figure. (a) shows contours with no added RC feedback and (b) is with RC feedback.](image2)
transistor in the design frequency band, the contribution from nonlinear capacitance seems to be negligible while the extent of negative feedback depends significantly on $\Gamma_L$.

Simulations show that with inductive source degeneration, the OIP$^3$ is maximized when the input is matched with a reactance of approximately $30 \, \Omega$, very close to the value optimizing gain and input return loss. Concerning the resistive part however, there is a trade-off between gain, input return loss and OIP$^3$, which improves for decreasing resistance (Fig 2a). The maximum OIP$^3$, achieved with an almost purely inductive match, is $3 \, \text{dB}$ higher than the figure that would be obtained with a conjugate input match. A compromise allowing for a return loss of -8 dB would increase OIP$^3$ by 1 dB while the gain is decreased by the same amount. In the first design, the input is matched with -15 dB return loss, resulting in 0.5 dB improvement of OIP$^3$ compared to the conjugate match.

E. Output matching network

On the output side, the choice of load impedance for maximum OIP$^3$ is highly dependent on the characteristics of the device nonlinearities. If the current has significant nonlinearities in its dependence not only of the gate-source voltage but also of the drain-source voltage, there may be several terms with important contributions to the IMD components.

As these depend on the output matching networks in different ways, small variations in the transistor characteristics around the bias point may then alter the optimum load, and the design may be very sensitive to changes in bias and inaccuracy in the transistor model. According to the large signal transistor model used in simulations, the OIP$^3$ contours in the $\Gamma_L$ plane are highly dependent on drain voltage. At 10 V bias, the OIP$^3$ contours move very close to the gain contours for the case with only inductive feedback. Adding RC feedback changes the optimum $\Gamma_L$ only very little (Fig. 3b).

III. RESULTS

The small signal and noise performance of the two circuits can be studied in Fig. 4. Noise figure has been measured using a noise figure analyzer, and was found to be 2-2.5 dB in the 6-8 GHz band for both amplifiers. An isolator was used on the amplifier output and the input and output mismatch of the amplifiers have been corrected for but some ripple still remains because of mismatch factors in the on-wafer measurement setup. In general, the large signal model predicts the small signal performance well.

OIP$^3$ was measured using two tone measurements at several power levels, with a tone spacing of 2 MHz. The results at 7 GHz are shown in Fig. 5 (a) for different bias. Agreement between measurements and the large signal model is good overall, but for high $V_{DS}$ and high current, the model overestimates OIP$^3$. For all drain bias, OIP$^3$ increases quickly with current up to around 50 mA and saturates below 100 mA. Between 10 V and 20 V, the OIP$^3$ increases almost proportionally to drain bias.

Fig. 4. Measured (solid lines) and simulated (dashed lines) small signal and noise performance of the two LNA’s.

Fig. 5 (a) Measured (LNA 1: o, LNA 2: x) and simulated OIP$^3$ (LNA 1: solid line, LNA 2: dashed line). (b) Measured OIP$^3$/PDC of LNA 1 (solid line) and LNA 2 (dashed line).
have been found to differ significantly from that at GaN devices, the I-V characteristics at RF frequencies are more important than the device S-parameters suggest. In many respects to the drain-source voltage could be more important than the device S-parameters suggest. This indicates that nonlinearities with respect to the dc-drain-source voltage could be more important than the device S-parameters suggest. In many GaN devices, the I-V characteristics at RF frequencies have been found to differ significantly from that at DC.

### IV. DISCUSSION

In general, GaN seems to offer linearity performance comparable to GaAs for the same amount of DC power consumption. The advantage of GaN becomes significant only when the linearity requirements become so high that scaling of GaAs devices would require impractical device sizes or current levels.

As GaN HEMTs have lower gain than many other technologies, improving the linearity using negative feedback is more difficult when working in GaN, as negative feedback always sacrifices gain. Also techniques relying on cancellation of several nonlinear components are difficult to implement in GaN, as very accurate knowledge of the device nonlinearities is required.

The fact that the maximum OIP<sub>3</sub> occurs at a current level as low as approximately 80 mA, around 20% of the maximum current level, is surprising considering the bias dependence of $g_{m}$ and $g_{ds}$, extracted from device S-parameter measurements. According to the Volterra model extracted from these measurements, the linearity should increase for currents above 100 mA due to a rapid decrease of $g_{ds}$, the second order derivative of $g_{m}$. The fact that this does not occur indicates that nonlinearities with respect to the drain-source voltage could be more important than the device S-parameters suggest. In many GaN devices, the I-V characteristics at RF frequencies have been found to differ significantly from that at DC.

### REFERENCES


### TABLE I

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[3]</td>
<td>10</td>
<td>5</td>
<td>0.8</td>
<td>46</td>
<td>6</td>
<td>6.6</td>
</tr>
<tr>
<td>[4]</td>
<td>10</td>
<td>7</td>
<td>1.45</td>
<td>24</td>
<td>0.15</td>
<td>1.7</td>
</tr>
<tr>
<td>[5]</td>
<td>18</td>
<td>2</td>
<td>1.8</td>
<td>54</td>
<td>24</td>
<td>11</td>
</tr>
<tr>
<td>[6]</td>
<td>11</td>
<td>6</td>
<td>1.6</td>
<td>24</td>
<td>0.12</td>
<td>2.1</td>
</tr>
<tr>
<td>[7]</td>
<td>20</td>
<td>2</td>
<td>3</td>
<td>52</td>
<td>30</td>
<td>5.3</td>
</tr>
<tr>
<td>[8]</td>
<td>15</td>
<td>1.8</td>
<td>1.5</td>
<td>49</td>
<td>16.5</td>
<td>4.8</td>
</tr>
<tr>
<td>[9]</td>
<td>15</td>
<td>12</td>
<td>1.2</td>
<td>30</td>
<td>0.06</td>
<td>17</td>
</tr>
</tbody>
</table>

LNA1  10  7  1.7  39  0.63  12  LNA2  7  7  2.6  41  0.65  19

The maximum OIP<sub>3</sub> occurs at a current level as low as approximately 80 mA, around 20% of the maximum current level, but it is almost independent of drain voltage for low currents.

OIP<sub>3</sub>/P<sub>dc</sub> is shown in Fig. 5 (b). The optimum OIP<sub>3</sub>/P<sub>dc</sub> of 19 is achieved at 63 mA and 10 V bias and is among the highest reported for Gallium Nitride amplifiers. Table I shows a comparison to some other published GaN LNA's and to commercial GaAs pHEMTs.
Paper E

Noise Temperature of an Electronic Tuner for Noise Parameter Measurement Systems


79th ARFTG Measurement Conference, Montreal, Canada, June 2012
Noise Temperature of an Electronic Tuner for Noise Parameter Measurement Systems

Olle Axelsson¹, Mattias Thorsell¹, Kristoffer Andersson¹, Jörgen Stenarson² and Yves Rolain³

¹Microwave Electronics Laboratory, Chalmers University of Technology, 412 96, Göteborg, Sweden
²SP Technical Research Institute of Sweden 501 15, Borås, Sweden
³Department ELEC, Vrije Universiteit Brussel, B-1050 Brussels, Belgium

Abstract—In this paper, the noise temperature of an electronic tuner is determined and its significance for the suitability of such tuners in noise parameter measurement systems discussed. The noise temperature of the tuner was found to be higher than the ambient room temperature in the laboratory and vary significantly between tuner states. For impedance states with small input reflections coefficients, the excess noise temperature is around 25 K. For some of the states with higher reflection coefficients, this figure increases, reaching around 45 K at \( |\Gamma| = 0.75 \). Unless accounted for, this leads to errors in noise parameter extraction when using an electronic tuner in noise parameter measurements.

I. INTRODUCTION

Measurement of noise parameters of active devices are commonly done by measuring the noise power at the output of the device under test (DUT) for several input terminations. As this method relies on the knowledge of the equivalent noise temperature of the input termination seen by the DUT, realizing these terminations using passive loads or mechanical tuners at room temperature provide the most accurate results. Nevertheless, in commercial systems electronic tuners are often used because of their speed and ease of use [1].

As these dissipate DC power and are not passive components, the noise temperature of the electronic tuner is not necessarily equal to the ambient room temperature, even though this is sometimes assumed. Both the temperature increase due to heat dissipated in the p-i-n diodes of the tuner and the noise current in the diodes themselves can potentially lead to excess noise in the tuner. As inaccurate knowledge of the input noise temperature leads to inaccuracy in the measured DUT noise parameters, characterization of the tuner noise is crucial for evaluating as well as improving accuracy of noise parameter measurement systems utilizing electronic tuners.

In this paper, the noise temperature of an electronic tuner is measured and is found to differ significantly from the ambient room temperature. Moreover, the excess noise is different for different impedance states, with most of the states presenting high reflection coefficients being noisier than states close to 50 \( \Omega \).

II. MEASUREMENTS

A. Setup and Procedure

The tuner under test is an ATN electronic tuner, used in a commercial noise parameter measurement system [1]. During the characterization, the tuner was used as a one port with the input port terminated in a reflectionless load, similar to the way it is used during noise parameter measurements. The 353 impedance states of this tuner are shown in the Smith Chart in Fig. 1. All measurements were done at 10 GHz.

The available noise power from the tuner is measured using the setup shown in Fig. 2. The noise power is measured using an Agilent N8975A Noise Figure Analyzer (NFA). Two isolators in cascade provide 54 dB of isolation between the tuner and the NFA to ensure the added noise and the gain of the NFA are independent of tuner input impedance.

In order to determine the tuner noise temperature \( T_i \) for each tuner state \( i \), we need to measure the noise powers \( P_i \) with the NFA when the tuner is connected as in Fig. 1. We also measure hot and cold state noise powers \( P_h \) and \( P_c \) when the tuner is replaced by a noise source with a known ENR value.
In addition, we need the input reflection coefficient of the tuner states, $\Gamma_i$, as well as of the isolator, $\Gamma_R$, and the noise source, $\Gamma_{NS}$. The measured noise power in Fig. 2 can be written:

$$P_t = kG\Delta f \left( \alpha(\Gamma_i)T_i + \left(1 - \alpha(\Gamma_i)\right)T_a + T_R \right)$$

with $G$ being the transducer power gain of the NFA measurement receiver, $\alpha(\Gamma_i)$ is the available gain of the isolator, $T_R$ the noise temperature of the NFA and $T_a$ the temperature of the isolator, assumed to be equal to the ambient temperature in the laboratory. Assuming $\Gamma_{NS}$ to be equal in the hot and the cold state, we can determine $G$ and $T_R$ from (1) using the result of the hot-cold measurement:

$$kG\Delta f = \frac{(P_h - P_c)}{\alpha(\Gamma_{NS})(T_h - T_a)}$$

$$T_R = \frac{P_c}{kG\Delta f} - T_a$$

Inserting (2) and (3) into (1) and calculating $\alpha(\Gamma_{NS})$ and $\alpha(\Gamma_i)$ from the isolator s-parameters assuming perfect isolation, we end up with the expression for $T_i$:

$$T_i = T_a + \frac{P_t - P_c}{P_h - P_c} \left(1 - |\Gamma_{NS}|^2\right) \left(1 - |\Gamma_i|^2\right) (T_h - T_a)$$

In order to have some statistics of the results and get an idea of the uncertainties and repeatability of measured results, the measurement series of $P_t$ for all tuner states was repeated 15 times, using the same values for the other parameters when calculating the noise temperature.

B. Results

The measured tuner noise temperatures for three measurements of all 353 tuner states are shown in Fig. 3, plotted versus $|\Gamma|$.

The noise temperature varies between approximately 315 K and 350 K, 25-55 K above the ambient room temperature of 295 K. There is a clear trend towards higher noise temperatures for states with higher reflection coefficients, starting at around 320 K at low $|\Gamma|$ and increasing to approximately 340 K at $|\Gamma| = 0.75$. However, the two outermost states on each branch (marked in red in Fig. 1 and with rings in Fig. 3) deviate from this trend and present noise temperatures similar to that of the low reflection states.

The uncertainty of the measurement seems to be dominated by drift in the gain of the measurement receiver. Even a very small drift between the measurement of noise power from the cold noise source and the tuner leads to a significant error in the factor $P_t - P_c$ in, as the differences in noise power are very small compared to the total noise power. The calculated noise temperatures of the same states vary by around 10 K between different measurement series, even though the change in measured noise power is as small as 0.02 dB.

III. CONCLUSIONS

We have measured the equivalent noise temperature of an electronic tuner and found that it significantly exceeds the ambient room temperature and is dependent of the impedance state of the tuner. Unless this is accounted for in the extraction of noise parameters when using electronic tuners as variable impedances in noise parameter measurement systems, this will lead to inaccuracy in the extracted parameters.

REFERENCES