July 2017

DYNAMIC RANGE LIMITATIONS OF LOW-NOISE MICROWAVE TRANSISTORS AT CRYOGENIC TEMPERATURES

Ahmet Hakan Coskun

DYNAMIC RANGE LIMITATIONS OF LOW-NOISE MICROWAVE TRANSISTORS AT CRYOGENIC TEMPERATURES

A Dissertation Presented

by

AHMET HAKAN COŞKUN

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2017

Electrical and Computer Engineering

DYNAMIC RANGE LIMITATIONS OF LOW-NOISE MICROWAVE TRANSISTORS AT CRYOGENIC TEMPERATURES

A Dissertation Presented

by

AHMET HAKAN COŞKUN

Approved as to style and content by:

Joseph C. Bardin, Chair

Robert W. Jackson, Member

K. Sigfrid Yngvesson, Member

Gopal Narayanan, Member

Christopher V. Hollot, Department Head
Electrical and Computer Engineering

ABSTRACT

DYNAMIC RANGE LIMITATIONS OF LOW-NOISE MICROWAVE TRANSISTORS AT CRYOGENIC TEMPERATURES

MAY 2017

AHMET HAKAN COŞKUN

B.Sc., YEDITEPE UNIVERSITY

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Joseph C. Bardin

Dynamic range is an important metric that specifies the limits of input signal amplitude for the ideal operation of a given receiver. The low end of dynamic range is defined by the noise floor whereas the upper limit is determined by large-signal distortion. While dynamic range can be predicted in the temperature range where compact transistor models are valid, the lack of large-signal models at temperatures below -55 C prevents the prediction and optimization of dynamic range for applications that require cryogenic cooling. For decades, the main goal concerning the performance of these applications was lowering the noise floor of cryogenic receiver front-ends. For this, linear small-signal noise models have been extensively studied and used for designs of low-noise amplifiers.

In this work, the existing small-signal noise modeling approach is extended to capture the weakly nonlinear properties of the transistors that are commonly used in cryogenic amplification. Indium phosphide high electron mobility transistors and

silicon germanium heterojunction bipolar transistors are considered. The goal of this work is to identify the fundamental dynamic range limitations of these transistors such that the results are not device specific, but applicable to the corresponding device families.

Identifying the fundamental limitations of dynamic range in a semiconductor device requires a broad understanding of physical properties of the transistors. For this, a theoretical analysis will be presented first as a function of temperature. The small-signal noise modeling will then be discussed using techniques that are well recognized in the literature. This will be followed by an explanation of the nonlinear modeling approach used in this work. This approach relies on the definition of Taylor series expansion coefficients of the dominant nonlinear mechanisms of the transistors. The modeling results will be interpreted with respect to the initially presented theoretical framework. Finally, the dynamic range performance will be studied as a function of source and load terminations. In addition to this systematic approach to understanding the physical limitations of dynamic range, model to measurement agreement of broadband cryogenic amplifiers will also be presented which will verify the accuracy of the modeling approach.

ACKNOWLEDGMENTS

I sincerely thank Professor Joseph Bardin for his support and guidance throughout this experience. This project would not have existed without his unique perspective in the first place. As much as his supervision, I benefited greatly from his earlier research, which will be evident throughout this dissertation. I consider myself privileged to be his first PhD graduate. I would also like to thank my committee members, Professors Robert Jackson, Sigfrid Yngvesson, and Gopal Narayanan, for their invaluable feedback.

I would like to thank my colleagues in Umass Amherst Radio Frequency Nanoelectronics Group, Mohammad Ghadiri Sadrabadi, Prasana Ravindran, Ajay Subramanian, Maruthi Yogeesh, Metin Ayata, Matt Conte, Su-Wei Chang, Shirin Montazeri, Wei-Ting Wong, Randy Kwende, and Mohsen Hosseini, for the great times we had together inside and outside of the university.

I would like to acknowledge Wei-Ting Wong for developing the proof of concept wideband high-linearity SiGe HBT amplifier, using the models presented in this work. I would also like to acknowledge Sander Weinreb for providing the InP HEMT amplifiers. These material enabled a strong validation of this work.

This work was funded by the Office of Naval Research. Any opinions, findings,and conclusions or recommendations expressed in this thesis are those of the author and do not necessarily reflect the views of ONR.

Finally, I would like to express my gratitude to my family, for their unconditional moral support.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

CHAPTER 1

INTRODUCTION

This dissertation is dedicated to understanding the limits to the dynamic range of transistors in the temperature range of 7 - 300 K. While dynamic range may be adequately discussed in the literature for room temperature operation, due to the lack of large-signal models, it is often not taken into account for applications that require cryogenic cooling at the circuit design level. Therefore the focus will be on the low temperature behavior and the changes observed with cooling.

Dynamic range is constrained by two fundamental phenomenons of solid-state devices; noise and distortion. To predict and evaluate dynamic range of a given transistor technology, the noise and non-linear behavior of the device needs to be understood and accurately modeled. The contribution of this thesis is the modeling, physical explanation, and circuit level optimization of these two limiting properties as a function of temperature for indium phosphide high electron mobility transistors (InP HEMT) and silicon germanium heterojunction bipolar transistors (SiGe HBT). An outline of the thesis is given below.

Theory: This section will be devoted to understanding fundamental physical properties of the device technologies that are subject to this research. DC transport mechanisms of each device will be discussed first. The discussion will then be expanded to include dynamic effects such as capacitances and time delays. With the knowledge of the transport properties, the mechanisms that limit the noise and linearity performance of each device will be described along with their predicted temperature dependence.

Modeling: Models that represent the physical operation are required to understand transistor characteristics and replicate them in a simulator for circuit designs. To characterize both ends of the dynamic range, a model that can predict both noise and weak nonlinearities is required. Following a description of the devices that are modeled in this work as well as the measurement test setup, the small-signal noise models will be presented. These models will then be expanded to capture nonlinear behavior of the devices under weak drive. The temperature dependence of the model parameters will be evaluated based upon the device physics.

Analysis: The noise and linearity performance of transistors depends upon the input and output termination impedance. Therefore, maximizing the linearity and dynamic range performance through optimization of the in-band and out-of-band impedance terminations will be studied. Two practical cases will be considered which correspond to designs of a narrowband and a wideband amplifier. Ultimately, the performance of each technology platform will be compared.

Applications: This section is dedicated to amplifier measurement results, which provide verification of the weakly nonlinear models that are presented and studied earlier in the thesis. The ultimate goal is to demonstrate that the linearity of a cryogenic amplifier can be predicted and optimized during the design process with the modeling approach presented in this work.

Aside from a detailed dynamic range analysis of InP HEMTs and SiGe HBTs, the noise analysis of alternative technologies such as short-channel MOSFET and compound-semiconductor HBT at cryogenic temperatures is also provided in this dissertation. The background and motivation of this work is explained in the rest of this chapter.

1.1 Background of Cryogenically Cooled High Sensitivity Receivers

Cryogenically cooled LNAs are required for applications where weak signals in a low background noise environment need to be recovered. By reducing the operating temperature of an LNA, the contribution of the LNA itself to the noise floor is greatly reduced, therefore improving receiver sensitivity. This improvement is attributed to both the reduction in thermal noise associated with passive losses as well as enhanced transport properties of semiconductors.

The idea of cooling a transistor to improve its noise performance dates back to 1980 [1], when it was realized that a gallium arsenide (GaAs) metal–semiconductor field-effect transistor (MESFET) cooled down to 20 K can provide noise performance close to what had been achieved by parametric amplifiers and masers. Since then, cryogenically cooled transistors have become an important part of systems used in deep space communication and radio astronomy in the form of front-end and IF amplifiers [2–5]. Other scientific applications include read-out circuits for microwave quantum computing [6, 7] and interface elements for various experiments in fundamental physics [8, 9]. Cryogenic LNAs have also been considered for military and commercial applications such as superconducting digital receivers [10], magnetic resonance imaging [11], and mobile base stations [12].

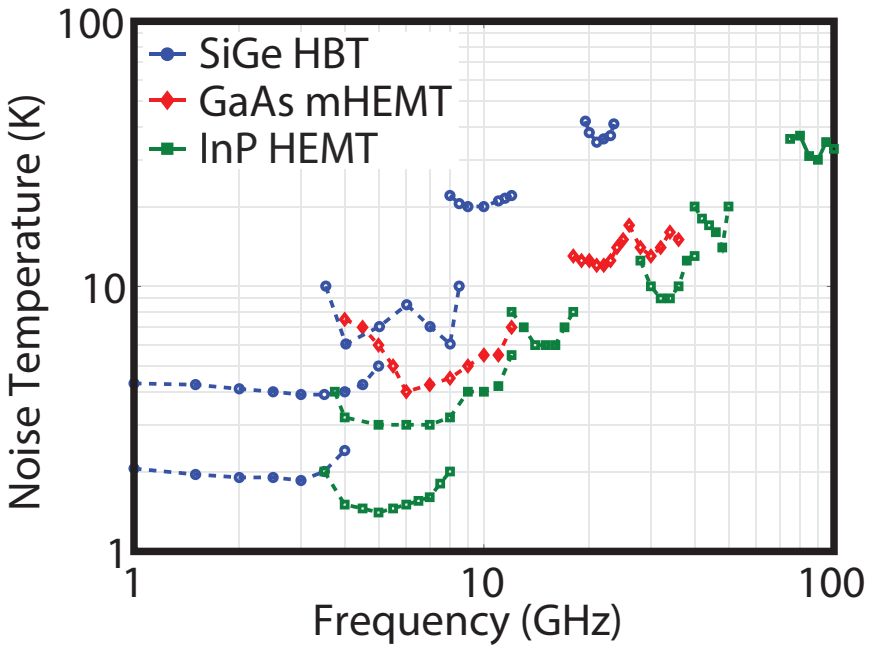
For a long period of time, cryogenic LNA designs were based solely on III-V technologies (e.g., GaAs and InP MESFETs and HEMTs) [13]. The SiGe HBT is an alternative technology that provides competitive noise performance at cryogenic temperatures. However, the first working example of a SiGe HBT was only available in 1987 [14], thirty years after its operating principle was first described in 1957 [15]. In the mean time, silicon bipolar transistors were neither as fast nor suitable for cryogenic cooling [16]. As agressive CMOS scaling due to the demand from digital applications reflected on the BiCMOS platform, the high frequency performance of

SiGe HBTs has become competitive with those of compound-semiconductor HEMT devices. The DOTSEVEN program aims to push $f_{\max}$ of SiGe HBTs to 700 GHz [17]. This goal has recently been met with a 0.1 μm SiGe HBT with $f_{\mathrm{t}}/f_{\max}$ values of 505/720 GHz [18]. For reference, state-of-the-art 30 nm InP HEMTs [19] and 20 nm GaAs mHEMTs [20] achieve $f_{\max}$ greater than 1 THz.

In addition to the ever increasing high frequency limits at room temperatue, significant improvements in small-signal and noise performance of SiGe HBT devices are observed with cooling [21]. Consequently, SiGe amplifiers with noise temperatures less than 6 K were reported within the sub 10 GHz frequency range [22–24]. Figure 1.1 shows noise performance of state of the art cryogenic LNAs found in the literature and their corresponding platform. The noise performance of a GaAs mHEMT is slightly worse than that of a InP HEMT due to the extra metamorphic buffer layer[1], which degrades the carrier confinement as a result of surface defects [25].

InP HEMTs provide better noise performance especially for frequencies beyond 10 GHz, which can be partially explained by the $f_{\max}$ difference of both technologies. However, there are multiple issues associated with HEMT processes that make SiGe HBTs a competitive candidate for lower frequency applications. HBTs have vertical structure. Therefore they do not suffer from the gain fluctuations that are observed in HEMT devices due to the surface trapping [35], which ultimately limits the radiometer performance of HEMT based amplifiers. Surface trapping also results in a high 1/f noise corner frequency of HEMTs, which makes broadband designs below 2 GHz challenging. Furthermore, the design of broadband matching networks in this frequency range is more complicated for HEMTs, as their input impedance is higher than that of HBTs. Wafer level variations are another issue with HEMT devices, as significant fluctuations of cryogenic noise performance are observed within a reticule [36]. On

[1]In GaAs mHEMTs, a metamorphic transition layer is required to be able to grow the InAlAs buffer layer on top of the GaAs substrate.



(a)

Figure 1.1: Noise performance of cryogenic LNAs. The results are compiled from [2, 13, 26–34].

the other hand, SiGe HBT process rely on high-yield CMOS fabrication methods. Another advantage of SiGe HBT is their low power consumption. Recently, it has been demonstrated that by reducing the collector voltage of HBT devices, LNAs with power consumption on the order of a couple hundred micro-watts can be realized [37]. This is very important for applications where thousands of cryogenic amplifiers are needed, as low DC power consumption will reduce the heat lift requirement for the cryocooler.

1.2 Dynamic Range Considerations

1.2.1 Room Temperature

The availability of large-signal models at room temperature enables the prediction and optimization of dynamic range performance for a given receiver. Furthermore, using weakly nonlinear small-signal models, researchers have also adequately investi-

gated the fundamental linearity limitations of HEMTs and HBTs at room temperature. These modeling approaches will be further explained in Chapter 6. Regarding the amplifier design, several different intermodulation cancellation techniques have been studied in the literature, which will be further reviewed in Chapter 7.

In addition to circuit and device level efforts concerning the intermodulation performance of the front-end amplifier, system level approaches have also been considered to enhance receiver linearity. As an example, frequency selective reconfigurable receiver architectures block interference by creating a sharp filtering response [38, 39]. For wideband receivers, harmonic rejection mixers have been demonstrated to be useful for eliminating third and fifth order distortion [40, 41].

1.2.2 Cryogenic Temperatures

Although InP HEMT and SiGe HBT technologies have been studied extensively in the literature in terms of linear small-signal and noise performance at cryogenic temperatures, little effort has been put into understanding the nonlinear behavior of semiconductor devices at these temperatures. To the best of the author's knowledge, the only articles that have been published to understand the cryogenic linearity of SiGe HBTs are [42, 43], where performance has only been reported down to 77 K. No analysis of HEMT linearity performance at low temperatures has been found in the literature.

The insufficient amount of research on this matter is not surprising. Receiver systems that involve cryogenic cooling are mostly expected to operate linearly, since they are assumed to be handling weak signals most of the time. Therefore, most of the research in cryogenic devices and amplifiers has focused on increasing the sensitivity of the receiver by lowering the noise temperature. However, linearity is becoming a more important metric as the frequency spectrum has become more populated and the likelihood of unwanted signals to appear in-band increases. Consequentially, the prob-

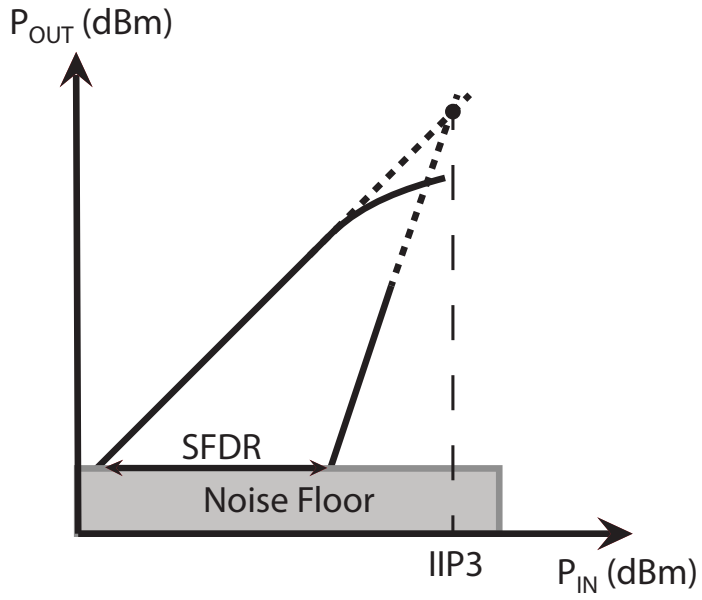
ability of intermodulation distortion (IMD) and gain compression increases depending on the linearity performance of the receiver. These non-linear distortion mechanisms are not desired for receivers, as they can interrupt communication links or reduce the accuracy of radiometer measurements. As technological advances in antenna designs and device fabrication enable higher bandwidth systems [44], nonlinearity has become an elevated concern. A wider spectrum is expected to be more populated, containing more spurious signals from the radiation sources of other applications. Interaction of these random signals with each other and with the signal of interest will create more IMD products in-band than what would exist in a narrow-band application.

In order for a receiver to operate properly, the incoming signal should have an amplitude above the noise floor and below the point where it generates detectable distortion. This range of useful signal level, also called as the *spurious-free dynamic range* (SFDR), is illustrated in Figure 1.2 and can be expressed in terms of the system noise temperature (T_{SYS}) and the input third order intercept point (IIP3) as

$$SFDR(\mathrm{dB}) = \frac{2}{3}\left[IIP_3(\mathrm{dBm}) - 10\log_{10}\left(kT_{\mathrm{SYS}}B\right) - 30\mathrm{dB}\right], \tag{1.1}$$

where k is the Boltzmann's constant and B is the system bandwidth. It should be noted that for applications that have bandwidths greater than 2:1, second order distortion is also as critical, since the resulting second order products fall within the band. Therefore, it must also be ensured that the SFDR is not limited by second order distortion.

As mentioned earlier, the likelihood of radio-frequency interference (RFI) is proportional to the receiver bandwidth. Depending on the linearity, RFI might necessitate reducing the instantaneous bandwidth. In this case, a higher integration time is required in order to maintain the specified receiver output fluctuation according to the radiometer equation, which is given as [45]



(a)

Figure 1.2: Spurious-free dynamic range.

$$\Delta T \approx \frac{T_{\mathrm{SYS}}}{\sqrt{B \times \tau_{\mathrm{int}}}}, \tag{1.2}$$

where ΔT is the receiver output fluctuation and τ_{int} is the integration time. Therefore, it is clear that a high dynamic range receiver requires less time to perform a continuum observation.

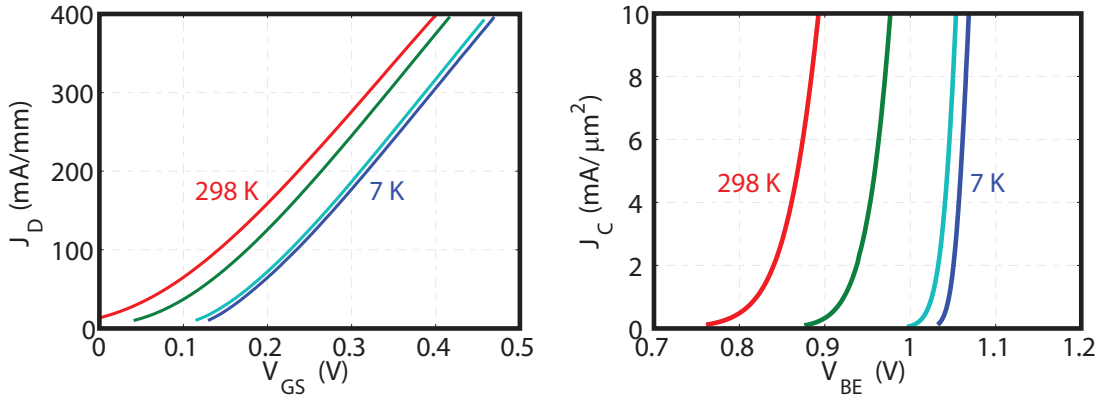
While high linearity receivers enable better RFI mitigation for continuum measurements, the quality of spectral line observations also correlates to the receiver dynamic range. In this case, the power ratio between the brightest and weakest feature in a radio image defines the dynamic range [46]. Therefore, maximizing the linearity of a receiver while minimizing its noise contribution is critical for obtaining high dynamic range images.

Aside from conventional receivers, there are emerging low-temperature applications that require linear amplification of a wide dynamic range signal. Microwave kinetic inductance detector (MKID) based cameras are one example. MKIDs relies

on the inductance of a superconducting photon detector which is effectively used in an LC resonator. Once a photon is incident on the detector, the kinetic inductance changes which results in a shift in the resonator frequency [47]. Many detectors with unique resonant frequencies can be multiplexed through a single cryogenic amplifier, given that the amplifier can handle the total excitation power used to sense the detector array [48]. Hence, high linearity cryogenic LNAs are required to read-out large arrays of MKIDs.

While cryogenic cooling of the receiver front-end LNA greatly reduces its noise contribution, changes in the amplifier linearity are not that clear. As a first order approximation, the linearity of a transistor is correlated to sharpness of its I-V curves. As a device is cooled down, the I-V curves get sharper, which is shown in Figure 1.3 for an InP HEMT and a SiGe HBT. The sharpness seems to increase significantly more with cooling for the HBT than it does for the HEMT, which indicates that SiGe HBTs may be more nonlinear at cryogenic temperatures. Thus, it is important to understand whether any of the competing technologies perform significantly better in terms of the IMD performance at cryogenic temperatures. Understanding this property will provide additional guidance when choosing the suitable platform for future applications and will also help with optimizing the future designs in terms of SFDR performance.

The expected increase in the transistor nonlinearity with cooling is captured by the reported linearity results of commercial amplifiers which are shown in Table 1.1. However, the results reported in [42] indicate that the output IP3 of a modern SiGe HBT increases upon cooling if the source and load terminations are optimized for maximum output power at a given temperature. This contradiction indicates that the impedance terminations play a significant role when determining the impact of intrinsic transistor nonlinearity on amplifier performance. While a traditional LNA design relies on matching the output of a transistor for maximum power transfer to



(a) (b)

Figure 1.3: (a) InP HEMT and (b) SiGe HBT DC I-V curves.

Table 1.1: Reported Linearity Results Of Commercial Cryogenic Amplifiers

Platform	Ref	Frequency Range	P_{DC}	IP_{1dB}	Gain
-	-	GHz	mW	dBm	dB
InP HEMT	[49]	4 - 8	56/4	-42/-51	42/39
InP HEMT	[50]	16 - 28	60/4	-32/-44	27/32
SiGe HBT	[51]	0.5 - 3	63/9	-29/-37	30/31

Room Temperature / Cryogenic

the load, it could be possible to follow a different approach in the design process to optimize the gain-linearity trade-off such that the degradation of IP3 with cooling is mitigated or prevented. Therefore, aside from a study of the intrinsic nonlinearities as a function of temperature, the impact of impedance terminations on linearity performance of the devices will also be presented in this work.

CHAPTER 2

DEVICE THEORY

The intrinsic transport properties of HEMTs and HBTs are investigated in this chapter. The band diagrams are analyzed first. The terminal currents are then explained in terms of semiconductor physics. Junction capacitances are discussed. Finally, based upon the initially presented theoretical framework, the mechanisms limiting the noise and linearity are studied.

2.1 High Electron Mobility Transistor

The high electron mobility transistor is based on the idea of creating a semiconducting region where the conduction band energy level (E_C) is below the Fermi level (E_F). This region is called a quantum well and is obtained by growing a large bandgap material (barrier) on top of a small bandgap material (e.g., AlGaAs on GaAs, InAlAs on InGaAs). As the quantum well provides the lowest energy state for electrons, it attracts electrons from the ionized impurities located in the doped large bandgap material. These electrons then form a two-dimensional electron gas (2DEG) at the heterojunction interface. This 2DEG serves as the channel of the HEMT and its carrier concentration is controlled by the gate voltage. Since the channel itself is undoped, high electron mobility is achieved in the channel due to the minimal ionized impurity scattering.

In this section, two distinct aspects of electron transport in HEMTs will be initially investigated. First, the channel formation as a function of gate potential will be presented. Second, the electron velocity as a function of drain-source electric field will

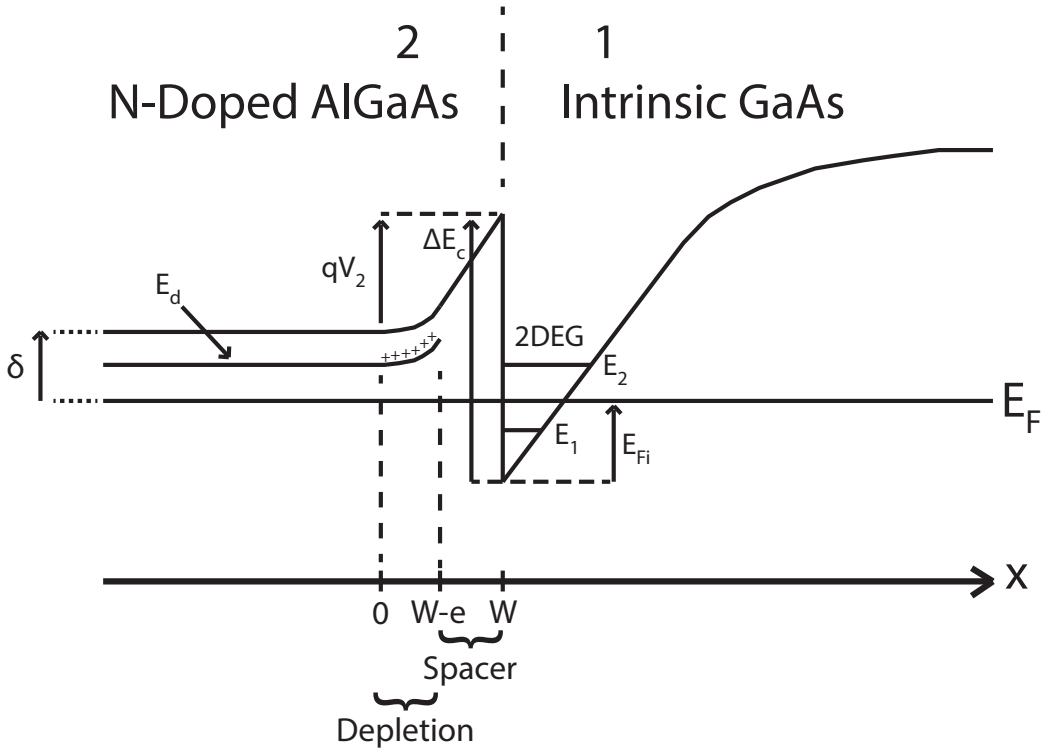


Figure 2.1: Conduction band diagram of AlGaAs - GaAs Heterostructure [53].

be discussed. Based on these fundamental properties, the I-V and C-V characteristics of HEMTs will be analyzed. Finally, the dominant sources of noise and distortion will be explained.

2.1.1 Equilibrium Concentration of the Quantum Well

The properties of the quantum well must be studied in order to understand HEMT operation. For a simple structure that only includes the large bandgap and the small bandgap material, electrons are inherently confined in the well, since the potential depth of the well is greater than the thermal voltage at room temperature [52]. Before moving on to the analysis of the complete HEMT structure, the quantum well carrier distribution of the simple heterostructure case will be explained.

A band diagram of a AlGaAs - GaAs heterostructure is shown in Figure 2.1. The AlGaAs layer is undoped near the heterojunction to ensure that the bulk mobility

is achieved in the channel. The heterojunction results in a depletion region in the AlGaAs layer, as the electrons are confined in the 2DEG. According to quantum mechanics, the electron distribution in a quantum well is partitioned into discrete sub-bands. The discrete energy levels of the first two subbands are denoted with E_1 and E_2. In [54], it was shown that the first two subbands correspond to 95% of the total electron concentration, which is assumed to be sufficient to describe the 2DEG system [53]. The total charge in the potential well can then be written as [55, 56]

$$n_{\mathrm{s}} = \frac{DkT}{q} \ln \left\{ (1 + e^{q/kT(E_{\mathrm{F}} - E_1)})(1 + e^{q/kT(E_{\mathrm{F}} - E_2)}) \right\}, \tag{2.1}$$

where D is the density of states given as

$$D = \frac{qm^*}{\pi\hbar^2}, \tag{2.2}$$

where m^* is the effective electron mass and $\hbar$ is the reduced Planck's constant. The subband energy levels E_1 and E_2 can be analytically obtained with the triangular well approximation as [57]

$$E_{\mathrm{i}} = \left(\frac{\hbar}{2m^*} \right)^{1/3} \left[3q^2 n_{\mathrm{s}} \left(i + \frac{3}{4} \right) /2 \right]^{2/3}. \tag{2.3}$$

Alternatively, E_1 and E_2 can be empirically defined as [56]

$$E_{\mathrm{i}} = \gamma_{\mathrm{i}} n_{\mathrm{s}}^{2/3}, \tag{2.4}$$

where γ is an experimentally obtained constant. The empirical approach is expected to be more accurate [53].

Equation (2.1) describes the charge as a function of subband energy levels with respect to the E_{F}. The equilibrium charge concentration is given as [56]

$$qn_{\mathrm{s0}} = \left\{2\epsilon_2 N_{\mathrm{D}} \left[\Delta E_{\mathrm{C}} - \delta - E_{\mathrm{Fi}}\left(n_{\mathrm{s}}\right)\right] + q^2 N_{\mathrm{D}}^2 e^2\right\}^{1/2} - qN_{\mathrm{D}}e, \tag{2.5}$$

where N_{D} is the doping concentration in the AlGaAs region, ϵ_2 is dielectric constant of the AlGaAs layer and e is the width of intrinsic AlGaAs region. The relationship between the n_{s} and E_{Fi} is transcendental. Therefore, a solution to the Equation (2.5) must be obtained numerically and must also satisfy Equation (2.1). The equilibrium charge is maximized for zero spacer (intrinsic AlGaAs region) width [56], however e is usually kept around 100 $\mathring{A}$ to minimize Coulombic scattering in the channel and to achieve high electron mobility [53]. The equilibrium concentration is proportional to N_{D}, but it saturates beyond a certain N_{D} that depends on the spacer width [58].

It should be noted that the AlGaAs region is assumed to be completely depleted for Equation (2.5). In the case of incomplete ionization, a more accurate expression for n_{s0} is reported in [59], where the effective position of 2DEG was also accounted for under the assumption that E_{Fi} varies linearly with n_{s}. This will be further discussed in the following analysis of 2DEG charge control.

2.1.2 Charge Control of the Two-Dimensional Electron Gas

Once the equilibrium properties of the quantum well are understood, the heterostructure can be studied in the presence of the gate metal contact. The band diagram for this case is shown in Figure 2.2. A depletion region arises at the interface of gate metal and barrier layer due to the Schottky junction formation. Another depletion region is present at the heterojunction interface. Depending on the thickness of the barrier layer, these two depletion regions overlap and the entire barrier layer is depleted. In this case, electrons provided by donors in the barrier layer are shared between the gate metal and the 2DEG [53]. This charge partitioning indicates

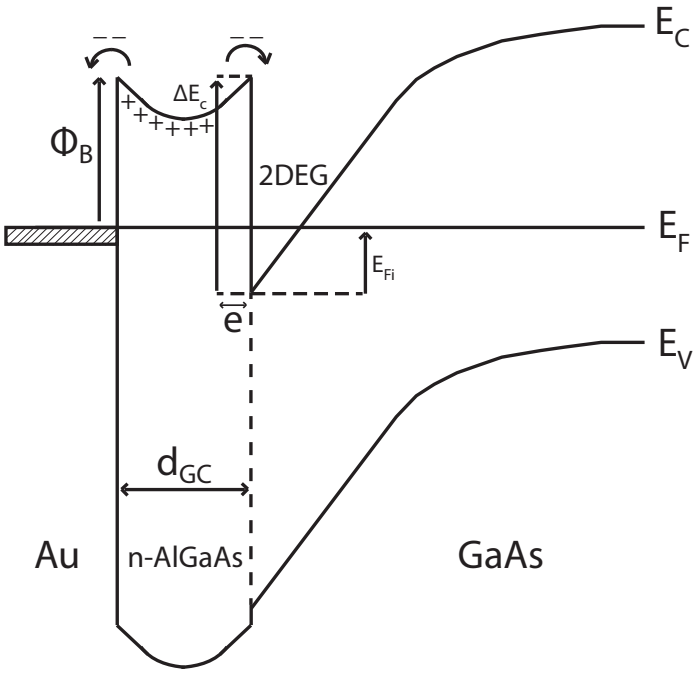


Figure 2.2: Equilibrium band diagram of the Au n - AlGaAs i - GaAs structure for a thin AlGaAs region [53].

that under equilibrium conditions with no gate bias, the device can already have a conducting channel.

A positive gate-channel voltage (V_{GC}) narrows the depletion region that covers the channel, which increases the 2DEG concentration. Once the equilibrium concentration (n_{s0}) is reached in the channel, the depletion region of the Schottky gate and 2DEG cease to overlap with any further increase in V_{GC}. At this point, the AlGaAs layer is no longer completely depleted, and a parasitic n-channel is formed. On the other hand, a negative gate voltage widens the depletion region, which deprives the channel of electrons.

For the case of a completely depleted AlGaAs layer, the relationship between the channel carrier concentration (n_{s}) and V_{GC} can be written as [52, 53]

$$n_{\mathrm{s}} = \frac{\epsilon_2}{q d_{\mathrm{GC}}} \left(V_{\mathrm{P}} - \phi_{\mathrm{B}} + V_{\mathrm{GC}} + \frac{\Delta E_{\mathrm{C}}}{q} - \frac{E_{\mathrm{Fi}}\left[n_{\mathrm{s}}\right]}{q} \right), \tag{2.6}$$

where V_{P} is the so-called pinch-off voltage given as

$$V_{\mathrm{P}} = \frac{1}{2} q \frac{N_{\mathrm{D}}}{\epsilon_2} \left(d_{\mathrm{GC}} - e\right)^2, \tag{2.7}$$

where d_{gc} is the thickness of barrier layer and e is the width of undoped barrier region. If the transcendental relationship between the E_{Fi} and n_{s} is neglected and a constant value of E_{F} is assumed, Equation (2.6) can be written in the following form:

$$n_{\mathrm{s}} = \frac{\epsilon_2}{q d_{\mathrm{GC}}} \left(V_{\mathrm{GC}} - V_{\mathrm{TN}}\right), \tag{2.8}$$

where V_{TN} is the so-called threshold voltage given as

$$V_{\mathrm{TN}} = \phi_{\mathrm{B}} - \frac{\Delta E_{\mathrm{C}}}{q} + \frac{E_{\mathrm{F}}}{q} - V_{\mathrm{P}}. \tag{2.9}$$

A more accurate description of the charge-control was proposed with the assumption that E_{Fi} varies linearly with n_{s} [55]. This dependence can be expressed as

$$E_{\mathrm{Fi}} = \Delta E_{\mathrm{F0}}(T) + a n_{\mathrm{s}}, \tag{2.10}$$

where $\Delta E_{\mathrm{F0}}(T)$ and a are constants. This relationships results in

$$n_{\mathrm{s}} = \frac{\epsilon_2}{q\left(d_{\mathrm{GC}} + \Delta d\right)} \left(V_{\mathrm{GC}} - V_{\mathrm{TN}}^{'}\right), \tag{2.11}$$

where Δd is the effective position of 2DEG given as

$$\Delta d = \frac{\epsilon_2 a}{q} \tag{2.12}$$

and $V_{\mathrm{TN}}^{'}$ is the modified threshold voltage given as

$$V_{\mathrm{TN}}^{'} = V_{\mathrm{TN}} + \Delta E_{\mathrm{F0}}(T). \tag{2.13}$$

The value of Δd was initially predicted to be around 80 $\mathring{A}$ [55]. Later, it was demonstrated that Δd is proportional to d_{GC} and depends strongly on the material system being used. For the lattice-matched AlInAs - InGaAs system investigated in this work, Δd was found to be 80 $\mathring{A}$ for d_{GC}=260 $\mathring{A}$ [60]. These values indicate a significant offset for effective distance of the 2DEG from the gate.

2.1.3 Velocity Saturation Under High Electric Field

At a given point in the channel, the drain-source current of a HEMT can be described as [52, 53]

$$I_{\mathrm{DS}} = q n_{\mathrm{s}} v_{\mathrm{d}} W_{\mathrm{G}}, \tag{2.14}$$

where v_{d} is the electron drift velocity and W_{G} is the gate-width. Equation (2.14) emphasizes the two distinct mechanisms that dictate I_{DS}. The carrier concentration (n_{s}) is controlled by V_{GC} as explained previously. The drift velocity is proportional to the drain-source electric field. For short-channel HEMTs ($L_{\mathrm{G}} \leq 2\ \mu m$) [61], the electric field induced by V_{DS} over the short drain-source distance is strong enough to saturate v_{d}. When the saturation velocity is reached (v_{sat}), I_{DS} becomes weakly dependent on V_{DS} and the device operates in the so-called saturation regime.

For single-valley semiconductors, such as Si or Ge, v_{d} is a monotonic function of the applied electric field [62]. However in compound semiconductors such as GaAs or InP, a two-valley structure is observed where the carrier mobility is different for each valley. When the carriers reach a certain energy, inter-valley transfer occurs and carriers propagate in the low mobility valley [63]. As a result, negative differential mobility is observed beyond a certain electric field and v_{d} starts to decrease with increasing electric field. For two-valley semiconductors, a transcendental relationship exists between v_{d} and the electron temperature (T_{e}) [53].

Carrier transport is limited by different mechanisms in the low and high electric field conditions. Under low electric field, the electron mobility is limited by ionized-

impurity scattering if the material is doped[1]. Since the dopants that contribute electrons to the 2DEG are located remotely from the channel, ionized impurity scattering is minimized in a 2DEG and mobility similar to that in the bulk material is achieved [52]. However, the electron velocity under high electric field is limited by optical phonon scattering, which is weakly dependent on doping levels [64]. Therefore, the advantage of achieving high electron mobility in 2DEG due to the elimination of ionized impurity scattering is only beneficial if the applied electric field is small. Furthermore, depending on the material system, the peak electron velocity (v_{p}) as a function of electric field can be lower in the 2DEG compared to the low-doped bulk case due to second order effects such as modified inter-valley transfer, real space transfer[2], and enhanced scattering via polar optical phonons [65].

To complicate the matter, non-uniform electric field across the channel result in significant spatial variations in the electron velocity [52]. The electric field is weak on the source side, where a high mobility corresponds to a higher velocity. As the electric field consistently increases towards the drain, the electron velocity reaches its peak and potentially drops back down depending on the energy level required for inter-valley transfer. Location dependent electron velocity can be predicted with Monte-Carlo simulations [66, 67]. A high low-field mobility, a high peak velocity, and a large inter-valley separation are desired for minimum transit time in short channel devices. A high mobilitiy is also critical for minimization of the source and drain access resistances in HEMT. For these reasons, $\mathrm{In_xGa_{1-x}As}$ has been the most favorable channel material to date for low-noise and high-frequency applications [68]. The low-field mobility and the peak electron velocity of InGaAs is proporitonal to the Indium content which is limited by strain [69]. Furthermore, the lattice-matched

[1]The amount of doping required for ionized-impurity scattering to be the dominant scattering mechanism depends on the material.

[2]This phenomena describes the electrons that have energy to overcome the heterojunction potential barrier and transfer to the barrier layer [52].

AlInAs - InGaAs system provides the second largest band offset among other HEMT material combinations, which leads to a high channel carrier concentration [70]. The AlGaN - GaN material system provides the largest band offset and has been favored for high-power applications.

Since the exact relationship between the electric field and the electron velocity is too complicated to be modeled analytically, a simple piecewise model was used for the rest of the analysis [53]:

$$v_{\mathrm{d}} = \begin{cases} \mu_{\mathrm{n}} E, & \text{if } E < E_{\mathrm{crit}} \\ v_{\mathrm{sat}}, & \text{if } E \geq E_{\mathrm{crit}} \end{cases} \tag{2.15}$$

where μ_{n} is the electron mobility, E is the applied electric field, and E_{crit} is the electric field at the onset of velocity saturation. This approach was successfully applied to real HEMTs with the assumption that v_{sat}=v_{p} [71].

2.1.4 I-V and C-V Characteristics of HEMTs

With the knowledge of underlying physical mechanisms of HEMT operation, current-voltage (I-V) and capacitance-voltage (C-V) characteristics of HEMT will be presented. First, the simplified case of a long-channel transistor will be considered. The analysis will then be expanded to evaluate the short-channel transistor operation.

2.1.4.1 I-V Model

For a long-channel HEMT, the I_{DS} expression in Equation (2.14) ignores the fact that 2DEG potential varies across the channel, which results in position dependence of n_{s}. When this is taken into account and the Poisson equation for charge-control is

solved in two dimensions, a quadratic relationship between I_{DS} and the gate-source voltage (V_{GS}) is obtained in the saturation mode as [53]

$$I_{\mathrm{DS}} = \mu_{\mathrm{n}} C_{\mathrm{G0}} \frac{W_{\mathrm{G}}}{L_{\mathrm{G}}} \frac{1}{2} (V_{\mathrm{GS}} - V_{\mathrm{TN}})^2, \tag{2.16}$$

where C_{G0} is the nominal gate - 2DEG capacitance per unit area given as

$$C_{\mathrm{G0}} = \frac{\epsilon_2}{d_{\mathrm{GC}} + \Delta d}. \tag{2.17}$$

For V_{DS}=V_{GS}-V_{TN}, the channel is pinched-off at the drain side meaning that n_{s}=0. When V_{DS} is further increased, the pinch-off point shifts towards the source and a depleted region is formed near the drain with a width of X_{S}. When $V_{\mathrm{DS}} \geq V_{\mathrm{GS}} - V_{\mathrm{TN}}$, the excess electric field is dropped across this saturation region so that the electric field at the point L_{G}-X_{S} stays constant. As V_{DS} is increased beyond the saturation, X_{S} becomes greater, which reduces the effective channel length. This phenomena is called channel length modulation and, for long channel devices, its impact on I_{DS} is negligible. Therefore the drain conductance (g_{ds}) is assumed to be practically zero.

The transconductance ($g_{\mathrm{m,HEMT}}$) of a long channel HEMT in the saturation regime is obtained by differentiating I_{DS} with respect to V_{GS}:

$$g_{\mathrm{m,HEMT}} = \mu_{\mathrm{n}} C_{\mathrm{G0}} \frac{W_{\mathrm{G}}}{L_{\mathrm{G}}} (V_{\mathrm{GS}} - V_{\mathrm{TN}}). \tag{2.18}$$

Equation (2.18) indicates that $g_{\mathrm{m,HEMT}}$ improves as the distance carriers travel under the gate is reduced. However for a certain L_{G}, velocity saturation in the channel will occur. In this case, I_{DS} in the saturation regime is given as [53]

$$I_{\mathrm{DS}} = \beta E_{\mathrm{crit}}^2 L_{\mathrm{G}} + \left[\left(\beta E_{\mathrm{crit}}^2 L_{\mathrm{G}} \right)^2 + \beta^2 E_{\mathrm{crit}}^2 \left(V_{\mathrm{GS}} - V_{\mathrm{TN}} \right)^2 \right]^{1/2}, \tag{2.19}$$

where $\beta = \mu_{\mathrm{n}} C_{\mathrm{G0}} W_{\mathrm{G}}$. The transconductance of the short-channel HEMT is then

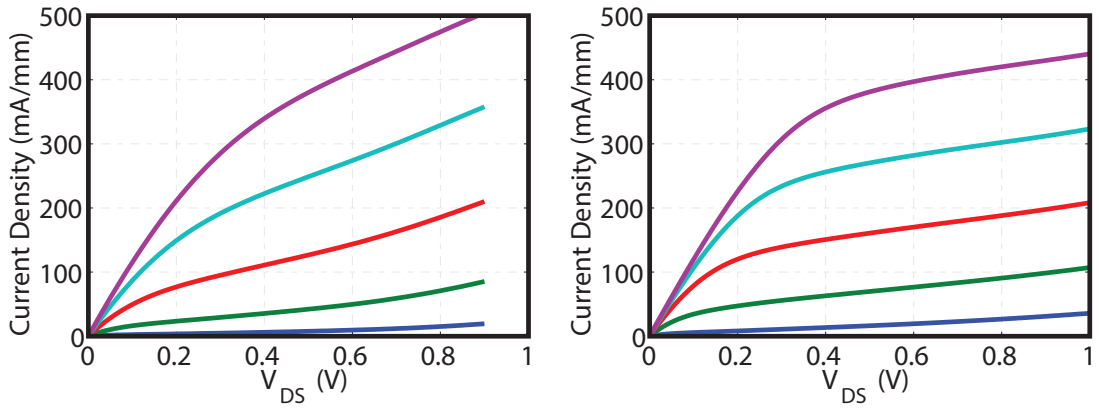
$$g_{\mathrm{m,HEMT}} = \mu_{\mathrm{n}} C_{\mathrm{G0}} \frac{W_{\mathrm{G}}}{L_{\mathrm{G}}} \frac{V_{\mathrm{GS}} - V_{\mathrm{TN}}}{\left(1 + \left(V_{\mathrm{GS}} - V_{\mathrm{TN}}\right)^2 / \left(E_{\mathrm{crit}} L_{\mathrm{G}}\right)^2\right)^{1/2}}. \tag{2.20}$$

This I-V model predicts an abrupt turn-on when V_{GS} exceeds V_{TN}. In reality, a smooth transition from sub-threshold to on-state is expected, with significant drain current flowing in the sub-threshold regime for short-channel devices. Around this transition, a HEMT is expected to have an I-V characteristic similar to that of BJT, as the channel under the gate forms an $\mathrm{n^+}$-n-$\mathrm{n^+}$ structure with the ungated channel regions [53].

Equation (2.20) indicates a linear relationship between $g_{\mathrm{m,HEMT}}$ and the overdrive voltage (V_{GS}-V_{TN}) for weak overdrive operation. This is similar to what is predicted by the long-channel I-V model. As the bias current is increased, the influence of velocity saturation becomes more substantial and the dependence of $g_{\mathrm{m,HEMT}}$ on the overdrive voltage becomes weaker.

Equation (2.19) still neglects the dependence of I_{DS} on V_{DS}, which can be significant for short-channel FETs (Fig. 2.3). For devices that are sized for operation at microwave frequencies, this dependence results in output resistances less than 1 kΩ. Therefore, a physical description is required for V_{DS} dependence of I_{DS} is required.

The finite output resistance is associated with the fact that the width of the 2DEG perpendicular to the gate (Δy) is non-zero. This concept is called the channel opening, which predicts that the channel-width is inversely proportional to the channel electron concentration (n_{s}) [72]. This can be visualized as follows. When n_{s} is increased with V_{G}, the electric field in the channel gets stronger, which increases the electron confinement and reduces the 2DEG width. However, it was reported that the channel opening principle does not fully account for the g_{ds} values observed in practical devices. A potential explanation is that the elevated temperature of electrons in the channel (T_{e}) causes n_{s} to be lower than what is expected, which in return



(a) (b)

Figure 2.3: Output curves of (a) 32 nm SOI MOSFET and (b) 100 nm InP HEMT.

causes the channel to be wider [53]. Nevertheless, with the assumption that Δy does not vary across the channel, an analytical model was provided which describes the relationship between V_{DS} and I_{DS} as [73]

$$V_{\mathrm{DS}} = V_{\mathrm{GS}} - V_{\mathrm{TN}} + \frac{I_{\mathrm{D}}}{\mu_{\mathrm{n}} C_{\mathrm{G0}} W_{\mathrm{G}} E_{\mathrm{crit}}} + \frac{I_{\mathrm{D}}}{2 W_{\mathrm{G}} \Delta y \epsilon_1 v_{\mathrm{sat}}} \left(L_{\mathrm{G}} - X_{\mathrm{S}}\right)^2 + E_{\mathrm{crit}} \left(L_{\mathrm{G}} - X_{\mathrm{S}}\right), \tag{2.21}$$

where ϵ_1 is dielectric constant of the small band-gap region. Equation (2.21) indicates that for a fixed V_{DS}, increasing I_{DS} through V_{GS} narrows the saturation region width. As a result, g_{ds} is expected to be proportional to I_{DS} for a constant V_{DS}. Ignoring the variations in the saturation region width under small-signal V_{DS} excitation, g_{ds} is not expected to change with V_{DS}.

2.1.4.2 C-V Model

Based on the parallel-plate approximation, Equation (2.17) predicts a constant gate capacitance. However, this is only valid for a constant carrier distribution across the channel. As discussed earlier, n_{s} varies across the channel as the channel potential increases towards the drain. Therefore, the gate capacitance is found by integrating n_{s} across the channel in a similar procedure to that of the calculation of I_{DS}.

When the device is in the triode regime, the total gate capacitance is expected to be roughly shared equally between the source and drain nodes. These two capacitances are denoted as C_{GS} and C_{GD}. In the saturation regime, the depleted channel region with width L_{G}-X_{S} results in a discontinuity between the channel and the drain node. Therefore, C_{GS} is mainly associated with the entire gate-2DEG capacitance whereas C_{GD} represents depletion capacitance of the Schottky barrier. C_{GS} can then be described as a function of g_{m} [74]:

$$C_{\mathrm{GS}} \approx \frac{g_{\mathrm{m,HEMT}} L_{\mathrm{G}}}{v_{\mathrm{avg}}}. \tag{2.22}$$

Where v_{avg} is the average electron velocity in the channel. It should be noted that parasitic coupling and Schottky depletion capacitance between the gate and source terminals also contribute to C_{GS}.

As mentioned earlier, a parallel conduction path is formed in the barrier layer when the depletion regions of the Schottky gate and the heterojunction interface cease to overlap. At the onset of this parallel conduction path, donor neutralization in the barrier layer is modulated with the gate voltage rather than the charge in the channel. As a result, charge-control in the channel degrades and the transconductance reduces with further increase of the gate voltage. This unique feature of HEMT results in a bell-shaped $g_{\mathrm{m,HEMT}}$-V_{GS} relationship. Interestingly, the measured C_{GS} curve also has the bell-shape despite the fact that the gate charge does not decrease when the parallel conduction path forms. This is because the donors in the barrier layer have a high resistance discharge path and can not respond to high frequency signals [54]. As capacitance extractions are usually obtained from high frequency s-parameter measurements, C_{GS} appears to decrease with gate voltage beyond the onset of the parallel conduction path formation.

The depletion capacitance of a simple metal-semiconductor junction is given as [75]

$$C_{\mathrm{depl}} = \sqrt{\frac{q\epsilon N_{\mathrm{D}}}{2\left(\phi_{\mathrm{bi}} + V_{\mathrm{R}} - kT/q\right)}}, \tag{2.23}$$

where ϕ_{bi} is the junction built-in potential and V_{R} is the applied reverse voltage. In a real HEMT, the formation of C_{GD} is far from ideal that of a metal-semiconductor junction, as lateral coupling between gate and drain contacts is expected and the barrier layer is pulse-doped. Therefore, use of a generalized depletion capacitance formula is more appropriate [76]:

$$C_{\mathrm{GD}} \approx \frac{C_{\mathrm{GD0}}}{\left[\left(V_{\mathrm{DG}}/\phi_{\mathrm{bi}}\right) + 1\right]^{m}}, \tag{2.24}$$

where V_{DG} is the drain-gate voltage, C_{GD0} is the value of C_{GD} for V_{DG}=0 and m is an empirical constant related to the junction profile.

The maximum current-gain frequency of a HEMT is a function of transconductance and the total gate capacitance and can be written as

$$f_{\mathrm{t}} = \frac{g_{\mathrm{m,HEMT}}}{2\pi\left(C_{\mathrm{GS}} + C_{\mathrm{GD}}\right)}. \tag{2.25}$$

Equation (2.22) indicates that improving the $g_{\mathrm{m,HEMT}}/C_{\mathrm{GS}}$ ratio requires reduction of the transit time ($L_{\mathrm{G}}/v_{\mathrm{avg}}$) in the channel. Thus, reducing the electron transit time is critical for the realization of high-speed HEMTs.

2.1.5 Dynamic Range Limitations

Having described the transport mechanisms of HEMT, the limiting factors of the device noise and nonlinearity will be discussed next.

2.1.5.1 Noise Performance

Aside from thermal noise generated by the parasitic access resistances, the broadband noise performance of HEMTs is limited by gate leakage current, gate induced

noise and channel noise. Low frequency noise is limited by the gate leakage current, whereas the gate induced noise and channel noise become dominant as the frequency increases. These individual sources of noise are uncorrelated and they will be reviewed in the following discussion.

The current density of a Schottky junction according to the thermionic emission-diffusion theory is written as [77]

$$J_{\mathrm{S}} = A^{**}T^{2}e^{-q\phi_{\mathrm{Bn}}/kT}\left(e^{qV/kT}-1\right), \tag{2.26}$$

Where V is the applied voltage, A^{**} is the electric field dependent effective Richardson constant, and ϕ_{Bn} is the peak barrier height in the semiconductor region. For the gate voltages HEMTs are typically operated at, the gate leakage current corresponds to voltage dependent saturation current of the Schottky junction. This current is due to electron tunneling and acts as a shot noise source [78]. If the leakage current is due to avalanche breakdown, the impact of this shot noise is multiplied. For the values of V_{DG} considered in this work, avalanche breakdown is unlikely to occur and the leakage currents are modeled as typical shot noise sources. The power spectral density of the gate leakage current can be written as

$$\overline{|i_{\mathrm{g}}|^{2}} = 2q\left(|I_{\mathrm{GS}}|+|I_{\mathrm{GD}}|\right), \tag{2.27}$$

where I_{GS} and I_{GD} are the currents flowing through gate-source and gate-drain junctions, respectively.

The gate induced noise corresponds to the thermal noise generated by the channel carriers that are capacitively coupled to the gate [79]. As will be shown in Chapter IV, this noise source is assigned to a non-quasi static (NQS) resistance, with the assumption that its equivalent temperature follows the ambient temperature.

The channel can be interpreted as a thermal noise source with its equivalent temperature varying across the channel. In the low-field region, the carriers are not velocity saturated and they are in thermal equilibrium with the lattice. As the electric field increases towards the drain side of the channel, the electron temperature starts to rise beyond the ambient temperature. This temperature deviation results in so-called hot-electron noise [80]. For practical proposes, the average channel noise is typically represented by assigning an empirical temperature (T_{D}) to the output conductance [81]. This temperature is proportional to the average electron energy in the channel [53].

It is well known that T_{D} is proportional to the bias current. The transconductance also improves with bias which mitigates the influence of channel noise on the overall performance. Therefore, the optimum bias for achieving minimum noise is expected to yield the lowest value for the following relationship [13]:

$$f\left(V_{\mathrm{DS}}, I_{\mathrm{DS}}\right) \propto \frac{\sqrt{I_{\mathrm{DS}}}}{g_{\mathrm{m,HEMT}}}. \tag{2.28}$$

For a HEMT, V_{DS} also has a significant impact on the noise performance. Increasing V_{DS} beyond a certain point results in impact-ionization and deactivation of traps, which increases the gate leakage current as well as the channel noise temperature and also reduces the output conductance [44, 82]. Therefore, V_{DS} of the HEMT must be selected carefully. As certain aspects of these second order phenomena become more pronounced at cryogenic temperatures, they will be discussed further in Chapter 3.

2.1.5.2 Linearity Performance

In order to evaluate how the transport properties of a HEMT limit its linearity performance, the derivatives of the I-V and C-V expressions are investigated. Since the nth order distortion is a function of all previous order nonlinearities [83], evaluating

the third-order intermodulation performance of a HEMT requires an understanding of all the derivatives up to third order[3].

The first derivatives of I_{DS} with respect to V_{GS} and V_{DS} are g_{m} and g_{ds}, respectively. Similarly, first derivative of the gate-source charge (Q_{GS}) with respect to V_{GS} is C_{GS} and first derivative of the gate-drain charge (Q_{GD}) with respect to the V_{GD} is C_{GD}. These first derivatives were already discussed in section 2.1.4. Therefore, the second and third derivatives of I_{DS}, Q_{GS}, and Q_{GD} remain to be explained. To start with, the second derivative of I_{DS} with respect to V_{GS} is given for short channel HEMT as

$$g_{\mathrm{m,HEMT}}^{'} = \mu_{\mathrm{n}} C_{\mathrm{G0}} \frac{W_{\mathrm{G}}}{L_{\mathrm{G}}} \frac{1}{\left(1 + \left(V_{\mathrm{GS}} - V_{\mathrm{TN}}\right)^2 / \left(E_{\mathrm{crit}} L_{\mathrm{G}}\right)^2\right)^{3/2}}. \tag{2.29}$$

A transistor is desired to have a high and weakly bias dependent transconductance. Therefore, $g_{\mathrm{m,HEMT}}^{'}$ can be normalized to $g_{\mathrm{m,HEMT}}$ in order to gain insight:

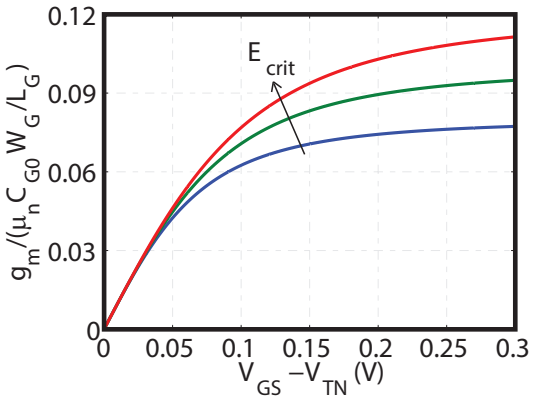
$$g_{\mathrm{m,HEMT}}^{'} / g_{\mathrm{m,HEMT}} = \frac{1}{\left(V_{\mathrm{GS}} - V_{\mathrm{TN}}\right)\left(1 + \left(V_{\mathrm{GS}} - V_{\mathrm{TN}}\right)^2 / \left(E_{\mathrm{crit}} L_{\mathrm{G}}\right)^2\right)}. \tag{2.30}$$

Similarly, the normalized second derivative of $g_{\mathrm{m,HEMT}}$ can be written as

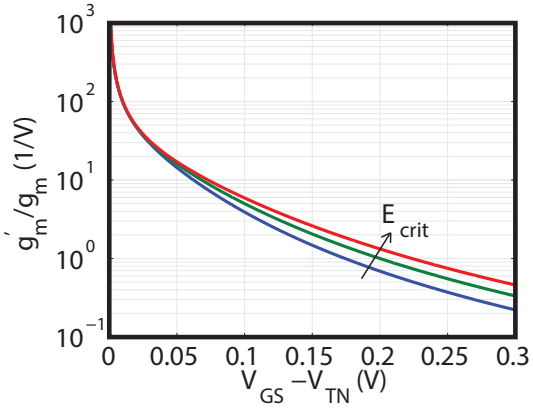
$$g_{\mathrm{m,HEMT}}^{''} / g_{\mathrm{m,HEMT}} = \frac{-3\left(E_{\mathrm{crit}} L_{\mathrm{G}}\right)^2}{\left(\left(V_{\mathrm{GS}} - V_{\mathrm{TN}}\right)^2 + \left(E_{\mathrm{crit}} L_{\mathrm{G}}\right)^2\right)^2}. \tag{2.31}$$

Values of Equations (2.20), (2.30) and (2.31) are computed and shown in Figure 2.4 for different E_{crit}. At high overdrive levels, $g_{\mathrm{m,HEMT}}$ and its normalized derivatives are proportional to E_{crit}. At low overdrive levels, $g_{\mathrm{m,HEMT}}$ and $g_{\mathrm{m,HEMT}}^{'}$ appear to be insensitive to E_{crit} whereas $g_{\mathrm{m,HEMT}}^{''}$ decreases with E_{crit}. Thus, at low overdrive levels, where the low noise amplifiers are expected to be operated, the linearity of the HEMT transconductance is proportional to E_{crit}.

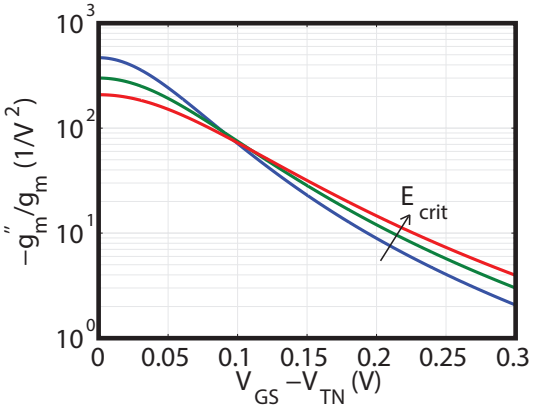
[3]While the fifth order distortion mathematically contributes to the third-order intermodulation, this contribution was observed to be negligible for the signal levels of interest. An analogy is the influence of third order nonlinearity on the compression of linear gain.



(a)



(b)



(c)

Figure 2.4: Theoretical normalized (a) g_{m}, (b) $g_{\mathrm{m}}^{'}$ and (b) $g_{\mathrm{m}}^{''}$ of short-channel HEMT. Traces correspond to E_{crit}=8,10 and 12 kV/cm.

Equation (2.21) and Figure 2.3 indicate that g_{ds} is expected to be weakly nonlinear provided that the device is kept in deep-saturation and higher-order effects are not present. However, due to channel width and length modulation, g_{ds} is also a function of V_{GS}. While it is hard to provide an analytical description, this dependence can be a significant source of nonlinearity, as the I_{DS}-V_{DS} relationship given by Equation (2.21) includes a square-law dependence on the effective channel length (L_{GS}-X_{S}).

Charge nonlinearity is considered next. As indicated by Equation (2.22), C_{GS} is related to g_{m} through the channel transit time ($L_{\mathrm{G}}/v_{\mathrm{avg}}$). Assuming that v_{avg} is weakly nonlinear for short channel devices, the bias dependence and nonlinearity of C_{GS} is expected to have a similar trend to that of $g_{\mathrm{m,HEMT}}$.

The feedback capacitance (C_{GD}) nonlinearity can have a significant impact on the overall performance due to the Miller effect. The first and second derivatives of C_{GD} with respect to V_{GD} can be written as

$$C'_{\mathrm{GD}} = \frac{-m}{V_{\mathrm{DG}} + \phi_{\mathrm{bi}}} C_{\mathrm{GD}} \tag{2.32}$$

and

$$C''_{\mathrm{GD}} = \frac{-m\,(m+1)}{(V_{\mathrm{DG}} + \phi_{\mathrm{bi}})^2} C_{\mathrm{GD}}. \tag{2.33}$$

Hence, higher V_{GD} and ϕ_{bi} reduces the feedback capacitance nonlinearity. As discussed earlier, the drain voltage of HEMT is limited at the high end due to the onset of impact ionization.

2.2 Heterojunction Bipolar Transistor

A bipolar transistor consists of two p-n junctions connected back-to-back. Since electron mobility is greater than hole mobility for direct and indirect band-gap semiconductors [84], n-p-n type transistors enable higher speeds and are used for high frequency applications. The shared p region is called base and the two n regions

are called the emitter and the collector. A homojunction bipolar transistor (BJT) is made entirely of a single semiconducting element (most commonly silicon), where the doping concentrations are different in each region. A heterojunction bipolar transistor (HBT) consists of at least two different bulk materials. For the device of interest in this work, a SiGe alloy is incorporated in the base, whereas the emitter and collector regions are silicon.

The impact of the germanium content in the base will be conceptually discussed first in this section. The base and collector currents will then be studied and the influence of the germanium content in the base will be quantitatively described. The dynamic charge properties will be investigated, which will enable the understanding of terminal capacitances and transit delays. Finally, the sources of noise and nonlinearity in SiGe HBTs will be presented.

2.2.1 Influence of SiGe Base

The band diagram of a SiGe HBT is shown in Figure 2.5. Since germanium has a lower bandgap (0.67 eV) than that of silicon (1.11 eV) [85], incorporating germanium in the base lowers the energy barrier for electrons in the emitter while the energy barrier for holes in the base remains unchanged. Therefore, for a given base-emitter voltage (V_{BE}), a SiGe HBT provides more collector current (I_{C}) than an identical Si BJT while their base currents (I_{B}) are the same. As a result, a SiGe HBT has a greater current gain (β) than that of Si BJT. Furthermore, a germanium gradient across the base dramatically improves the Early voltage (V_{A}) [86].

Having a graded germanium profile across the base also improves dynamic properties of the device. The emitter charge storage time (τ_{e}) is inversely proportional to β, which is enhanced with a SiGe base. The base transit time is reduced by the quasi-electric field that is generated by the germanium grade across the base [26, 87]. The reduction in these time constants enables a faster response time of I_{C} to V_{BE}.

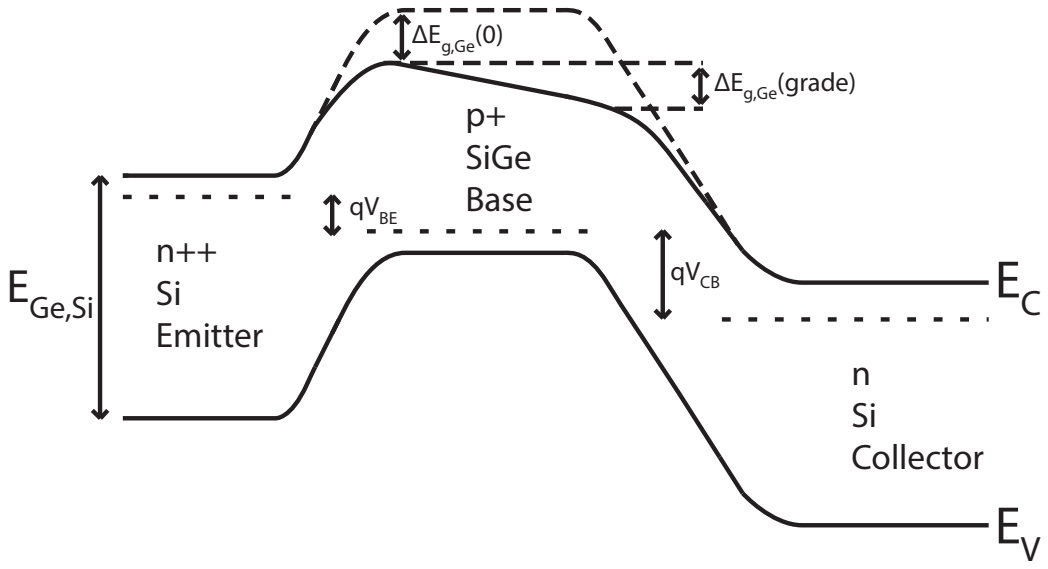


Figure 2.5: Band diagram of SiGe HBT [26]. Dotted lines between E_{C} and E_{F} correspond to the Fermi levels in each region.

Finally the germanium content in the base improves the sheet resistance of the base access path, which is given as [26]

$$R_{\mathrm{B,sheet}} = \frac{1}{q\mu_{\mathrm{pb}} N_{\mathrm{AB}}^{-} W_{\mathrm{B}}} (\Omega/\square), \tag{2.34}$$

where μ_{pb} is the hole mobility in the base, N_{AB}^{-} is the ionized acceptor concentration in the base, and W_{B} is the base-width between the emitter and collector boundaries. The hole mobility of germanium is four times greater than that of silicon, which inherently reduces $R_{\mathrm{B,sheet}}$ for SiGe HBTs.

2.2.2 Terminal Currents

With the consequences of the germanium content in the base understood, the terminal currents of SiGe HBT can be studied quantitatively. Assuming that the electron distribution in the conduction band follows Boltzmann statistics, the Ge grading is strong, the base doping profile is constant, and bulk and surface combi-

nation is negligible, the collector current density of SiGe HBT under low-injection is given as [26, 88]

$$J_{\mathrm{C}} \approx J_{\mathrm{C0}} \left(e^{qV_{\mathrm{BE}}/kT} - 1 \right), \tag{2.35}$$

where the collector saturation current is given as

$$J_{\mathrm{C0}} = n_{\mathrm{io,Si}}^2 \tilde{\gamma} \tilde{\eta} \frac{\mu_{\mathrm{nb,Si}}}{N_{\mathrm{AB}}^- W_{\mathrm{B}}} \Delta E_{\mathrm{g,Ge}}(grade) e^{\left(\Delta E_{\mathrm{g,app}} + \Delta E_{\mathrm{g,Ge}}(0)\right)/kT}, \tag{2.36}$$

where $\tilde{\eta} > 1$ is the electron mobility ratio between SiGe and Si in the base, $\tilde{\gamma} < 1$ is the effective density-of-states ratio between SiGe and Si, $\mu_{\mathrm{nb,Si}}$ is the Si electron mobility in the base, $\Delta E_{\mathrm{g,app}}$ is the doping related bandgap narrowing, and $n_{\mathrm{io,Si}}$ is the intrinsic carrier concentration in Si which is [77]

$$n_{\mathrm{io}} = \sqrt{N_{\mathrm{C}} N_{\mathrm{V}}} e^{-E_{\mathrm{g}}/(2kT)}, \tag{2.37}$$

where N_{C} and N_{V} are the conduction and valence band density of states. The apparent bandgap narrowing occurs due to the heavy doping of emitter and is given as [89]

$$\Delta E_{\mathrm{g,app}} \approx 18.7 \times 10^{-3} \ln \left\{ \frac{N_{\mathrm{AB}}^-}{N_{\mathrm{DE}}^+} \right\}, \tag{2.38}$$

where N_{DE}^+ is the ionized donor concentration in the emitter. Equation (2.38) is valid if the base and emitter doping concentrations are greater than $7{\times}10^{17}$ cm^{-3} which is satisfied for the modern SiGe HBTs.

The base current density of a SiGe HBT is given as [26]

$$J_{\mathrm{B}} \approx J_{\mathrm{B0}} \left(e^{qV_{\mathrm{BE}}/kT} - 1 \right), \tag{2.39}$$

where the base saturation current is given as

$$J_{\mathrm{B0}} = \frac{kT \mu_{\mathrm{pe}} n_{\mathrm{io,Si}}}{L_{\mathrm{PE}} N_{\mathrm{DE}}^+}, \tag{2.40}$$

where μ_{pe} and Ł$_{\mathrm{PE}}$ are the hole mobility and diffusion length in the emitter, respectively. Thus, J_{B} is independent of the germanium content in the base. Assuming that the $\tilde{\gamma}\tilde{\eta}$ product is not too far off from unity, Equations (2.36) and (2.40) clearly illustrate that J_{C} and β are enhanced due the bandgap offset achieved with the germanium content in the base.

With the terminal currents understood, the transconductance and input conductance of a SiGe HBT can be easily derived. The transconductance is obtained by differentiating I_{C} with respect to V_{BE} and given as

$$g_{\mathrm{m,HBT}} = \frac{qI_{\mathrm{C}}}{kT}. \tag{2.41}$$

Similarly, the input conductance is found by differentiating I_{B} with respect to V_{BE} and can be written as

$$g_{\mathrm{be}} = \frac{g_{\mathrm{m,HBT}}}{\beta_{\mathrm{AC}}}, \tag{2.42}$$

where β_{AC} is the AC current gain of the device. Equations (2.41) and (2.42) imply that, for a given $g_{\mathrm{m,HBT}}$, g_{be} reduces significantly with increasing germanium content in the base.

The terminal currents were derived under low-injection conditions where the base hole concentration (p_{b}) in the valence band does not significantly deviate from N_{AB}^{-}. However, p_{b} is proportional to the number of injected electrons in the base which degrades the theoretical β as J_{C} is increased. The increase in p_{b} with J_{C} is given as

$$\Delta p_{\mathrm{b}} = \frac{J_{\mathrm{C}}}{qv_{\mathrm{d}}}, \tag{2.43}$$

where v_{d} is the electron drift velocity in the base. Taking this increase into account, the effective current gain can be written as

$$\beta' = \beta \frac{1}{1 + J_{\mathrm{C}} / \left(qN_{\mathrm{AB}}^{-} v_{\mathrm{d}}\right)}. \tag{2.44}$$

This phenomena was first described by Webster [90] and has been known as the Webster effect since then.

2.2.3 Transit Times and Capacitances

With the static transport properties of SiGe HBT described, the dynamic properties, which ultimately relate to the device speed, can be discussed. The intrinsic emitter-collector transit time, which describes the response time of I_{C} to V_{BE}, is given as [16, 91]

$$\tau_{\mathrm{ec}} \approx \tau_{\mathrm{e}} + \tau_{\mathrm{b}} + \tau_{\mathrm{cbd}} + \frac{kT}{qI_{\mathrm{C}}} \left(C_{\mathrm{BE,depl}} + C_{\mathrm{BC,depl}} \right), \tag{2.45}$$

where τ_{e} is the emitter charge storage time, τ_{b} and τ_{cbd} are the transit times in base and collector-base depletion regions, and $C_{\mathrm{BE,depl}}$ and $C_{\mathrm{BC,depl}}$ are the depletion regions corresponding to the base-emitter and base-collector regions, respectively. According to the drift-difussion theory, the time constants are given as [26]

$$\tau_{\mathrm{e}} \approx \frac{q}{2kT} \frac{W_{\mathrm{E}}^2}{\mu_{\mathrm{pe}} \beta_{\mathrm{DC}}}, \tag{2.46}$$

$$\tau_{\mathrm{b}} \approx \frac{qW_{\mathrm{B}}^2}{\tilde{\eta} \mu_{\mathrm{nb,Si}}} \frac{1}{\Delta E_{\mathrm{g,Ge}}(grade)}, \tag{2.47}$$

and

$$\tau_{\mathrm{cbd}} \approx \frac{W_{\mathrm{CBD}}}{2v_{\mathrm{sat}}}, \tag{2.48}$$

where W_{E} is the emitter-width along the current path, W_{CBD} is the width of base-collector depletion region, and v_{sat} is the electron saturation velocity in the base-collector depletion region. Hence, τ_{e} improves proportionally to the β_{DC} increase with the germanium content in the base. The quasi-electric field generated by the germanium gradient in the base improves τ_{b}.

The depletion capacitances in Equation (2.45) remain to be explained in order to fully account for the total forward transit time. The depletion capacitance of a p-n junction with abrupt doping transition is given as [85]

$$C_{\text{depl}} = \sqrt{\frac{q\epsilon}{2\left(\phi_{\text{bi}} + V_{\text{R}}\right)} \frac{N_{\text{A}} N_{\text{D}}}{N_{\text{A}} + N_{\text{D}}}}, \tag{2.49}$$

which is very similar to Schottky depletion capacitance given by Equation (2.23). The depletion capacitance can be generalized for different doping profiles as [76]

$$C_{\text{depl}} \approx \frac{C_{\text{depl,0}}}{\left[\left(V_{\text{R}}/\phi_{\text{bi}}\right) + 1\right]^{m}}, \tag{2.50}$$

where $C_{\text{depl,0}}$ is the value of C_{depl} at V_{R}=0 and m is a constant describing the doping profile ($m = 1/2$ for abrupt doping transition and $m = 1/3$ for linearly graded doping). Therefore, the depletion capacitances are mainly determined by the doping levels.

The base-emitter capacitance ($C_{\text{BE,diff}}$) is a result of charge neutrality. As the minority carriers traverse across the base and emitter, the majority carriers are modulated in order to preserve the net charge value in these regions [16]. The diffusion capacitance resulting from this phenomena can be defined with respect to the time constants given in Equations (2.46) - (2.48). Hence, it can be written as [16]

$$C_{\text{BE,diff}} = g_{\text{m,HBT}}\left(\tau_{\text{b}} + \tau_{\text{e}} + \tau_{\text{cbd}}\right) = g_{\text{m,HBT}}\tau_{\text{f}}. \tag{2.51}$$

This relationship is strikingly similar to that between C_{GS} and g_{m} of HEMT where the ratio $C_{\text{GS}}/g_{\text{m}}$ is equal to channel transit time.

Equation (2.45) indicates that τ_{ec} is inversely proportional to I_{C}. As I_{C} is increased, the base-collector depletion region gets wider on the collector side as the collector electron concentration is reduced by $J_{\text{C}}/\left(qv_{\text{sat}}\right)$. When the boundary of

the depletion region reaches the extrinsic collector, which has a higher doping concentration than that of the intrinsic collector, the electric field gradient across the space-charge region becomes zero. With further increase of I_{C}, the base-collector depletion region shifts to the intrinsic-extrinsic collector boundary, which effectively widens the base region. As a result, τ_{b} significantly increases since it is inversely proportional to the base-width. This phenomena causes f_{t} roll-off at high current densities and was first described by Kirk [92].

2.2.4 Non-Equilibrium Transport

The HBT transport properties presented above are based on the classical drift-diffusion theory, which assumes that electron velocity in the base is limited by scattering. This assumption is valid if the base-width is significantly longer than the carrier mean free path length [93]. If the base-width is much shorter than the carrier mean free path length, electron transport in the base becomes ballistic, meaning that the electrons do not experience scattering in the base. In this case, electrons in the base are subject to velocity-overshoot, which would result in a collector current higher than expected. For modern SiGe HBTs, the base-width can be comparable to the carrier mean free path length at lower temperatures. In this case, quasi-ballistic transport is expected where the electrons encounter only a few collisions across the base.

Similar to the hot-electron concept of a HEMT, electrons that are subject to ballistic transport in the base region of HBT will have a different temperature than that of the lattice. This can be taken into account in the form of an ideality factor and the resulting collector current density can be written as [16, 26]

$$J_{\mathrm{C}} \approx J_{\mathrm{C0}} \left(e^{qV_{\mathrm{BE}}/(n_{\mathrm{c}}kT)} - 1 \right), \tag{2.52}$$

where $T_{\text{eff}} = n_{\text{c}}T$ describes the effective electron temperature. The saturation current is also observed to be higher than its ideal value at cryogenic temperatures due to an increase in the inrinsic carrier concentration ($n_{\text{io,Si}}$). The physical reason behind this trend has not been fully understood as $n_{\text{io,Si}}$ is an equilibrium property [94]. It should be noted that the effective temperature of holes will be different than that of electrons. This nondeality associated with the effective hole temperature is due to tunneling and field-asissted recombination process rather than non-equilibrium transport [16]. Regarding the small-signal performance, g_{m} is directly related to n_{c} as

$$g_{\text{m,HBT}} \approx \frac{qI_{\text{C}}}{n_{\text{c}}kT}. \tag{2.53}$$

Thus, any increase in n_{c} degrades the device transconductance. The base transit time (τ_{b}) can also be expressed to capture the delay associated with the ballistic transport [94]

$$\tau_{\text{b}} \approx \frac{W_{\text{B}}}{v_{\text{T}}} + \frac{qW_{\text{B}}^2}{\tilde{\eta}\mu_{\text{nb,Si}}} \frac{1}{\Delta E_{\text{g,Ge}}(grade)}, \tag{2.54}$$

where v_{T} is the thermal velocity which is inversely proportional to the ambient temperature.

Depending on the bias, the ideality factor also captures high-injection effects. Therefore, the existence of non-equilibrium transport is uniquely determined by observing the value of n_{c} in the low-injection regime ($J_{\text{C}} \approx$ 1-10 $\mu\text{A}/\mu\text{m}^2$). As it turns out, non-equilibrium transport becomes more pronounced at cryogenic temperatures. Therefore, it will be discussed further in Chapter 3.

2.2.5 Dynamic Range Limitations

Physical limitations to the HBT dynamic range will be analyzed next with respect to the previously studied transport mechanisms.

2.2.5.1 Noise Performance

Similar to HEMTs, series access resistances of HBTs generate thermal noise. Aside from this, the intrinsic noise performance of an HBT is limited by the shot noise generated by junction currents. Shot noise in p-n junctions is due to the randomness of carriers that have enough energy to overcome a potential barrier [95].

Shot noise in an HBT relates to the diffusion currents. The diffusion takes places in the base-emitter junction, whereas current across base-collector junction is due to the drift created by the electric field. Therefore, electrons and holes diffusing through base-emitter junction generate shot noise. The power spectral density of the total emitter shot noise current can then be written as [96]

$$\overline{|i_{\mathrm{e}}|^2} = 2q\left(I_{\mathrm{Ep}} + I_{\mathrm{En}}\right), \tag{2.55}$$

where I_{Ep} and I_{En} are the hole and electron components of the emitter current, respectively. A delayed version of I_{En} appears at the collector terminal which can be expressed as [97]

$$I_{\mathrm{C}} = I_{\mathrm{En}} e^{j\omega\tau_{\mathrm{n}}}, \tag{2.56}$$

where τ_{n} is the noise transit time which is different from the emitter-collector forward transit time τ_{ec}. In this case, the relationship between the collector and emitter noise currents can be written as [97]

$$i_{\mathrm{n,c}} = i_{\mathrm{n,e}} e^{-j\omega\tau_{\mathrm{n}}}. \tag{2.57}$$

The implications of this relationship is quite interesting. As the $\omega\tau_{\mathrm{n}}$ product increases with frequency, the collector noise current deviates from the emitter noise current. This has a positive effect on achievable noise performance of the device [26]. However, τ_{n} can only be determined from high frequency noise measurements which

are very challenging at cryogenic temperatures. At room temperature, it was demonstrated that τ_{n} is typically equal to 65% of the transit delay time [98, 99]. Unfortunately, such a direct relationship between τ_{n} and a small-signal element that can be defined from s-parameter measurements is not available at cryogenic temperatures. Therefore, the logical approach is to ignore τ_{n} for frequencies $f \ll f_{\mathrm{t}}$, where the impact of τ_{n} is expected to be minimal. With this, the noise parameters can be entirely predicted with the knowledge of the terminal currents and the small-signal model. In the DC frequency limit, the achievable noise can be predicted as [26]

$$T_{\mathrm{MIN,LF}} \approx T\frac{n_{\mathrm{cx}}}{\sqrt{\beta_{\mathrm{DC}}}} \tag{2.58}$$

where n_{cx} is the extrinsic collector current ideality factor. Thus, a low n_{cx} is as critical as a high β_{DC} for optimization of the noise performance. At high frequencies, the achievable noise of HBT is inversely proportional to f_{max}.

2.2.5.2 Linearity Performance

An HBT is inherently nonlinear due to its exponential I_{C}-V_{BE} relationship. As was done for HEMT, derivatives of the I-V and C-V expressions can be investigated for an HBT to evaluate its linearity limitations.

The V_{BE} dependence of the collector-current ideality factor prevents an exact derivation of g_{m} and its derivatives over a wide bias range. Even at room temperature where the V_{BE} dependence of n_{c} is weakest, n_{c} is a strong function of V_{BE} beyond 1 mA/μm^2, where a device is likely to be operated [26]. Therefore, the exact influence of n_{c} on the transconductance linearity can only be understood with experimental results. In the mean time, n_{c} can be assumed bias independent for the sake of analysis and derivatives of $g_{\mathrm{m,HBT}}$ with respect to V_{BE} can be obtained as

$$g_{\mathrm{m,HBT}}^{'} \approx \frac{I_{\mathrm{C}}}{\left(n_{\mathrm{c}}V_{\mathrm{TH}}\right)^2} \tag{2.59}$$

and

$$g_{\mathrm{m,HBT}}^{''} \approx \frac{I_{\mathrm{C}}}{\left(n_{\mathrm{c}} V_{\mathrm{TH}}\right)^3} \tag{2.60}$$

where $V_{\mathrm{TH}} = q/kT$ is thermal voltage. Therefore, the $g_{\mathrm{m,HBT}}$ nonlinearity is a strong function of the $n_{\mathrm{c}}V_{\mathrm{TH}}$ product. Ultimately, this relationship directly translates into the fact that the value of g_{m} and its nonlinearity are tightly coupled for HBTs.

The nonlinearity of the input conductance can be evaluated by taking derivatives of g_{be} with respect to V_{BE}:

$$g_{\mathrm{be}}^{'} \approx \frac{I_{\mathrm{B}}}{\left(n_{\mathrm{b}} V_{\mathrm{TH}}\right)^2} \tag{2.61}$$

and

$$g_{\mathrm{be}}^{''} \approx \frac{I_{\mathrm{B}}}{\left(n_{\mathrm{b}} V_{\mathrm{TH}}\right)^3}, \tag{2.62}$$

where n_{b} is the base current ideality factor. Therefore, the nonlinearity of g_{be} has a trend similar to that of the $g_{\mathrm{m,HBT}}$ nonlinearity. Since $g_{\mathrm{be}} \approx g_{\mathrm{m}}/\beta_{\mathrm{AC}}$, an HBT with higher β is expected to have a smaller g_{be} nonlinearity for a given $g_{\mathrm{m,HBT}}$.

Recalling that $C_{\mathrm{BE,diff}} = g_{\mathrm{m,HBT}}\tau_{\mathrm{f}}$, the nonlinearity of the base-emitter diffusion capacitance is also expected to have a trend similar to that of g_{m}. Since the time constants that constitute τ_{f} are mainly determined by physical parameters, τ_{f} is expected to be weakly bias dependent. Therefore, V_{BE} dependence of $C_{\mathrm{BE,diff}}$ is similar to that of g_{m}.

Finally, the nonlinearity of the depletion capacitance C_{BC} follows the same trend with that of the HEMT C_{GD} nonlinearity. For the completeness, derivatives of the C_{BC} with respect to V_{CB} are given as

$$C_{\mathrm{BC}}^{'} = \frac{-m}{V_{\mathrm{CB}} + \phi_{\mathrm{bi}}} C_{\mathrm{BC}} \tag{2.63}$$

and

$$C_{\mathrm{BC}}^{''} = \frac{-m\left(m+1\right)}{\left(V_{\mathrm{CB}} + \phi_{\mathrm{bi}}\right)^2} C_{\mathrm{BC}}. \tag{2.64}$$

An interesting point is that a p-n junction with linearly graded doping (m=1/3) is expected to have a more linear depletion capacitance than a junction with abrupt doping transition (m=1/2), assuming that ϕ_{bi} is similar for both cases.

2.3 Summary

The theoretical device operation and its limitation on the dynamic range performance has been presented in this chapter. For HEMTs, the channel charge-control mechanism and the electron velocity-electric field relationship in the channel were discussed in detail and the latter principle was found to be highly influential on the small-signal, noise, and linearity performance of the device. For HBTs, the effect of the germanium content in the base on static and dynamic properties of the transistor has been extensively studied. While the noise and high frequency performance significantly improve with the germanium content in the base, the transconductance nonlinearity of HBT was found to be mainly dependent on the thermal voltage and the ideality factor. For both devices, it was shown the nonlinearities of the charge-control capacitance and the transconductance are expected to have similar trends. With knowledge of the theoretical background, the expected temperature dependence of device performance will be discussed in the next chapter.

CHAPTER 3

EXPECTED PERFORMANCE AT CRYOGENIC TEMPERATURES

The expected temperature dependence of device performance is investigated in this chapter. As the focus of this work is the dynamic range of transistors at cryogenic temperatures, it is necessary to build an understanding of how the device operation changes in 7-300 K temperature range, where the experimental data is available. The general transport mechanisms and dynamic range limitations were studied in Chapter 2 and the temperature dependence of these properties is emphasized in this chapter. First, changes in the physical properties of the relevant materials are discussed. Second, these changes are linked to the temperature dependence of the I-V and C-V relationships. Finally, the limitations to the dynamic range performance at cryogenic temperatures are investigated.

3.1 Equilibrium Properties

In order to understand the overall temperature dependence of the static and dynamic transport properties of the devices, the key physical bias independent parameters are studied as a function of temperature in this section. First, the carrier concentration is considered, which provides the terminal currents. Second, the mobility and saturation velocity properties are investigated, which relate to the high frequency performance. Finally, the junction built-in potential is studied for both devices, which is essential for understanding the temperature dependence of the turn-on voltage, the feedback capacitance, and its nonlinearity.

3.1.1 Carrier Concentration

As demonstrated in Chapter 2, the terminal currents of HEMTs and HBTs depend strongly on the mobile carrier concentration. Therefore, understanding how the carrier concentration changes with temperature is critical for predicting the I-V relationship of the devices at cryogenic temperatures.

3.1.1.1 Equilibrium Channel Concentration of a HEMT

Referring back to Equation (2.5), the equilibrium channel concentration (n_{s0}) of a HEMT is a function of the conduction band offset (ΔE_C), the doping concentration (N_D), and the Fermi level (E_F), which has a transcendental dependence on n_s. The conduction band offset is not expected to change significantly with cooling, as the bandgap of both alloys that form the heterojunction is inversely proportional to the temperature [100, 101]. The Fermi level for a given n_s, increases with cooling, which would effectively reduce n_{s0} [55, 102]. However, the drop in the electric field at the heterojunction interface due to incomplete ionization is proportional to the thermal voltage, which compensates the increase in E_F with cooling [59]. As a result, it turns out that n_{s0} is weakly temperature dependent according to the analytical model presented in [59]. This outcome agrees well with the experimental results of an InP HEMT obtained with Hall measurements [103].

An interesting influence of cooling on the carrier concentration is that the 2DEG shifts towards the heterojunction interface [102]. This can be traced back to the assumption that the average carrier distance from the interface (Δd) is proportional to the slope of E_F with respect to n_s, which decreases with cooling [55]. An alternative explanation is that Δd is inversely proportional to thermal diffusion across the heterojunction [103]. As a result of the reduction in Δd with cooling, the gate capacitance and transconductance are expected to increase with cooling.

3.1.1.2 Intrinsic Carrier Concentration of an HBT

The base and collector saturation current of a SiGe HBT is proportional to the intrinsic carrier concentration of silicon ($n_{\mathrm{io,Si}}$), which is is proportional to $T^{3/2}e^{-E_{\mathrm{g}}/(2kT)}$ [26]. The linear term arises from the temperature dependence of the density of states and the exponential term describes the temperature dependence of the bandgap factor. The bandgap of silicon increases by 4% with cooling from 300 K to 0 K, which also contributes to the temperature dependence of n_{io} [104]. As a result, $n_{\mathrm{io,Si}}$ decreases by about a factor of 10^{-250} with cooling from 300 K to 10 K [26].

3.1.2 Mobility and Saturation Velocity

As demonstrated in Chapter 2, the velocity-field relationship is critical for static and dynamic performance of the devices. As this relationship is significantly dependent on temperature, it is important to investigate this aspect in order to understand the physical limitations of the dynamic performance at cryogenic temperatures.

3.1.2.1 Channel Electron Velocity of HEMT

The low-field mobility of the two-dimensional electron gas (2DEG) is critical for the channel transit time in a HEMT. As ionized impurity scattering in the 2DEG is minimized, the mobility follows that of the bulk case and is limited by polar-optical phonon scattering, which is inversely proportional to temperature [68]. Thus, the 2DEG mobility is greatly improved with cooling [64, 65, 105]. However, the 2DEG mobility can degrade or saturate below a certain temperature, as ionized-impurity and alloy scattering become dominant [68, 103].

As the electrons are subject to velocity saturation in the channel towards the drain, the peak electron velocity (v_{p}) and saturation velocity (v_{sat}) as a function of electric field are also critical. The temperature dependence of v_{sat} is provided as [68]

$$v_{\mathrm{sat}}(T) = v_{\mathrm{sat}}(0) - KT, \tag{3.1}$$

where $v_{sat}(0)$ and K are given as 7.7×10^6 cm/s and 5.3×10^3 cm/(s·K) for $In_{0.53}Ga_{47}As$, respectively. Thus, v_{sat} and v_p are expected to improve with cooling, although by a factor much less than that of the mobility improvement. This trend was experimentally verified in [65, 105]. As v_p improves less than the mobility with cooling, the critical electric field, which marks the onset of velocity saturation (E_{crit}), is expected to reduce with cooling.

3.1.2.2 Base-Emitter Transit Velocity of HBT

The terminal currents, transit time and base resistance of a SiGe HBT depend heavily on electron and hole mobilities in the base and emitter regions. Electron and hole mobilities in silicon were computed as a function of doping density and temperature in [26], assuming that the mobility is limited by the ionized-impurity scattering and phonon scattering. The results indicate that the temperature dependence of the mobility is inversely proportional to the doping density.

For doping densities where the overall mobility is limited by phonon scattering, the mobility improves with cooling [106]. However, beyond a certain doping level, the ionized-impurity scattering is strong enough to degrade the mobility according to Matthiessen's rule [84]. At this point, the overall mobility becomes weakly temperature dependent, as the temperature dependence of the ionized-impurity and phonon scattering mechanisms cancel out. Therefore the electron and hole mobilities in modern SiGe HBTs are expected to be weakly temperature dependent, as the doping levels are on the order of 10^{18} [26].

Aside from mobility, the saturation velocity is also critical, as it determines the transit time across the base-collector space-charge region. Drift velocity measurements and monte carlo simulation results indicate that the saturation velocity of silicon and germanium should improve upon cooling [107], which would reduce the total transit time at low temperatures.

3.1.3 Junction Built-in Potential

As demonstrated in Chapter 2, the junction built-in potential (ϕ_{bi}) is critical for the turn-on voltage and the depletion capacitance. As ϕ_{bi} is expected change with temperature, understanding this dependence is critical for predicting the device performance at cryogenic temperatures.

3.1.3.1 Schottky Potential

The gate metal and barrier layer of a HEMT forms a metal-semiconductor junction whose built-in potential can be written as [108]

$$\phi_{\mathrm{bi}} = \phi_{\mathrm{Bn}} - \frac{kT}{q} \ln \left\{ \frac{N_{\mathrm{C}}}{N_{\mathrm{D}}} \right\}, \tag{3.2}$$

where ϕ_{Bn} is the peak barrier height in the semiconductor region, N_{C} is the conduction band density of states and N_{D} is the donor concentration. Thus, ϕ_{bi} has a negative slope with respect to the thermal excitation. Therefore, the threshold voltage (V_{TN} of the channel and gate Schottky diode is expected to increase with cooling.

3.1.3.2 p-n Junction Potential

The built-in potential of an abruptly doped p-n junction is given as [77]

$$\phi_{\mathrm{bi}} = \frac{kT}{q} \ln \left\{ \frac{N_{\mathrm{D}} N_{\mathrm{A}}}{n_{\mathrm{io}}^2} \right\}, \tag{3.3}$$

where N_{A} is the acceptor concentration. For the doping levels considered in modern SiGe HBTs, the ionization ratio is not expected to drop significantly with cooling [109]. However, as described earlier, the intrinsic carrier concentration reduces significantly with cooling, which results in an increase of ϕ_{bi}. Therefore, the turn-on voltage of a SiGe HBT is higher at cryogenic temperatures.

3.2 DC and AC Performance

Once the key changes to the device physical properties with cooling are understood, the expected temperature dependence of the intrinsic DC and RF performance is summarized. For both transistors, the terminal currents and the associated conductances are evaluated first. The charge-control and feedback capacitances are then discussed. Since the extrinsic resistances mainly depend on the device geometry, their temperature dependence will be discussed once the model extraction results are presented.

3.2.1 I-V and C-V Characteristics of HEMT

For a given overdrive voltage ($V_{\mathrm{GS}} - V_{\mathrm{T}}$), the drain-source current (I_{DS}) is expected to increase with cooling, as the mobility and nominal gate capacitance (C_{G0}) increases. The gate current (I_{G}) is expected to decrease significantly with cooling as the saturation current of a Schottky diode is proportional to $T^2 e^{-q\phi_{\mathrm{Bn}}/kT}$ [77].

For a given ($V_{\mathrm{GS}} - V_{\mathrm{T}}$), the transconductance ($g_{\mathrm{m,HEMT}}$) is expected to improve with cooling. However, under cryogenic operation, $g_{\mathrm{m,HEMT}}$ will saturate at a lower ($V_{\mathrm{GS}} - V_{\mathrm{T}}$) as E_{crit} decreases with cooling. For a given I_{DS}, the output conductance (g_{ds}) is expected to increase with cooling as the saturation velocity improves [73]. The electron velocity improves with cooling, which results in less charge being required in the channel for a fixed I_{DS}.

The gate-source capacitance (C_{GS}) is expected to increase with cooling proportional the improvement in $g_{\mathrm{m,HEMT}}$ as $C_{\mathrm{GS}} \approx g_{\mathrm{m,HEMT}} L_{\mathrm{G}}/v_{\mathrm{avg}}$ [74]. However, the increase in C_{GS} with cooling is expected to be less than that of $g_{\mathrm{m,HEMT}}$, as the channel transit time reduces with cooling. For a given V_{GD}, the gate-drain capacitance (C_{GD}) is expected to decrease with cooling due to an increase in the junction built-in potential increases.

3.2.2 Terminal Currents and Junction Capacitances of HBT

The performance of an HBT is strongly coupled to the ambient temperature. Thus, changes in the performance of an HBT with cooling are expected to be much greater than that of a HEMT. The base saturation current (J_{B0}) is proportional to n_{io} and reduces greatly with cooling [85]. The hole diffusion length in the emitter (L_{PE}) is proportional to $\sqrt{T}$ which contributes to the decrease of J_{B0} with cooling [77]. The collector saturation current (J_{C0}) is also proportional to n_{io} but it is exponentially enhanced by the band offset created by the germanium content in the base [88]. The temperature dependent influence of the band offset on J_{C0} can be written as [26]

$$\frac{J_{\mathrm{C0}}(T)}{J_{\mathrm{C0}}(300)} \propto \frac{e^{\left(\Delta E_{\mathrm{g,app}}+\Delta E_{\mathrm{g,Ge}}(0)\right)/kT}}{e^{\left(\Delta E_{\mathrm{g,app}}+\Delta E_{\mathrm{g,Ge}}(0)\right)/k300}}. \tag{3.4}$$

While the exact temperature dependence of $\Delta E_{\mathrm{g,app}}$ and $\Delta E_{\mathrm{g,Ge}}(0)$ is hard to predict [26], Equation (3.4) indicates a much greater enhancement of J_{C0} due to the band offset at lower temperatures. As the DC curent gain (β) is ideally equal to $J_{\mathrm{C0}}/J_{\mathrm{B0}}$, it is expected to improve proportionally to the band offset enhancement of J_{C0} with cooling.

For a given collector current (I_{C}), the transconductance of an HBT is inversely proportional to the collector current ideality factor (n_{c}) and the thermal voltage (V_{TH}). As V_{TH} reduces significantly with cooling, changes in n_{c} are critical for predicting $g_{\mathrm{m,HBT}}$ at cryogenic temperatures. As it turns out, the $g_{\mathrm{m,HBT}}$ improvement with cooling is limited by an increase in n_{c} [26], which will be discussed in the next section. For a given I_{C}, the input conductance of HBT (g_{be}) is expected to reduce with cooling, as the β_{AC} improvement is greater than that of $g_{\mathrm{m,HBT}}$.

Similar to the C_{GS}-$g_{\mathrm{m,HEMT}}$ relationship, the base-emitter diffusion capacitance ($C_{\mathrm{BE,diff}}$) relates to $g_{\mathrm{m,HBT}}$ through a delay term ($C_{\mathrm{BE,diff}} \approx g_{\mathrm{m,HBT}}\tau_{\mathrm{f}}$). For devices with high β, τ_{f} is dominated by the base transit time (τ_{b}), which is mainly determined by the electron mobility in the base (μ_{nb}) [16]. As μ_{nb} is nearly temperature independent

due to the high doping levels [26], τ_f is expected to be insensitive to cooling. As a result, the temperature dependence of $C_{BE,diff}$ is expected to be very similar to that of $g_{m,HBT}$. For a given value V_{BC}, the base-collector capacitance (C_{BC}) is expected to decrease with cooling, as the junction built-in potential increases and the impurity ionization decreases [109].

3.3 Nonideal Transport Phenomena

As discussed in Chapter 2, the special effects such as surface trapping and impact-ionization in a HEMT and non-equilibrium base transport in an HBT results in deviations from the ideal device performance. As it turns out, these properties have a significant temperature dependence, which will be studied here to have a better understanding of limitations to the device performance at cryogenic temperatures.

3.3.1 Kink Effect in HEMTs

At high drain-source voltages (V_{DS}), an abnormal increase in I_G and g_{ds} is observed for short-channel HEMTs [110]. This behavior has been attributed to a combination of effects such as surface trapping and impact ionization. Recently, a study on understanding the temperature and epitaxy dependence of these individual effects using pulsed-DC measurements revealed the source of these effects [82].

The increase in I_G with V_{DS} in a certain V_{GS} range (bell shape) is due to impact ionization [111]. Impact ionization describes the generation of electron-hole pairs when electrons in the channel with high energy levels collide with the lattice. The holes are then swept to the gate by the electric field perpendicular to the channel. As a result, I_G increases. While the increase in I_G becomes less pronounced with cooling, the increase in g_{ds} becomes stronger, which points to a separate mechanism [44, 82].

The kink in the I_{DS}-V_{DS} relationship at cryogenic temperatures was found to be mainly due to the traps in the buffer or barrier layer interface [112]. To support this

claim, a GaAs mHEMT with an additional metamorphic buffer layer was compared to an InP HEMT with identical features [82]. The kink was found to be less significant for the GaAs mHEMT and could be completely eliminated with a pulse length that is not short enough for the InP HEMT [82].

The impact of this phenomena on this work was that V_{DS} had to be chosen carefully in order to prevent the increase of I_{G} at room temperature and g_{ds} at cryogenic temperatures.

3.3.2 Non-Equilibrium Transport in HBT

If the base-width is shorter than the carrier mean-free path, ballistic (scattering-free) transport is expected in the base region. The carrier mean-free path is inversely proportional to temperature [94]. Thus, even if the base-width is not short enough for ballistic transport at room temperature, the device might be subject to ballistic transport at cryogenic temperatures, as the carrier mean-free path can exceed the physical base-width.

The experimental data presented in [26] suggests that ballistic transport takes place at cryogenic temperatures for modern SiGe HBTs. The indication of ballistic transport is that the value of ideality factor (n_{c}) at very-low injection increases upon cooling. The implication of this increase is that the electrons gain enough energy during the ballistic transport such that their effective temperature (T_{eff}) is higher than the lattice temperature. As a result, the device operates as if it was subject to a higher ambient temperature. Regarding the small-signal noise performance, the biggest drawback of non-equilibrium transport is the limited improvement to g_{m} with cooling. Furthermore, the rate at which n_{c} increases with bias becomes worse with cooling which can further limit the device performance at high current densities.

Although the holes are not subject to high electric-field transport while diffusing from the base to emitter, the experimental data indicates that the effective hole

temperature exceeds that of the electrons at cryogenic temperatures [26]. As discussed in Chapter 2, the increase in the effective hole temperature is explained by tunneling and field-assisted recombination mechanisms [16].

3.4 Dynamic Range

In this section, changes in the transport properties and terminal characteristics with cooling are utilized to evaluate temperature dependence of the dynamic range performance.

3.4.1 Noise

For both devices, cooling results in a reduction of thermal noise generated by the series access resistances, resulting in an improvement of the broadband noise performance. Aside from this, the temperature dependence of the intrinsic noise sources will further determine how the noise performance changes with cooling.

The limit to low-frequency noise ($T_{\mathrm{MIN,LF}}$) achievable by HEMT and HBT devices are dominated by the gate and base currents, respectively. As these drop significantly with cooling for a given drain or collector current, $T_{\mathrm{MIN,LF}}$ is expected to improve significantly with cooling.

For HEMT, the gate-induced noise is expected to decrease with cooling as it is assumed to be thermal [81]. However, the temperature dependence of the hot-electron noise due to the velocity saturated carriers is difficult to predict and expected to be partly influenced by inter-valley scattering [80]. Accounting for the inter-valley scattering, an empirical fit to the Monte Carlo simulations provided in [113] for MESFETs predicts the effective hot-electron noise temperature as

$$T_{\mathrm{n}} = T\left(1 + \gamma\left(\frac{E}{E_{\mathrm{crit}}}\right)^{n}\right), \tag{3.5}$$

where E is the applied electric-field, γ is a constant of 6-10 (GaAs-InP) at 300 K and 20-38 (GaAs-InP) at 77 K, and n is a constant of 3 at 300 K and 2 at 77 K. The result of this fit and the Monte Carlo simulations indicate that the hot-electron noise is weakly temperature dependent for $E \approx E_{\mathrm{crit}}$. However, since the electric-field in the channel greatly varies from the source to drain, the exact temperature dependence of T_{n} can not be estimated without the knowledge of the electric-field profile in the channel.

Recalling from Chapter 2 that the overall noise of HEMT is expected to be proportional to $\sqrt{I_{\mathrm{DS}}}/g_{\mathrm{m}}$, an improvement in the overall noise performance with cooling is expected as the device will have a higher g_{m} for a given I_{DS}. The high frequency noise of an HBT is expected to decrease with cooling parallel to the f_{max} improvement [26].

3.4.2 Linearity

While the reduction of noise with cooling is beneficial for receiver sensitivity, the temperature dependence of the dynamic range also depends on the temperature dependence of transistor linearity. With the knowledge of how the physical parameters, I-V curves, and C-V curves change with cooling, the temperature dependence of the nonlinear behavior can be estimated.

As discussed in Chapter 2, the nonlinearity of the charge-control capacitance (C_{GS}, C_{BE}) is expected to have a similar bias and temperature dependence to that of g_{m}. As long as a HEMT is biased in deep-saturation and its V_{DS} is not at the kink point, no significant change to the g_{ds} nonlinearity with cooling is expected since I_{DS} varies linearly with V_{DS} as demonstrated in Chapter 2. Although the bias dependence of β increases with cooling, the nonlinearity of g_{be} is also expected to have a fairly similar temperature trend to that of g_{m}.

For a given overdrive voltage and fixed gate-length (L_{G}), the nonlinearity associated with the transconductance of a HEMT is only a function of E_{crit} as shown in

Chapter 2. Referring back to Figure 2.4, the nonlinearity associated with $g_{\mathrm{m,HEMT}}$ is expected to increase with cooling at low overdrive levels as E_{crit} decreases. Oppositely, at high overdrive levels, the transconductance of a HEMT is expected to be more linear at cryogenic temperatures as E_{crit} decreases with cooling. In general, HEMTs are typically operated at low current densities for low-noise applications which suggests that the $g_{\mathrm{m,HEMT}}$ and C_{GS} nonlinearities increase upon cooling.

The nonlinearity associated with the transconductance of an HBT is expected to inevitably increase with cooling, as V_{TH} is reduced. However, the increase in n_{c} with cooling is expected to limit the degradation in $g_{\mathrm{m,HBT}}$ linearity. Furthermore, the increasing bias sensitivity of n_{c} with cooling can potentially influence the overall shape of the $g_{\mathrm{m,HBT}}$ nonlinearity as a function of I_{C}. The nonlinearities associated with C_{BE} and g_{be} are expected to increase proportionally to the $g_{\mathrm{m,HBT}}$ nonlinearity with cooling.

For both devices, the feedback capacitance (C_{GD}, C_{BC}) is expected to become more linear with cooling due to the ϕ_{bi} increase. However as the turn-voltages are higher at cryogenic temperatures, the applied reverse voltage on the feedback capacitance will decrease with cooling if the output voltages (V_{DS}, V_{CE}) are kept constant. This may somewhat neutralize the benefit of the increase in ϕ_{bi} with cooling.

Thus, the general trend is that the nonlinearities that are modulated by the input voltage (V_{GS}, V_{BE}) should become stronger with cooling whereas the nonlinearities that are modulated by the output voltage should become weaker. The effect of these two sets of nonlinearities on the overall intermodulation performance depend on the termination impedances. Therefore, it might be possible to tweak the termination impedances such that the trade-off between the input and output modulated nonlinearities is optimized to yield the best possible performance at different temperatures.

3.5 Summary

The temperature dependence of the key physical parameters and the expected dynamic range performance has been discussed in this chapter. For HEMTs, the intrinsic noise performance is expected to improve with cooling as the gate current and electron transit time in the channel reduces with cooling. However, the decrease in the critical electric field with cooling results in a more nonlinear transconductance. For HBTs, the noise performance is greatly improved with cooling due to the exponential enhancement of the DC current gain and the enhancement of transconductance with reducing thermal voltage. However, the reducing thermal voltage also results in a significantly more nonlinear transconductance. For both devices, the feedback capacitance and its nonlinearity is expected to decrease with cooling, as the junction built-in potential increases. In the following chapters, these expectations will be compared against the experimental modeling results.

CHAPTER 4

EXPERIMENTAL SETUP

The device specific details and the measurement setup are presented in this chapter. As the fundamental principles of HEMTs and HBTs regarding the dynamic range limitations were studied previously, the purpose of this chapter is to enable a smooth transition from the theoretical analysis to the experimental results.

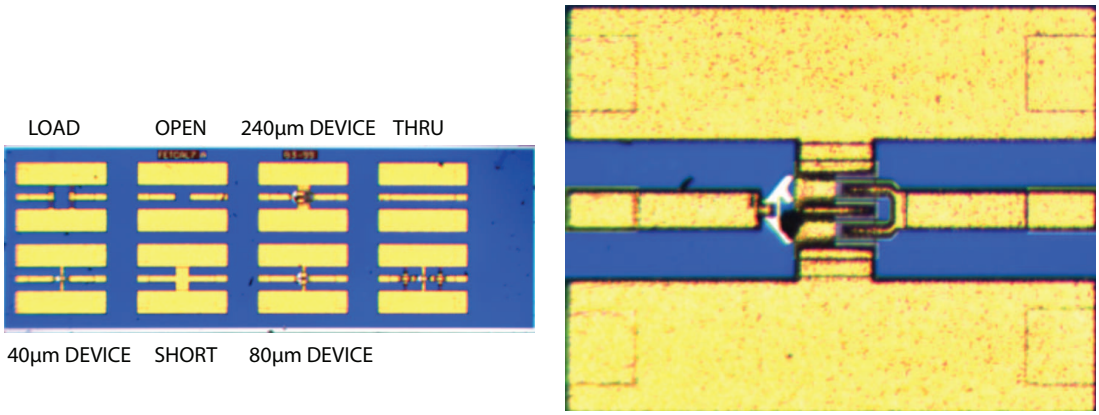
4.1 Device Description

In this section, the physical structures of the devices that were characterized are explained. Since the previous discussion of general HEMT and HBT operation was mostly notional, it is important to provide a visual perspective into the key features of each transistor that will be evaluated.

4.1.1 HEMT

0.1 μm gate-length InP HEMTs fabricated by Northrop Grumman Space Technology (formerly TRW Space Technology Division) were used for the HEMT analysis. This fabrication process enables the design of cryogenic low-noise amplifiers with 35 K average noise temperature at W-band [114] and achieves an f_{t} and f_{max} of 180 and 350 GHz at room temperature, respectively [115]. The test coupon which includes open, short, load and thru calibration structures is shown in Figure 4.1.

Layer profile and cross-sectional STEM images of the HEMT are shown in Figure 4.2. A series of trade-offs in the process parameters regarding the device performance can be summarized as follows:



(a) (b)

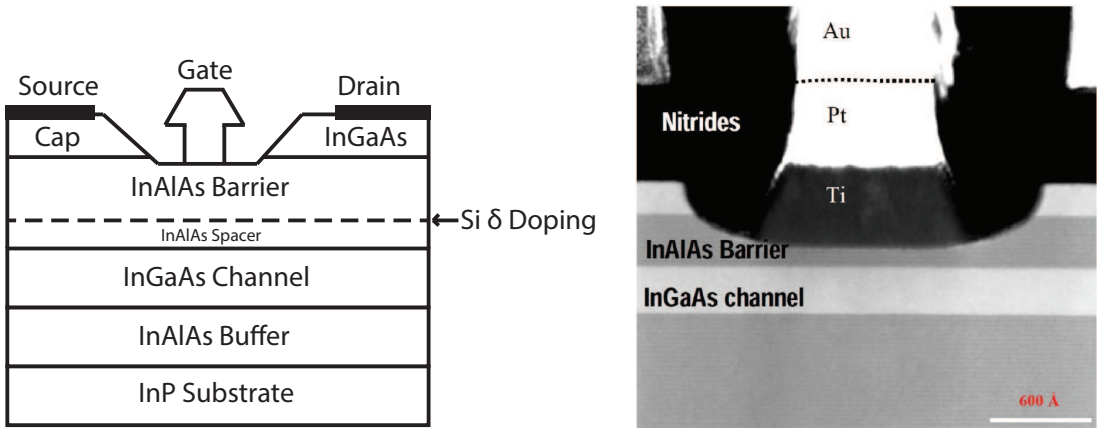
Figure 4.1: (a) Photo of the HEMT test coupon. (b) Close-up photo of the 240 μm gate-width device test structure.

- Reducing the barrier thickness improves the transconductance and the access resistances at the cost of degradation in other metrics such as gate capacitance, breakdown voltage, and gate leakage current [116].

- The thickness of the spacer layer is critical for separating the delta doping from the 2-DEG while maximizing the sheet carrier concentration (n_s) [53] and preventing parasitic channel formation [103].

- Higher indium content in the channel increases the mobility [117]. However, there is an upper limit due to the mismatch induced strain and lattice dislocation.

The NGST HEMT has a 60 % indium mole fraction employed in the channel with n_s equal to 3.5×10^{12} cm^{-2} and an electron mobility greater than 9,000 cm^2/V-s at room temperature [119]. Compared to the low-noise InP HEMT presented in [103] which had a 65 % Indium mole fraction, the mobility of the NGST HEMT is lesser, while the n_s is greater. Although increasing the Indium mole fraction helps with the f_t of the devices, no improvement was observed in the noise performance [120].



(a) (b)

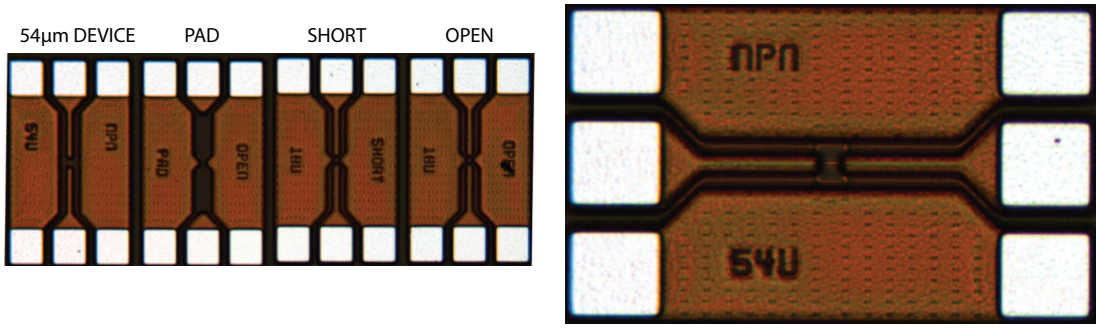
Figure 4.2: (a) Typical layer profile of state-of-the-art InP HEMT. Layers are not drawn to scale. (b) Cross-sectional STEM image of gate region of the NGST InP HEMT [118] © 2005 IEEE.

4.1.2 HBT

0.12 μm emitter-width SiGe HBTs from IBM's (now Global Foundries) BiCMOS8HP process were characterized. The process also offers integrated 0.13 μm CMOS transistors which can be used to realize digital control and processing on the same chip with RF building blocks. The HBT has a room temperature f_t and f_{max} of 200 and 280 GHz, respectively [121]. Photographs of the device and de-embedding structures, which include open, short and pad-open test sites, appear in Figure 4.3.

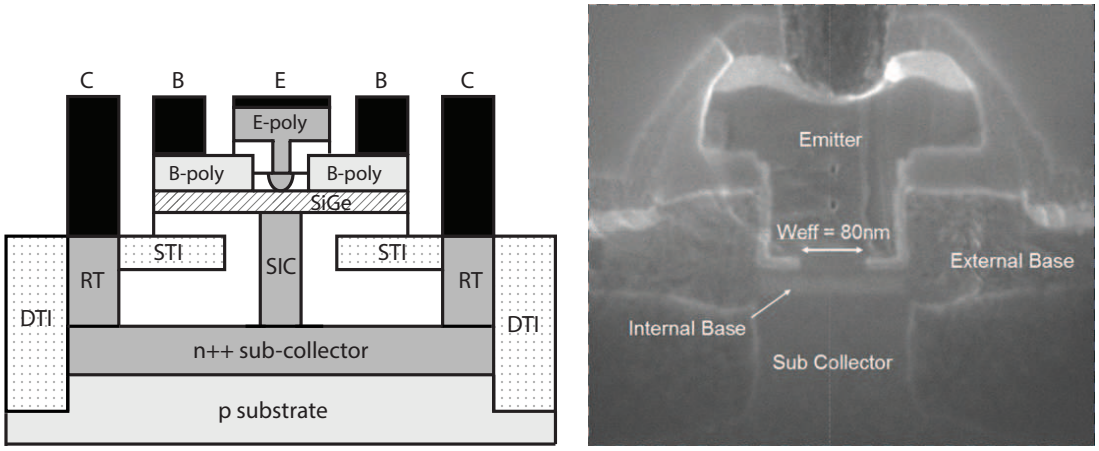
The layer profile and cross-sectional STEM image of the HBT are shown in Figure 4.4. Since an HBT has current flow in vertical direction, vertical scaling of the physical device parameters are very critical to the device performance. Reducing the thickness of the intrinsic device layers lowers the transit times, which improves f_t [16]. Increasing the Ge gradient across the neutral base further reduces the base transit time [86].

The most basic trade-off in the fabrication of HBTs is base doping. A higher base doping reduces the base resistance, which is beneficial for the high frequency small-signal and noise performance. However, it also reduces the DC current gain which



(a) (b)

Figure 4.3: (a) Photo of the SiGe HBT test structures. (b) Close-up photo of the 54 μm emitter length device test structure.



(a) (b)

Figure 4.4: (a) Typical layer profile of state-of-the-art SiGe HBT [26]. Layers are not drawn to scale. (b) Cross-sectional STEM image of a SiGe HBT [122] © 2011 IEEE.

would increase the low frequency noise floor. As explained in Chapter 2, germanium content in the base significantly alleviates this trade-off.

Doping of the selectively implanted collector (SIC) determines the trade-off between the peak f_t, C_{CB}, and the collector-emitter breakdown voltage B_{VCEO}. A higher doping level reduces the thickness of the collector-base space-charge region (SCR) and delays the onset of Kirk effect which improves the peak f_t [92]. However, it also increases the C_{CB} which is detrimental to the low current density f_t. Increasing the SIC doping reduces the collector-emitter breakdown voltage due to a higher impact ionization rate [16].

In addition to vertical scaling, lateral scaling of the HBT is also critical for reducing the base resistance [86], which significantly influences the noise performance and f_{max}. However, overlap between the extrinsic base and the SIC is expected to increase with lateral scaling due to out-diffusion that occurs during the high temperature fabrication process steps [123]. This will result in a higher C_{CB}. Ultimately, lateral and vertical scaling is a joint effort to optimize the device performance.

4.2 Measurement Setup

A block diagram of the measurement setup appears in Figure 4.5. On-wafer device measurements were performed using Lakeshore CRX-4K cryogenic probe station [124], which is shown in Figure 4.6. Closed-cycle helium circulation, provided by a Sumitomo F-70 compressor [125], cools the chuck to 4 K. Precise temperature monitoring and control in the range of 4 - 300 K is achieved by Lakeshore Model 336 cryogenic temperature controllers, which are linked to the sensors and heaters inside the station. Probe mounts are heat-sunk to the chuck in order to minimize the temperature difference between the reticule and the probe. Despite their higher loss, stainless-steel cables were used between the probe and the external connector to maintain a better thermal isolation between the temperature extremes.

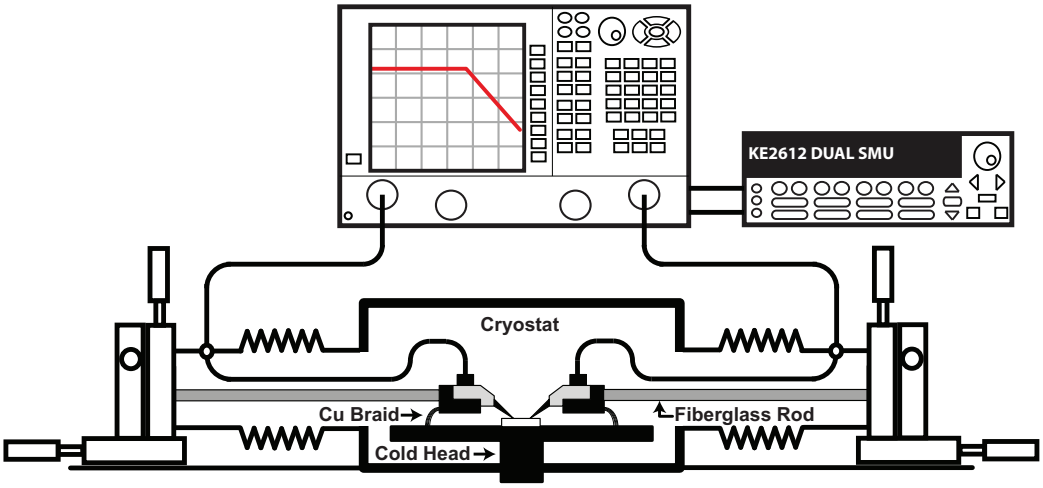
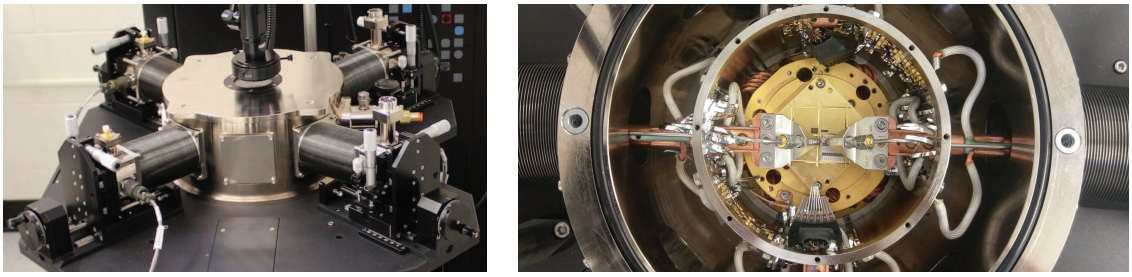


Figure 4.5: Block diagram of the measurement setup.



(a) (b)

Figure 4.6: (a) Side and (b) top view of the cryogenic probe station.

Scattering parameter (s-parameter) measurements were performed from 0.01–67 GHz with an Agilent (now Keysight) N5247A PNA-X vector network analyzer (VNA) [126]. The biases were provided with a Keithley 2612 dual SMU [127] through bias-tees internal to the network analyzer. Cryogenic probes manufactured by SUSS (now Cascaded Microtech) were used for RF probing [128]. For HEMT measurements, broadband SOLT calibrations were performed with the short, open, load and thru structures located on the test coupon. This sets the measurement reference plane to the device terminals. For the HBT measurements, broadband LRRM calibrations [129] were performed with a CSR-8 impedance standard substrate (ISS) [130], bringing the measurement reference plane to the probe tips. Following the calibration, a pad-open-short (POS) [131] deembedding procedure was performed to move the reference plane to the device terminals. Resistance values obtained from measurements of the short structure were used to correct DC voltages at the device terminals for the series and ground resistances.

OIP3 measurements of devices terminated in 50 Ω were performed on-wafer and compared to the model simulations to validate the non-linear models. The models and comparison results will be presented in Chapter 6. The OIP3 measurements were also done with the VNA [132]. For this, a separate channel was created on the VNA and the s-parameter calibration was repeated in order to account for the cable and probe losses and to correct power levels at the device terminal. Measuring the IP3 with the VNA provides significant advantages over a spectrum analyzer based measurement, which also requires signal generators. Using a single instrument rather than three dramatically reduces the complexity of the measurement setup and the amount of time required. Since no physical modifications were required to the setup between the s-parameter and the OIP3 measurements, both sets of data can be obtained without lifting the probes, which increases the lifetime of the test structures and also the consistency of the data. Furthermore, the VNA inherently provides the power

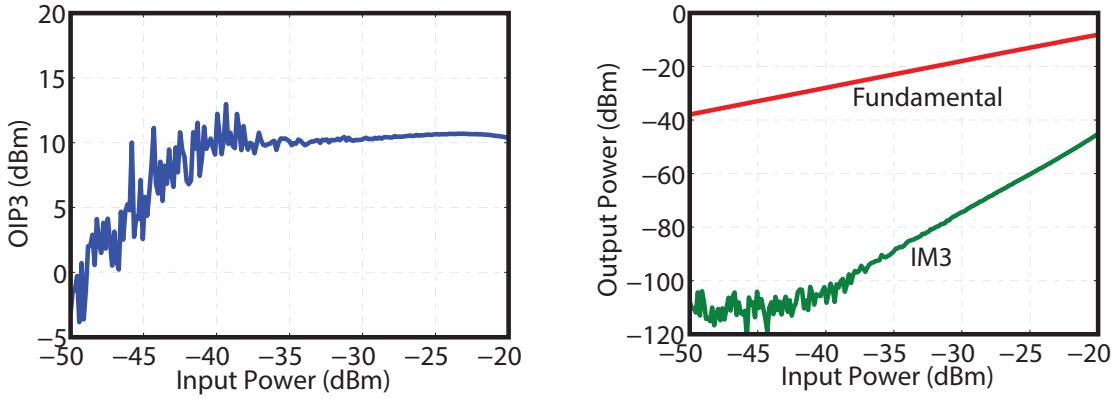


Figure 4.7: Example power sweep for IP3 measurement of the SiGe HBT.

sweep required to verify slope of the fundamental and IM3 tones, which minimizes the effort to repeat the IP3 measurements over a wide range of bias, frequency, and temperature points. An example power sweep appears in Figure 4.7.

Measurements were automated with MATLAB scripts. Device characterization was performed at 7, 77, 200 and 298 K. For all measurements in this work, the source, emitter, and substrate terminals were tied to ground. Model simulations were performed in the AWR Microwave Office environment. Device characterization was performed for a drain-source voltage (V_{DS}) and collector-emitter voltage (V_{CE}) equal to 0.6 V and 1 V, respectively. The V_{DS} of the HEMT was limited at the high end due to the significant kink effect at cryogenic temperatures. The V_{CE} of the HBT was limited at the low end due to forward biasing of the base-collector junction at cryogenic temperatures.

CHAPTER 5

SMALL-SIGNAL AND NOISE MODELING

In this chapter, the extracted small-signal noise models are presented. These models will form the basis of the non-linear models that will be presented in the next chapter. This chapter is divided into three sections. First, the model topologies and their parameter extraction techniques are explained. Second, the agreement between the models and s-parameter measurements is demonstrated. Finally, changes in the model parameters with cooling are discussed.

5.1 Model Extraction Procedure

HEMT and HBT small-signal noise models that can accurately predict the cryogenic noise performance are well established in the literature [26,133]. In this section, an overview of these models and their extraction approach are provided.

5.1.1 HEMT Small-Signal Noise Model

The small-signal noise model used for the HEMT is shown in Figure 5.1. The series gate, source, and drain access resistances (R_{G}, R_{S}, R_{D}) and the thermal noise sources ($v_{\mathrm{n,g}}$, $v_{\mathrm{n,s}}$, $v_{\mathrm{n,d}}$) associated with them are bias independent extrinsic elements. The remaining bias dependent intrinsic network describes the dynamic operation of the device and consists of the following:

1. A transconductance ($g_{\mathrm{m,HEMT}}$) with a time delay (τ) and an output conductance (g_{ds}) that describe dependence of the output small-signal current on the input and output voltages.

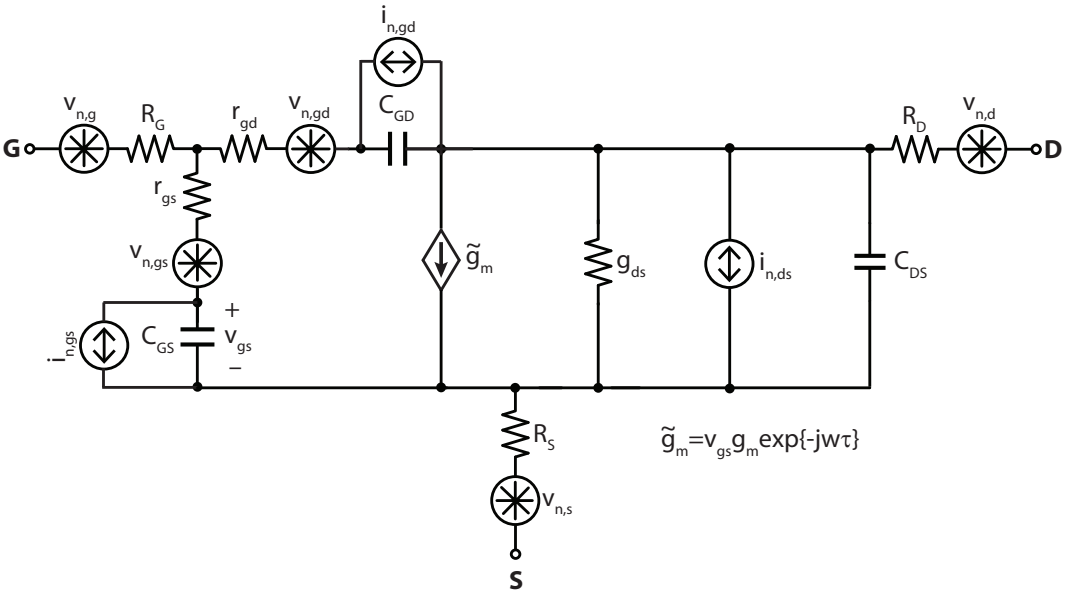


Figure 5.1: HEMT small-signal noise model.

2. Capacitances associated with charge-control (C_{GS}), depleted Schottky junction (C_{GD}), and channel-drain coupling over the depleted saturation region (C_{DS}).

3. Non-quasi-static (NQS) resistances (r_{gs}, r_{gd}) that describe the delay in charging/discharging of the corresponding capacitances and gate induced noise associated with them ($v_{\mathrm{n,gs}}$, $v_{\mathrm{n,gd}}$).

4. Shot noise sources ($i_{\mathrm{n,gs}}$, $i_{\mathrm{n,gd}}$) associated with the gate leakage current [134].

5. Channel thermal noise source ($i_{\mathrm{n,ds}}$) associated with g_{ds}.

Thermal noise associated with the parasitic resistances is proportional to the ambient temperature. The rest of the noise sources are defined empirically based on Pospieszalski's seminal work [81]. The ambient temperature (T_{A}) is assumed for the noise sources that are associated with the NQS resistances. For the channel thermal noise, an unknown elevated temperature is assumed that is often denoted as T_{drain} or T_{D}. For the specific device characterized in this work, T_{D} is estimated as

400+6×T_A [110,135]. This estimation aligns well with the previously published results of other short-channel FETs [136–138].

To extract the small-signal model, the extrinsic resistances are initially determined which will be described next. Once the extrinsic resistances are subtracted from the model, the bias dependent intrinsic elements are determined from y-parameters of the remaining network.

5.1.1.1 Determination of the Extrinsic Resistances

For model extraction, the series parasitic resistances were first determined using the cold-FET method [139]. For zero drain current, s-parameter measurements were performed as a function of gate current (I_G). Once the channel is fully turned-on, the relationship between the extrinsic resistances and the z-parameters can be written as

$$\Re\{Z_{11}\} \approx R_S + R_G + \frac{nkT}{qI_G}, \tag{5.1}$$

$$\Re\{Z_{12}\} \approx R_S, \tag{5.2}$$

and

$$\Re\{Z_{22}\} \approx R_S + R_D, \tag{5.3}$$

where n is the Schottky gate diode ideality factor. To determine the gate resistance, $\Re\{Z_{11}\text{-}Z_{12}\}$ was extrapolated to infinite I_G, where the series resistance contribution from the gate Schottky diode becomes zero. Figure 5.2 shows the z-parameters as a function of I_G. It can be seen that R_S and R_D are independent of I_G as expected.

5.1.1.2 Determination of the Intrinsic Elements

Once the extrinsic resistances have been determined and subtracted from the small-signal equivalent circuit that corresponds to the active measurement, the in-

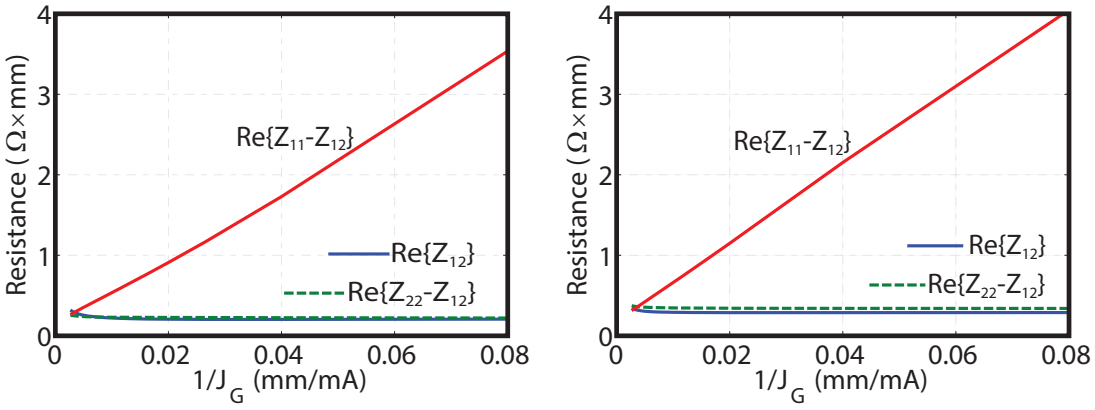


Figure 5.2: Cold-FET z-parameters at (a) 7 K and (b) 298 K.

trinsic parameters can be determined from the y-parameters of the remaining network as

$$g_{\mathrm{m,HEMT}} = |Y_{21} - Y_{12}|, \tag{5.4}$$

$$g_{\mathrm{ds}} = \Re\{Y_{22} + Y_{12}\}, \tag{5.5}$$

$$C_{\mathrm{GS}} = \frac{-1}{\omega\Im\{1/(Y_{11} + Y_{12})\}}, \tag{5.6}$$

$$C_{\mathrm{GD}} = \frac{1}{\omega\Im\{1/(Y_{12})\}}, \tag{5.7}$$

$$C_{\mathrm{DS}} = \frac{\Im\{Y_{22} + Y_{12}\}}{\omega}, \tag{5.8}$$

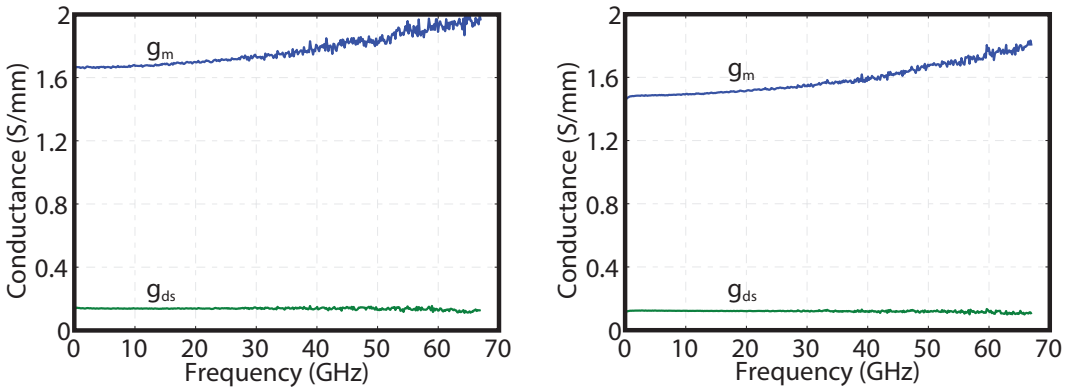
$$r_{\mathrm{gs}} = \Re\left\{\frac{1}{Y_{11} + Y_{12}}\right\}, \tag{5.9}$$

$$r_{\mathrm{gd}} = -\Re\left\{\frac{1}{Y_{12}}\right\} \tag{5.10}$$

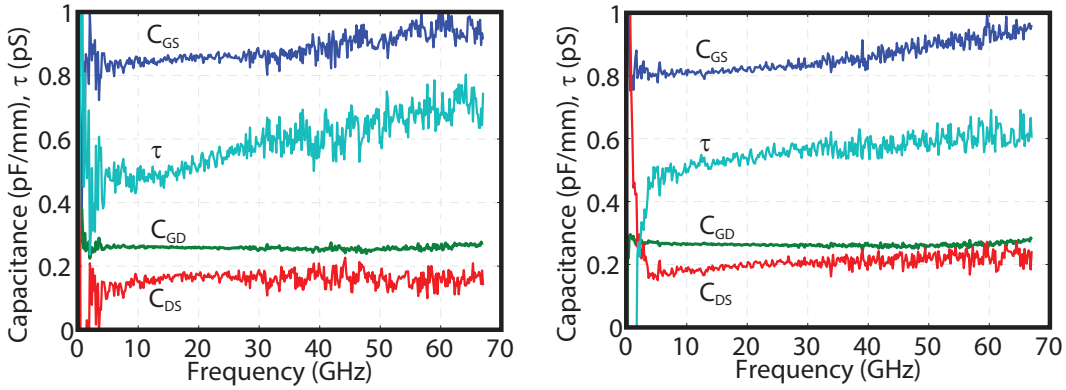
and

$$\tau = \frac{-\angle(Y_{21} - Y_{12})}{\omega}. \tag{5.11}$$

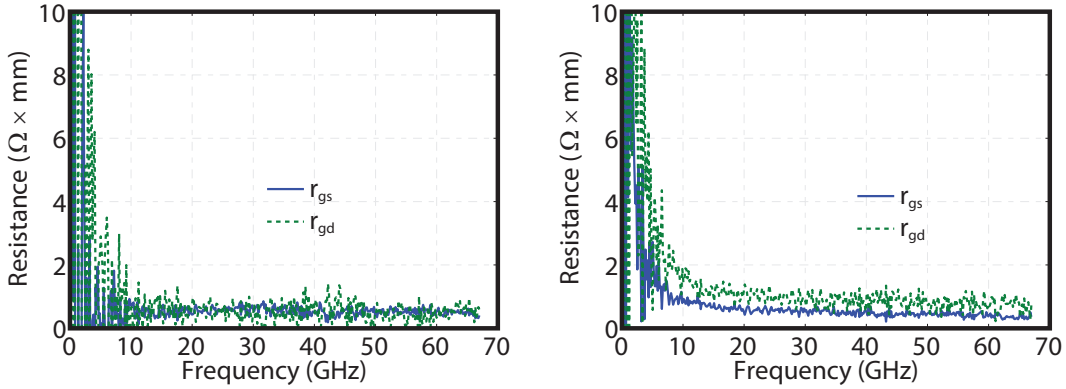
Example extraction results of the intrinsic parameters as a function of frequency are shown in Figure 5.3 for a current density of 125 mA/mm. The absolute values were determined through averaging over 15-45 GHz frequency range.



(a)



(b)



(c)

Figure 5.3: Frequency dependence of the HEMT intrinsic parameters (a) g_{m} and g_{ds}, (b) capacitances and τ, (c) NQS resistances for a current density of 125 mA/mm. The left column is 7 K and the right column is 298 K.

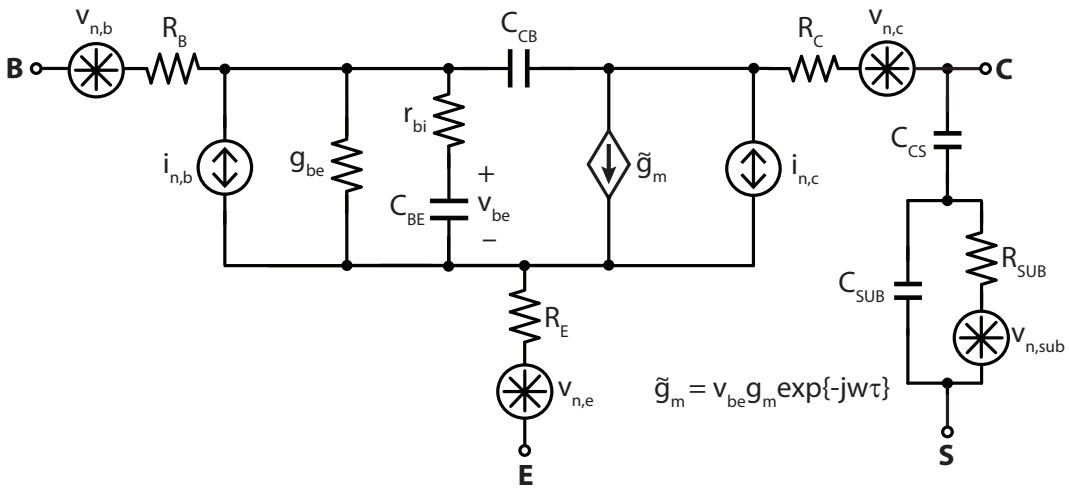


Figure 5.4: HBT small-signal noise model.

5.1.2 HBT Small-Signal Noise Model

The small-signal noise model used for HBT modeling is shown in Figure 5.4. The model includes a substrate network ($C_{\rm CS}$, $C_{\rm SUB}$, $R_{\rm SUB}$), which represents the coupling between the collector and substrate. Similar to the HEMT, the parasitic emitter and collector resistances ($R_{\rm E}$, $R_{\rm C}$) are bias independent, whereas the base resistance ($R_{\rm B}$) is bias dependent. $R_{\rm B}$, $R_{\rm E}$ and $R_{\rm C}$ generate thermal noise which are included in the model ($v_{\rm n,b}$, $v_{\rm n,e}$, $v_{\rm n,c}$).

Compared to HEMT, the HBT has a non-zero input conductance ($g_{\rm be}$) at DC due to the forward-biased base-emitter junction. On the other hand, the HBT has a high output impedance which is negligible for the $V_{\rm CE}$ considered in this work. $C_{\rm CB}$ is a depletion capacitance that represents the reverse-biased base-collector junction, whereas $C_{\rm BE}$ represents the diffusion and depletion capacitances of the base-emitter junction. A noiseless NQS resistance ($r_{\rm bi}$) is included in the model to capture the input NQS effect [140,141]. It was observed that $r_{\rm bi}$ is critical for accurately predicting the noise and high frequency s-parameters simultaneously. Finally, shot noise due to the base and emitter junction currents ($i_{\rm n,b}$, $i_{\rm n,c}$) is included in the model.

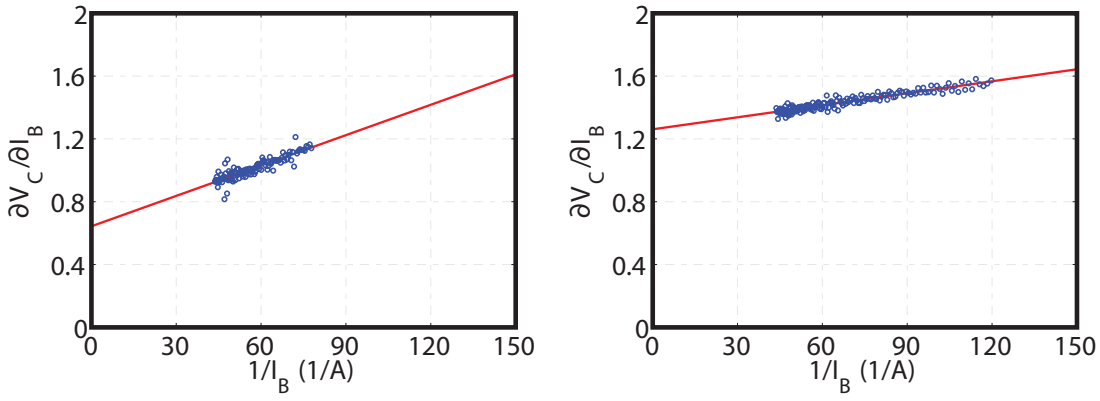


Figure 5.5: R_E extraction of a 2.16 μm^2 device using the open-collector method at (a) 7 K and (b) 298 K.

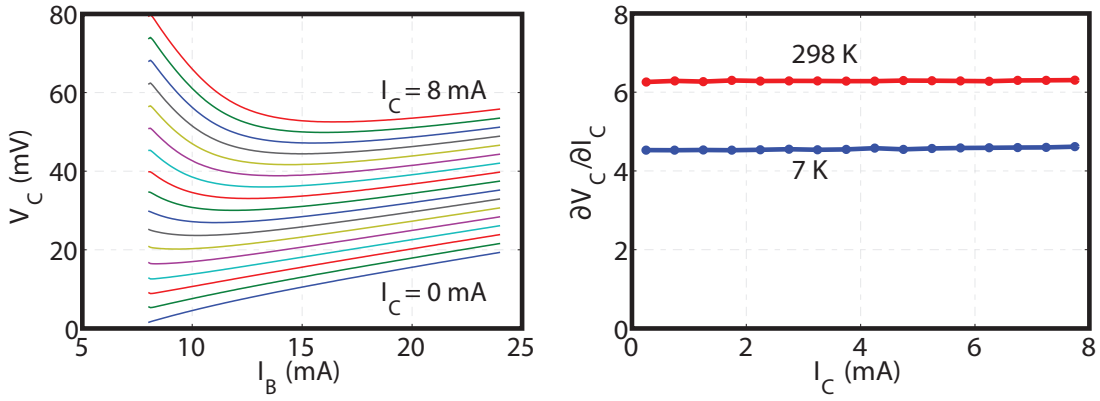
5.1.2.1 Determination of the Emitter and Collector Resistances

Extraction of the HBT small-signal model starts with determining R_E and R_C using the DC open-collector method [142]. To find R_E, the collector current (I_C) is forced to zero and the base current (I_B) is swept. The slope of the collector voltage (V_C) as a function of $1/I_B$ is then extrapolated to infinite I_B. This intercept is approximately equal to R_E. Example extraction results are shown in Figure 5.6.

The extrinsic collector resistance is also determined using the DC open-collector method [142]. To find R_C, I_C is swept while I_B is kept constant at a high current density to ensure that the device is in deep saturation. The slope of V_C is then approximately equal to the sum of R_C, R_E and cable resistance (R_{cable}), where R_E and R_{cable} are already known.

5.1.2.2 Determination of the Substrate Network

The substrate network is determined prior to extraction of the intrinsic elements. Since the collector-substrate junction is a reverse biased p-n junction, its depletion capacitance is bias dependent and thus needs to be determined for the exact V_C at which the device is operated at. Therefore, s-parameter measurements were performed



(a) (b)

Figure 5.6: R_{C} extraction of a 2.16 μm^2 device using the open-collector method. (a) V_{C} as a function of I_{B} at different forced I_{C} levels. (b) Slope of V_{C} with respect to I_{C} for I_{B}=24 mA.

for the off device with nominal V_{CE} and 0 V_{BE}. The small-signal model corresponding to this biasing scheme is shown in Figure 5.7. The substrate network can then be defined from y-parameters of the measurement as

$$Y_{\mathrm{SUB}} = Y_{12} + Y_{22}. \tag{5.12}$$

At low frequencies, imaginary part of Y_{SUB} is dominated by C_{CS} [143]. Therefore

$$C_{\mathrm{CS}} = \lim_{\omega \to 0} \frac{1}{\omega} \Im \{Y_{\mathrm{SUB}}\} \tag{5.13}$$

Once C_{CS} is subtracted from the network, the remaining elements can be defined as

$$C_{\mathrm{SUB}} = \frac{1}{\omega} \Im \{Y'_{\mathrm{SUB}}\} \tag{5.14}$$

and

$$R_{\mathrm{SUB}} = \frac{1}{\Re \{Y'_{\mathrm{SUB}}\}}. \tag{5.15}$$

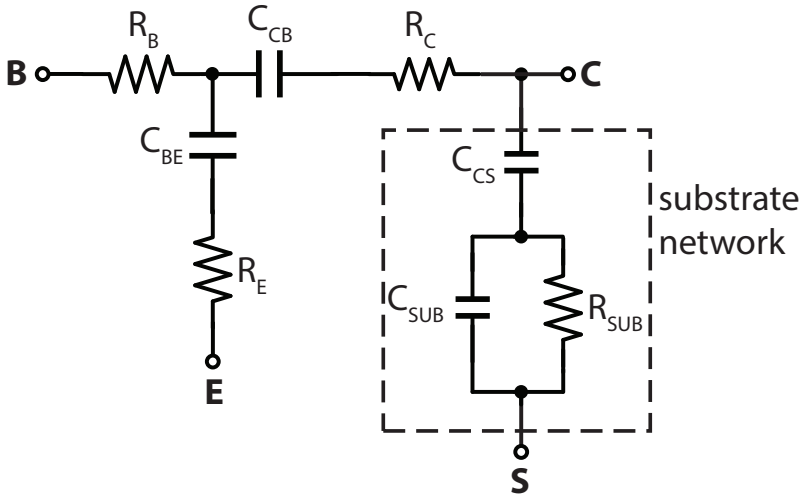


Figure 5.7: HBT small-signal model in the off-state.

The total capacitance of the substrate network ($\Im\{Y_{\mathrm{SUB}}\}/\omega$) and R_{SUB} as a function of frequency are shown in Figure 5.8. It can be seen that, at room temperature, the total capacitance rapidly drops with frequency due to the high substrate resistance. As the device is cooled down, R_{SUB} reduces which results in a milder total capacitance roll-off with frequency. This reduction in R_{SUB} is counterintuitive since the resistance of silicon is expected to increase with cooling if the doping level is below the Mott transition [144], which holds true for the substrate [26].

5.1.2.3 Determination of the Base Resistance and Input Conductance

Once R_{E}, R_{C}, and the substrate network are subtracted from the small-signal equivalent circuit that corresponds to the active measurement, R_{B} can be determined as

$$R_{\mathrm{B}} = \Re\{Z_{11} - Z_{12}\}. \tag{5.16}$$

Alternatively, R_{B} can be extracted from y-parameters of the intrinsic network as

$$R_{\mathrm{B}} = \lim_{\omega\to\infty} \Re\{\frac{1}{Y_{11}}\}, \tag{5.17}$$

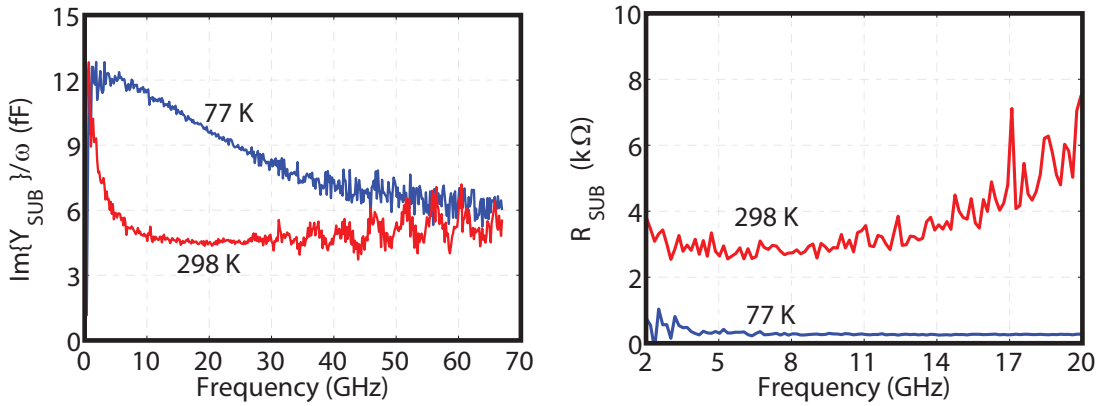


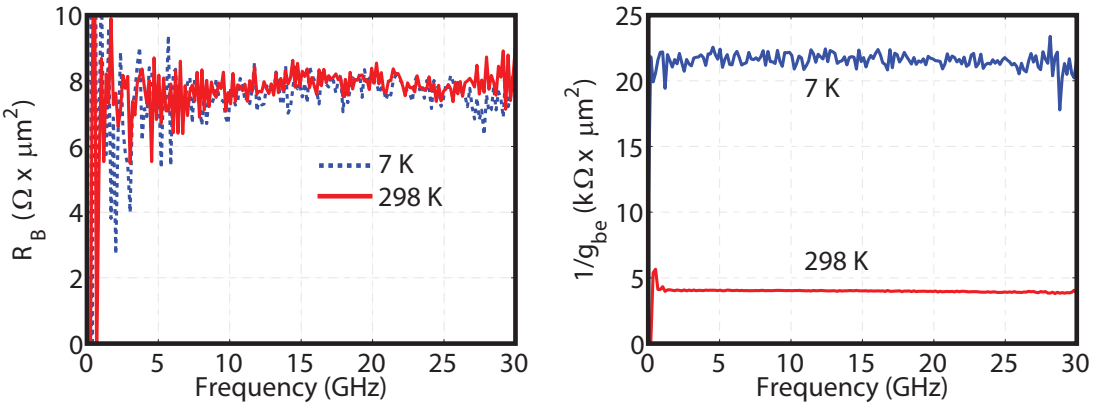
Figure 5.8: (a) Total capacitance of the substrate network and (b) R_{SUB} extraction as a function of frequency for a 2.16 μm^2 device.

which is a common approach in the literature. However, this equation integrates the NQS resistance r_{bi} into R_{B} which results in overestimation of R_{B} and yields a nonphysical R_{B} trend with respect to the current density [140].

Following extraction of R_{B}, g_{be} can be directly determined from the y-parameters of the remaining network. However, this approach can be problematic for the very small g_{be} values which are observed at cryogenic temperatures. Alternatively, g_{be} can be determined from DC data by calculating the slope of V_{BE} with respect to I_{B}. Another approach is to find g_{be} from time delay of the base-emitter network with the knowledge of AC current gain [145]. It was observed that this method gives the same result with the DC approach. Example extraction results of R_{B} and g_{be} with the time delay method are shown in Figure 5.9.

5.1.2.4 Determination of the Intrinsic Capacitances, Transconductance, Time Delay and NQS Resistance

Following the subtraction of the series resistances and substrate network from the HBT small-signal model that corresponds to the active measurement, the remaining unknown parameters can be determined from the y-parameters of the corresponding



(a) (b)

Figure 5.9: (a) R_{B} and (b) g_{be} extraction as a function of frequency for a current density of 2.2 mA/μm^2.

network as

$$g_{\mathrm{m,HBT}} = |Y_{21} - Y_{12}|, \tag{5.18}$$

$$C_{\mathrm{BE}} = \frac{-1}{\omega \Im\{1/(Y_{11}+Y_{12})\}}, \tag{5.19}$$

$$C_{\mathrm{BC}} = \frac{-\Im\{Y_{12}\}}{\omega}, \tag{5.20}$$

$$r_{\mathrm{bi}} = \Re\left\{\frac{1}{Y_{11}+Y_{12}}\right\}, \tag{5.21}$$

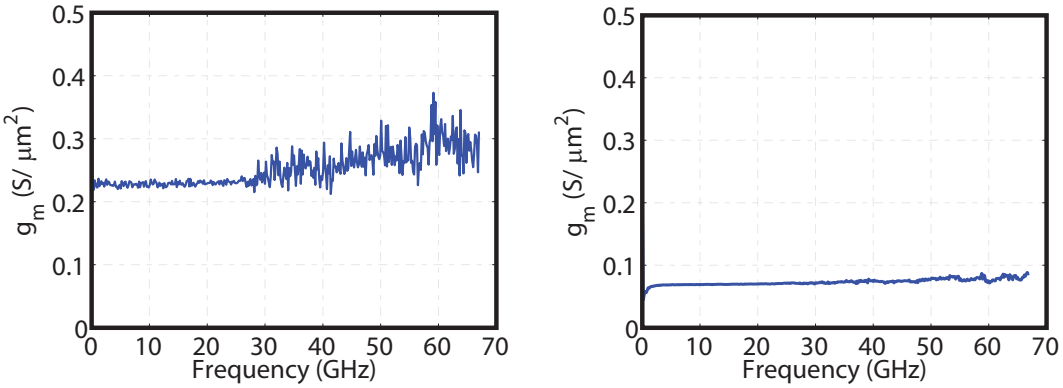
and

$$\tau = \frac{-\angle(Y_{21}-Y_{12})}{\omega}. \tag{5.22}$$

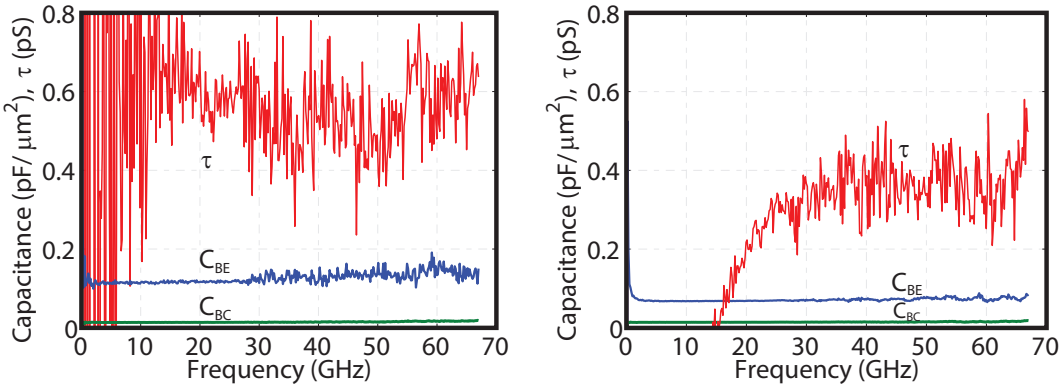
Example extraction results of the intrinsic parameters as a function of frequency are shown in Figure 5.10 for a current density of 2.2 mA/μm^2. The constant values of model parameters were determined through averaging over 10-30 GHz frequency range.

5.2 Model to Measurement Agreement

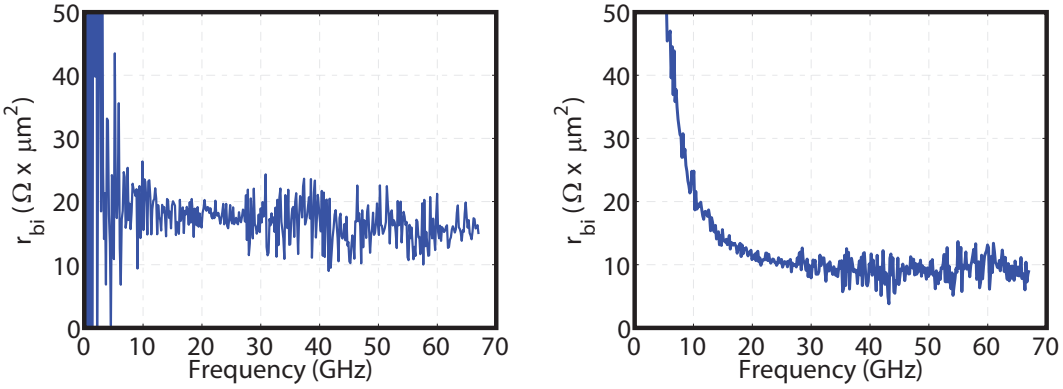
The accuracy of the models was verified following the extraction. For this, f_{t}, f_{max}, and s-parameter results of the models were compared to those of the de-embedded



(a)



(b)



(c)

Figure 5.10: Frequency dependence of the HBT intrinsic parameters (a) g_{m}, (b) capacitances and τ, (c) NQS resistance for a current density of 2.2 mA/μm^2. The left column is 7 K and the right column is 298 K.

measurements. The results are shown in Figures 5.11 and 5.12 for the HEMT and HBT devices, respectively. Good agreement is observed up to 67 GHz.

5.3 Modeling Results

The temperature and bias dependence of the model parameters are discussed and explained in terms of the device physics in this section.

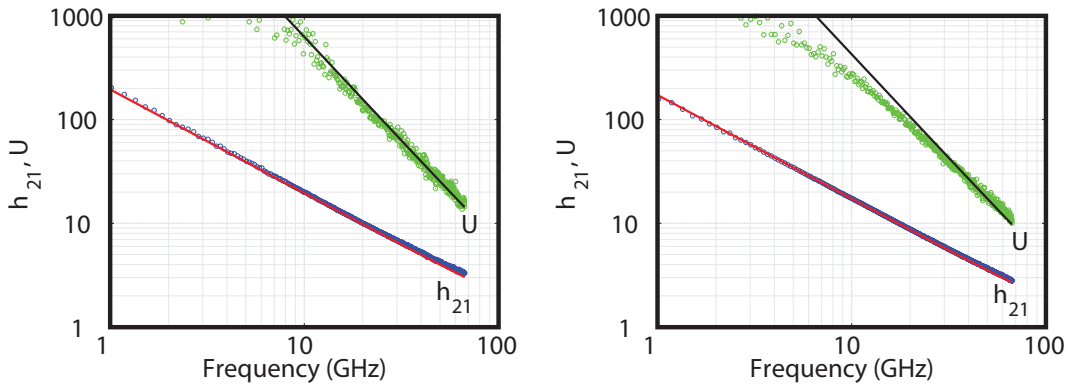
5.3.1 Gate and Base Currents

The gate current of HEMTs and the base current of the HBTs dominate low frequency noise of the devices. Therefore, understanding their temperature dependence is a critical step towards predicting how the noise floor of the devices will change with cooling.

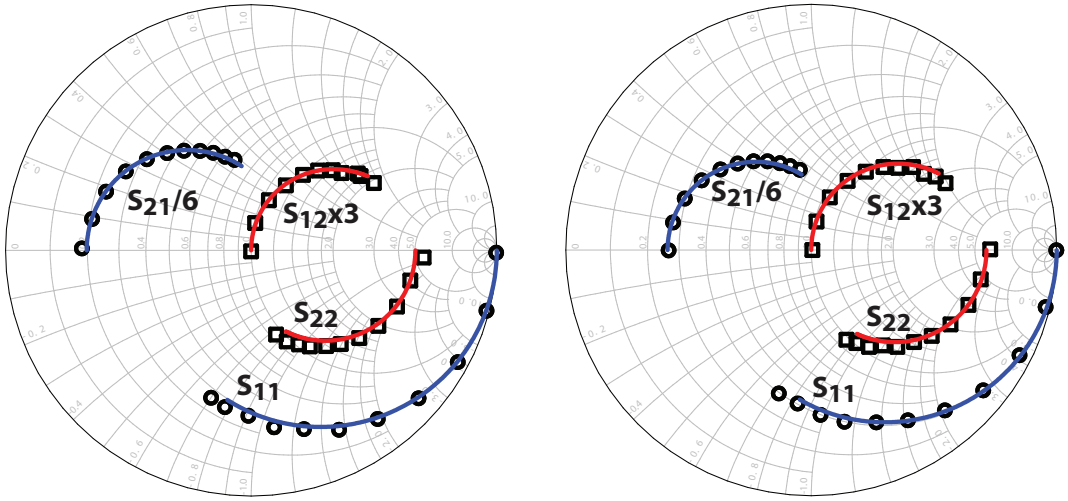
5.3.1.1 HEMT Gate Current

The gate current of HEMTs is due to the Schottky diode between the gate metal and the barrier layer. However, this diode is turned on at a gate voltage higher than what is required to turn on the channel. Therefore, only the diode saturation current flows through the gate unless the device is operated at very high drain current densities. However, the saturation current of a Schottky diode depends significantly on the applied voltage and can be greater than that of a p-n junction [77].

I_G is the sum of currents passing through the gate-source (I_{GS}) and gate-drain (I_{GD}) junctions. To determine I_{GS}, I_G was measured for V_{GD}=0 V. To determine I_{GD}, I_G was measured for V_{GS}=0 V while V_{GD} was kept the same with that of the active measurement. Figure 5.13 shows measurements results of I_{GS} and I_{GD} and it can be seen that both components of I_G decrease with cooling as the thermal excitation drops. This reduction in I_G is expected to improve the low frequency noise performance.

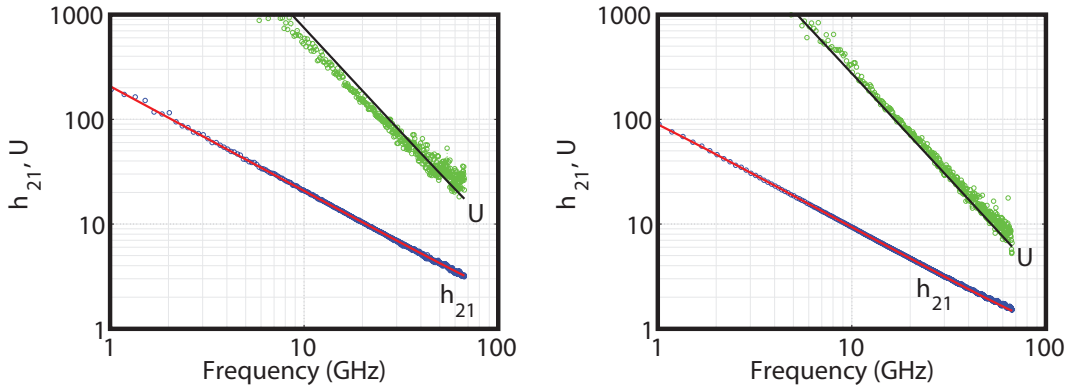


(a)

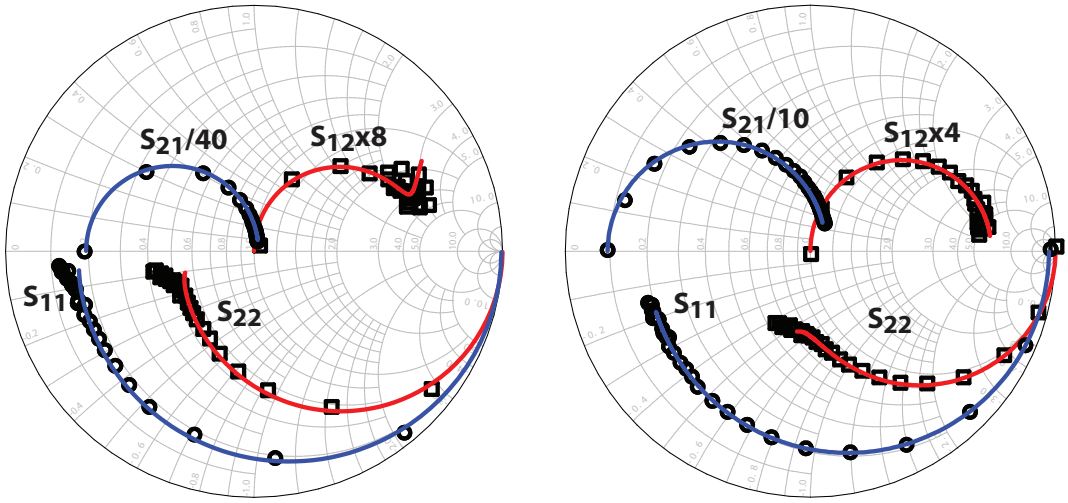


(b)

Figure 5.11: Model to measurement (a) H_{21}/U and (b) s-parameters comparison of the HEMT device. The current density is 125 mA/mm. Solid lines represent the model and symbols represent the measurement. Frequency range is 0.01 - 67 GHz. The left column is 7 K and the right column is 298 K.

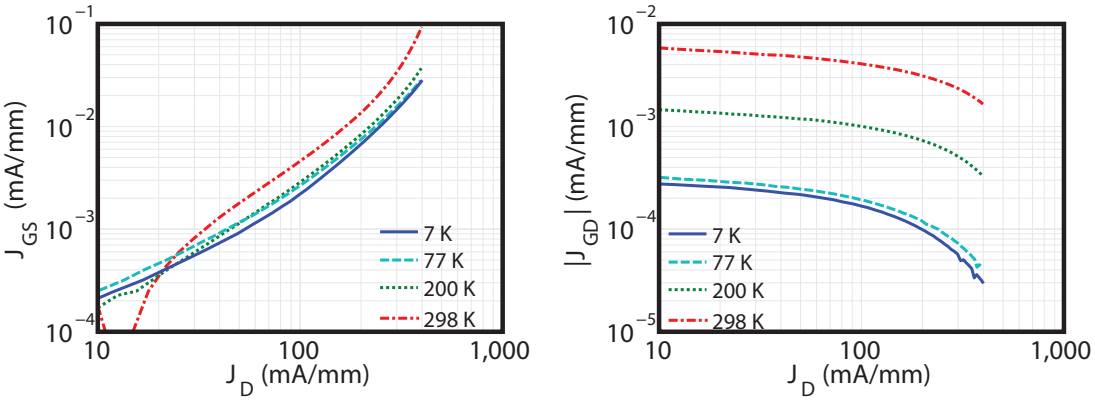


(a)



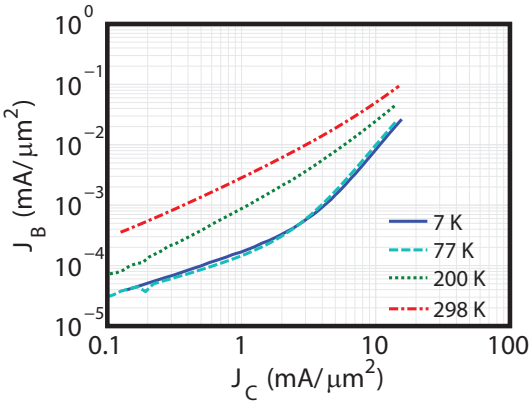
(b)

Figure 5.12: Model to measurement (a) H_{21}/U and (b) s-parameters comparison of the HBT device. The current density is 1.25 mA/μm^2. Solid lines represent the model and symbols represent the measurement. Frequency range is 0.01 - 67 GHz. The left column is 7 K and the right column is 298 K.



(a) (b)

Figure 5.13: Measured (a) I_{GS} and (b) I_{GD} of the HEMT.



(a)

Figure 5.14: Measured I_B of the HBT.

5.3.1.2 HBT Base Current

The base current of SiGe HBT is dominated by the hole diffusion current [16]. Therefore, I_B is expected to reduce significantly with cooling due to the exponential enhancement of the bandgap effect with decreasing temperature. The measurement results are shown in Figure 5.14 and confirm the expectations.

Table 5.1: Extracted HEMT Series Resistances

Temperature K	R_G $\Omega \times mm$	R_S $\Omega \times mm$	R_D $\Omega \times mm$
7	0.164	0.19	0.204
77	0.167	0.194	0.222
200	0.177	0.244	0.284
298	0.194	0.282	0.34

5.3.2 Access Resistances

Input (gate, base) and degeneration (source, emitter) resistances are critical to the noise and f_{max} performance of the devices. Series access resistances arise from the metal contacts and doped semiconductor regions that provide the transition from the device terminals to the intrinsic device. While the resistivity of metals decreases with cooling [146], the temperature dependence of the semiconductor resistivity depends on the doping levels [68, 106, 147].

5.3.2.1 HEMT Gate, Source, and Drain Resistances

The extracted extrinsic resistances of the HEMT are summarized in Table 5.1. In general, the resistances decrease with cooling. Compared to the temperature dependence of R_G that was reported in [103,148], the improvement with cooling is observed to be significantly smaller in this work. The improvement in R_G with cooling saturates in the 7 - 77 K range, which is attributed to boundary scattering being the limitation for carrier mobility rather than electron-phonon scattering [103].

5.3.2.2 HBT, Base, Emitter, and Collector Resistances

The extracted series resistances of the HBT are summarized in Table 5.2. While R_B reduces with increasing bias due to current crowding effects [91], it was observed to be weakly bias dependent in the 2 - 10 mA/μm^2 range. The base resistance improves with cooling at low bias although the temperature dependence is not monotonic and

Table 5.2: Extracted HBT Series Resistances

Temperature K	R_B@J_C=1 mA/μm^2 $\Omega \times \mu$m^2	R_B@J_C=5 mA/μm^2 $\Omega \times \mu$m^2	R_E $\Omega \times \mu$m^2	R_C $\Omega \times \mu$m^2
7	7.9	7.5	1.4	3.1
77	7.5	7.5	1.7	3.4
200	7.2	7.3	2.1	4
298	8.6	7.5	2.7	4.9

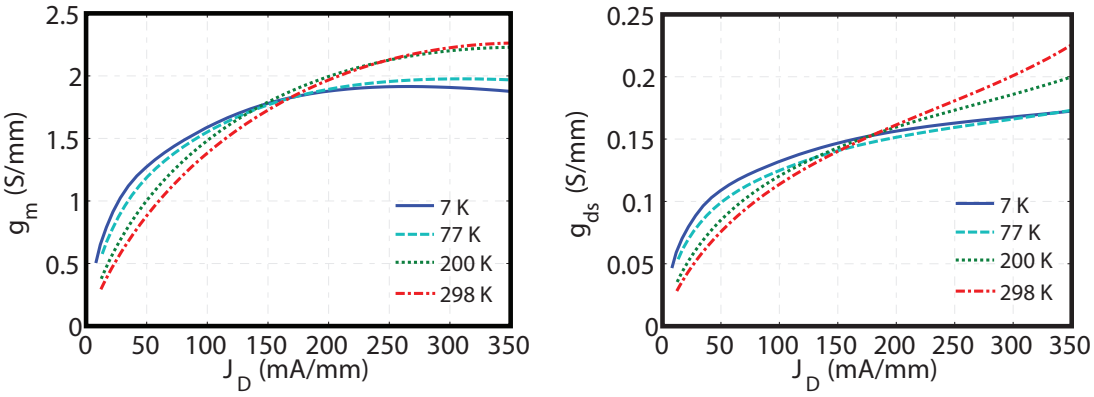
it weakens as the bias current is increased. While the temperature dependence of R_B varies significantly between foundries [26,149], the trend observed for the IBM device used in this work is consistent with previously reported results of the same device [26]. The emitter and collector resistances both improve significantly with cooling which indicates a high dopant concentration in the corresponding regions.

5.3.3 Transconductance and Input/Output Conductances

The transconductance of the devices are expected to improve with cooling, which is beneficial to their small-signal and noise performance. At low frequencies, where the capacitances have a relatively higher impedance, the output conductance limits the voltage gain of HEMT, whereas the input conductance of the HBT relates to the current gain. Therefore, it is important to study temperature dependence of g_{ds} and g_{be}.

5.3.3.1 HEMT Transconductance and Output Conductance

The extracted values of g_m and g_{ds} for the HEMT are shown in Figure 5.15 as a function of bias and temperature. The transconductance improves with cooling for current densities below 150 mA/mm. This improvement is a two step process. First, the mobility improves with cooling, which results in an increased velocity of electrons in the channel [64]. Second, the mean carrier distribution in the channel shifts towards the gate with cooling due to the reduction of thermal diffusion in the channel [103].



(a) (b)

Figure 5.15: Extracted (a) g_{m} and (b) g_{ds} of the HEMT.

This leads to a stronger charge control in the channel, and is often referred as improved quality of pinch-off. Beyond 150 mA/mm, the peak $g_{\mathrm{m,HEMT}}$ degrades with cooling as the onset of a parallel conduction path in the barrier layer shifts to lower current densities and the critical field corresponding the velocity saturation (E_{crit}) decreases [54, 105].

The temperature and bias dependency of g_{ds} is observed to be similar to that of $g_{\mathrm{m,HEMT}}$. For current densities below 150 mA/mm, g_{ds} increases with cooling. This can be explained by the increase in the saturation velocity with cooling [68], which results in I_{DS} being more sensitive to variations in V_{DS}, as demonstrated in Chapter 2.

5.3.3.2 HBT Transconductance and Input Conductance

The extracted values of g_{m} and g_{be} for the HBT are shown in Figure 5.16 as a function of bias and temperature. The transconductance improves with cooling at any current density, as the thermal voltage (V_{TH}) drops. However, the improvement is limited by the degradation in the collector current ideality factor (n_{c}), which, as discussed in Chapter 3, is attributed to the non-equilibrium effects [94]. The input conductance decreases with cooling, as the AC current gain (β_{AC}) improves more than $g_{\mathrm{m,HBT}}$.

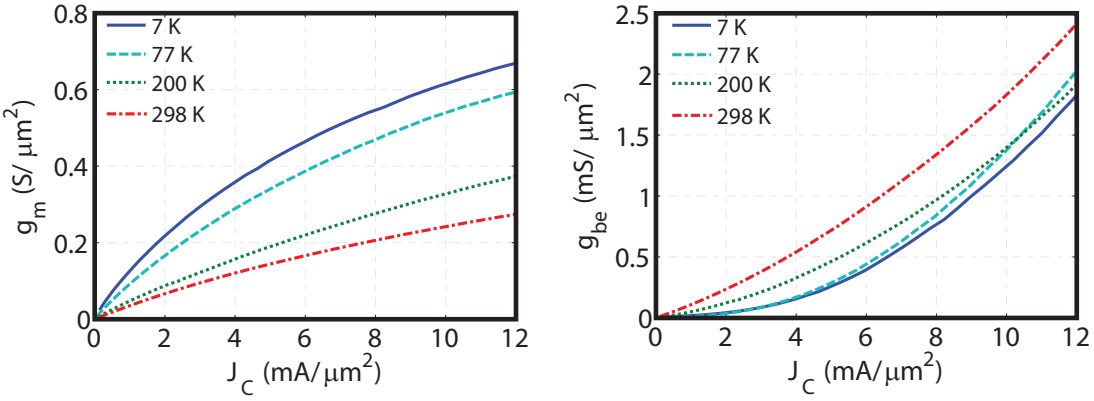


Figure 5.16: Extracted (a) g_{m} and (b) g_{be} of the HBT.

5.3.4 Capacitances

For both devices investigated in this work, dominant capacitances in the small-signal model are expected to have a significant bias and temperature dependence. As the capacitances ultimately determine frequency dependence of the device behavior, understanding their properties is a crucial step towards characterizing the limits of each device.

5.3.4.1 HEMT Gate-Source, Gate-Drain, and Drain-Source Capacitances

At nominal current densities, the gate-source capacitance of a HEMT can be described as a parallel plate capacitor between the gate charge and the channel electron concentration. As the gate voltage is increased, more electrons populate the channel and the mean carrier distribution shifts towards the gate [53].

The extracted values of C_{GS} are shown in Figure 5.17(a). It can be seen that C_{GS} has a similar profile to that of $g_{\mathrm{m,HEMT}}$. At low current densities, C_{GS} increases with cooling, as the channel electron concentration is positioned closer to the gate [103]. As V_{GS} is further increased, the parallel conduction path starts to turn on, which reduces C_{GS} and degrades the charge control [54]. As the formation of the parasitic channel starts at lower current densities when the device is cooled down, C_{GS} and

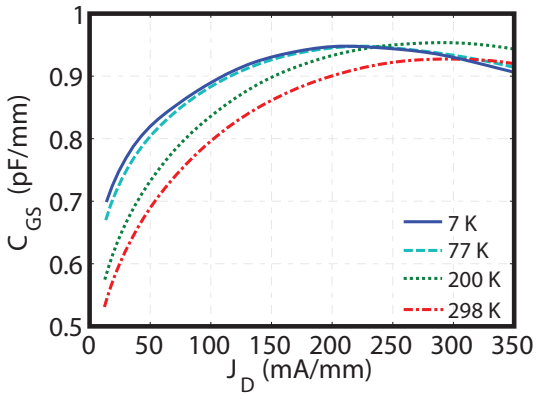
$g_{m,HEMT}$ roll-off earlier at lower temperatures before reaching the peak values observed at higher temperatures.

The gate-drain capacitance of a HEMT is a Schottky depletion capacitance that is formed with reverse biasing of the gate-drain junction. As the device is cooled down, C_{GD} is expected to decrease as the junction built-in potential (ϕ_{bi}) increases. However for a constant current density and V_{DS}, V_{GD} decreases with cooling as the channel turn-on requires a higher gate potential. This is expected to limit the C_{GD} improvement with cooling to some extent. The extracted C_{GD} is shown in Figure 5.17(b) and it can be seen that C_{GD} improves with cooling.

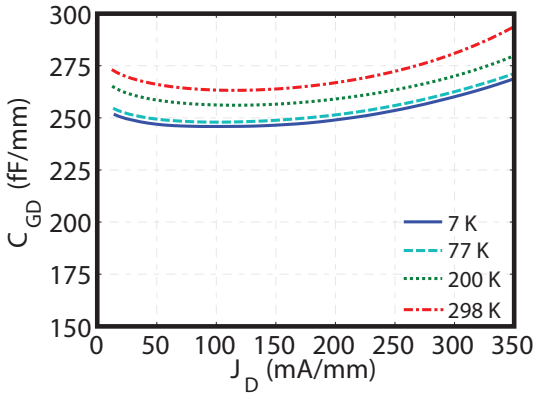
Finally, the extracted C_{DS} of the HEMT is shown in Figure 5.17(c). C_{DS} describes the capacitance between the channel and drain terminal across the depleted saturation region [53]. C_{DS} reduces with cooling down to 77 K and reduces with bias at all temperatures up to 200 mA/mm. It appears that C_{DS} is inversely proportional to the channel conductance (g_{ds}), hence it reduces with cooling as g_{ds} increases.

5.3.4.2 HBT Base-Emitter, Base-Collector, and Collector-Substrate Capacitances

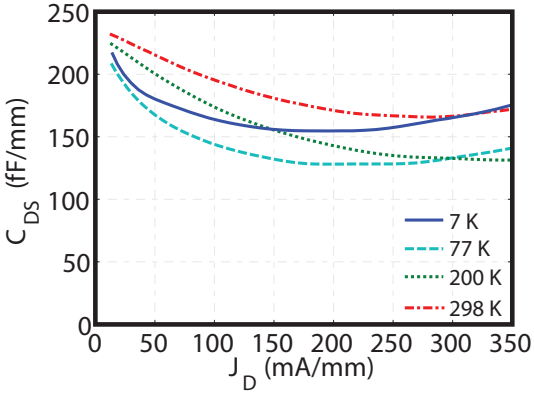
As mentioned earlier, C_{BE} is the parallel combination of depletion and diffusion capacitances in the base-emitter junction. While the depletion component is expected to be weakly temperature dependent, the diffusion capacitance is expected to increase significantly with cooling as $C_{BE,diff} \approx g_{m,HBT}\tau_f$ where τ_f is the total transit time in base and emitter regions [26]. Figure 5.18(a) shows the C_{BE} increase with cooling. For J_C=2 mA/μm^2, the C_{BE} increases by a factor of 1.7 with cooling from 298 K to 7 K, whereas the increase in $g_{m,HBT}$ is a factor of 3.3 under the same conditions. This difference clearly demonstrates that only the diffusion component of C_{BE} increases proportional to $g_{m,HBT}$ with cooling.



(a)

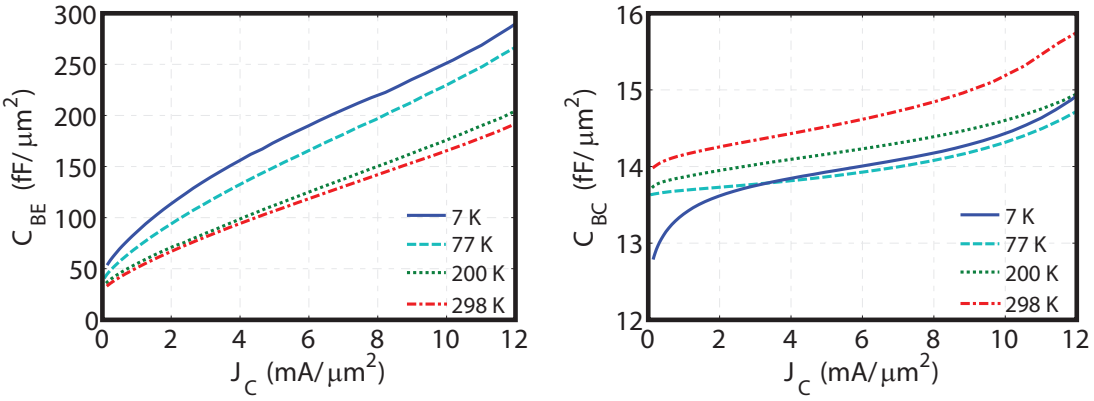


(b)



(c)

Figure 5.17: Extracted (a) C_{GS}, (b) C_{GD} and (c) C_{DS} of the HEMT.



(a) (b)

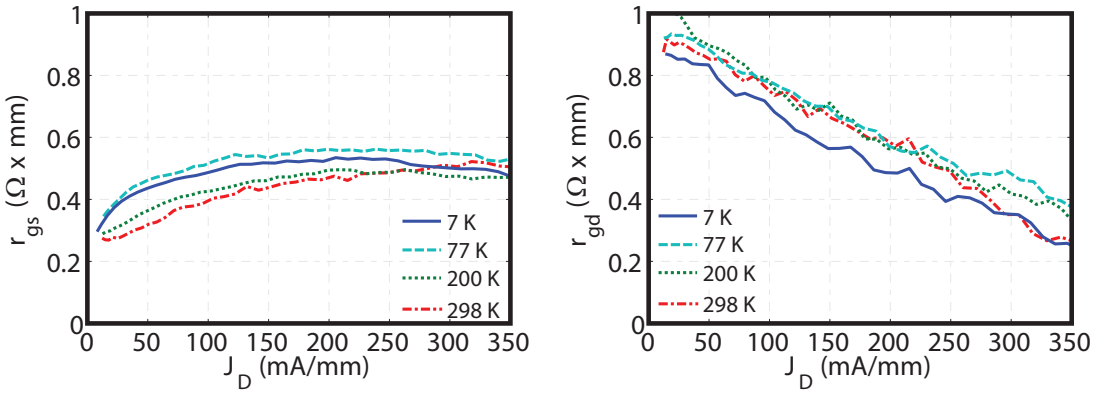
Figure 5.18: Extracted (a) C_{BE} and (b) C_{BC} of the HBT.

The base-collector capacitance of an HBT is the depletion capacitance arising from the reverse-biased base-collector junction. Similar to C_{GD} of the HEMT, C_{BC} is expected to decrease with cooling, as the junction built-in potential goes up. Extracted C_{BC} values appear in Figure 5.18(b), and verify the expectations. It should be noted that, for a fixed V_{CE}, the decrease in C_{BC} with cooling is limited by the reducing V_{BC}, as the device requires higher V_{BE} at lower temperatures for a given current density.

Finally, the total capacitance of the substrate network is found to be weakly temperature dependent in the 77 - 298 K range with an R_{SUB} decreasing with cooling. As the device is further cooled down to 7 K, a sharp decline in the C_{CS} is observed as the substrate freezes out due to being doped below the Mott-transition. Regarding the absolute values of the substrate network, C_{CS}, C_{SUB} and R_{RUB} was found to be 5.4 fF/μm^2, 3.2 fF/μm^2 and 620 $\Omega \times \mu$m^2 at 77 K, respectively. The impact of the substrate network on the device performance was observed to be negligible at the frequencies of interest.

5.3.5 NQS Resistances

NQS resistances are inherently difficult to extract, as they are in series with a capacitor in the small-signal model. Unfortunately, this does not make them any less



(a) (b)

Figure 5.19: Extracted (a) R_{GS} and (b) R_{GD} of the HEMT.

important regarding device performance. While it is hard to predict their temperature and bias dependence, the experimental results can still be linked to the overall performance.

5.3.5.1 HEMT Gate-Source and Gate-Drain Resistance

The HEMT model consists of two NQS resistances, r_{gs} and r_{gd}. The extraction results are shown in Figure 5.19. The gate-source resistance increases slightly with cooling and has a bell shaped bias dependence, in a similar trend to that of C_{GS}. Furthermore, the peak value of r_{gs} is reached at lower bias points with cooling, which was also observed for C_{GS}. The gate-source resistance has a diminishing impact on the noise performance with cooling as its noise contribution is inversely proportional to temperature.

The gate-drain resistance appears to be temperature independent and inversely proportional to the bias current. At low bias, it is significantly higher than r_{gs}. Fortunately, the overall noise performance is not as sensitive to r_{gd} as it is to r_{gs} since r_{gd} is in series with a relatively high impedance feedback capacitance.

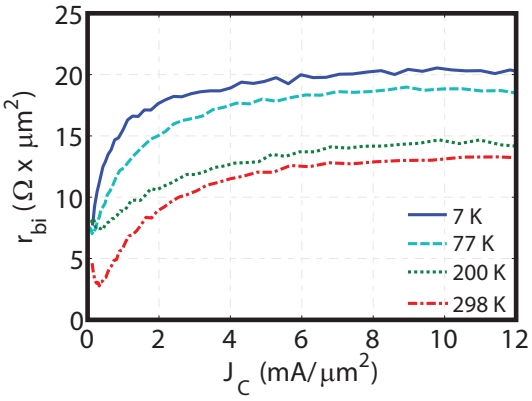


Figure 5.20: Extracted r_{bi} of the HBT.

5.3.5.2 HBT Non-Quasi Static Base Resistance

The non-quasi static base resistance represents the delay between the base charge and V_{BE} [150]. The extraction results are shown in Figure 5.20 and it can be seen that r_{bi} increases significantly with bias and cooling. Fortunately, the impact of r_{bi} on the noise performance is minimal as it is non-dissipative and it represents a delay phenomena. Thus, the non-quasi static base resistance does not generate thermal noise. Regardless, the extracted r_{bi} values are too high to be ignored for the small-signal modeling. The increase in r_{bi} with cooling demands further investigation of the physical origins, which is beyond the scope of this work.

5.4 Summary

The small-signal noise model extraction techniques and modeling results have been presented in this chapter. The extraction of individual model parameters were described step by step. Model to measurement agreement was demonstrated. An overall improvement in the transport properties with cooling enable better high frequency and noise performance. In the next chapter, the models presented so far will be expanded to capture the nonlinear behavior of the devices.

CHAPTER 6

NONLINEAR MODELING

In this chapter, the extracted nonlinear models, which are used to characterize intermodulation performance of the devices, are presented. First, different approaches that are commonly used for large-signal modeling are discussed. Second, the extraction procedure of the modeling approach used in this work is presented. Simulation results of the extracted models are compared to on-wafer IP3 measurements to verify the models. Finally, the temperature and bias dependence of the model parameters are explained in terms of the device physics that were presented in Chapters 2 and 3.

6.1 Comparison of Different Approaches

Empirical and physical compact models can predict the I-V and C-V characteristics of transistors in a continuous manner through the use of closed-form expressions. Empirical models are not limited by the physical origins of the device operation; they are defined with functions that provide the best fit. Physical models are based on semiconductor principles, although fitting factors can be present.

To the best of author's knowledge, there is no commonly used physical model for a HEMT. The difficulty of developing a physical model for a HEMT can be attributed to several different mechanisms that were discussed in Chapter 2. First, the transcendental relationship between the Fermi level and sheet carrier concentration is challenging to solve [59, 102]. Donor neutralization at the onset of the parallel conduction path formation in the barrier layer can also be difficult to predict [54].

Finally, modeling the velocity - field relationship in the channel in the presence of inter-valley transfer can be almost impossible without Monte Carlo simulations [66, 67].

In the absence of physical models, empirical models have become successful in the modeling of HEMTs. Among them, the Angelov model is probably the most popular [151]. It relies on the hyperbolic tangent function to define the relationship between the drain current and the gate voltage, which can predict the smooth roll-off in the drain current outside the linear charge-control regime. This model is particularly known for its success in predicting the derivatives of the transconductance, which is crucial for accurate modeling of the intermodulation distortion [152]. This model was also demonstrated to be suitable for use at croygenic temperatures [153].

For an HBT, physical models are readily available. Some of these models are VBIC [154], HICUM [155] and MEXTRAM [156]. These models rely on the classical drift-diffusion theory, which does not account for non-equilibrium base transport. As the non-equilibrium base transport describes the device operation at cryogenic temperatures, these models are only accurate in a relatively narrow temperature range (-55 – -125 C) [157]. Recent efforts based on the modification of the MEXTRAM model have enabled accurate prediction of the SiGe HBT characteristics at 93 K [158].

Empirical and physical compact models enable the continuous description of I-V and C-V relationships over a wide range of biases. However, this is not necessarily the primary goal within the context of this work. What is essential for intermodulation characterization is the accurate knowledge of local derivatives of the I-V and C-V relationships. Therefore, the available small-signal models can be extended to include this information, which will be called a weakly nonlinear model.

The weakly nonlinear model requires Taylor series expansion coefficients of the individual nonlinearities. Since only the weak nonlinearities are of interest, the bias point is not expected to be modulated by the RF drive and the expansion coefficients

can be defined independently at any bias point. This is important, as it eliminates the need for analytical functions to describe the I-V and C-V characteristics; the Taylor series expansion coefficients can be obtained and defined numerically. Thus, as long as the sources of nonlinearity are known, the intermodulation performance of a device can be predicted with ease at a given temperature or bias point wtih this approach. For this reason, the weakly nonlinear models have often been used to study the intermodulation performance of FETs [159–161] and HBTs [162–164]. This approach has also been systematically applied in this work.

6.2 Extraction Procedure

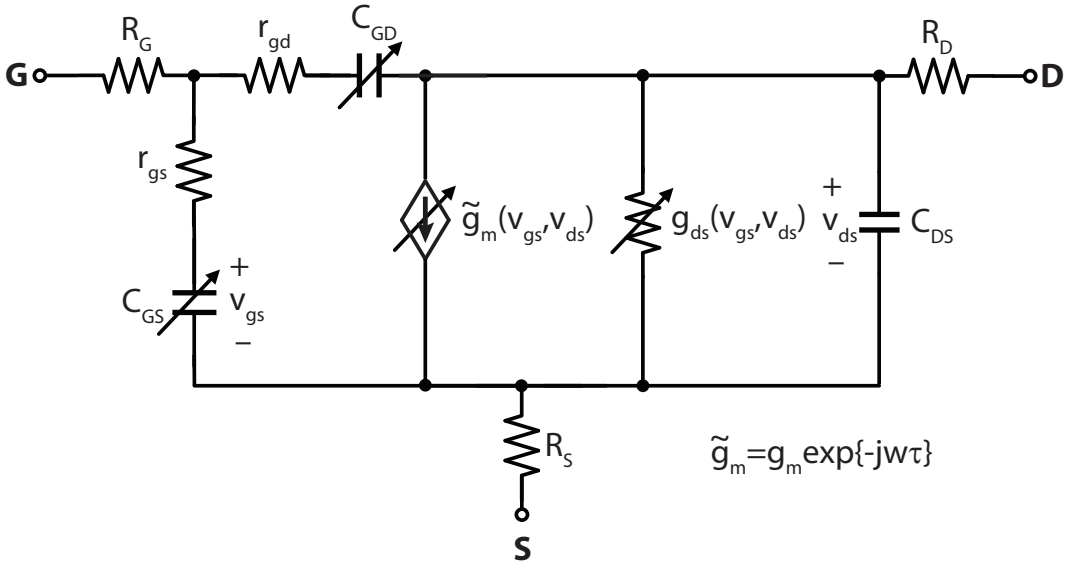
Figure 6.1 shows the weakly nonlinear models that were acquired by expanding the small-signal models. The Taylor series coefficients were defined up to third order for the strong nonlinearities which were discussed in Chapter 2. Similar to a small-signal model, the weakly nonlinear model is defined at discrete bias points.

A Taylor series expansion of HEMT drain-source current up to third order can be written as [165]

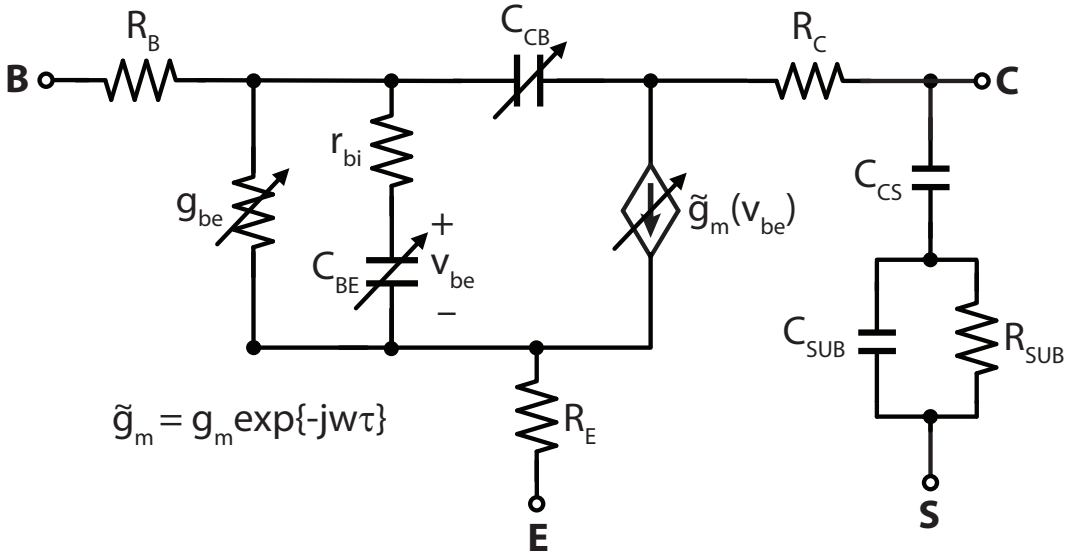
$$\begin{aligned} I_{\mathrm{ds}}\left(V_{\mathrm{gs}}, V_{\mathrm{ds}}\right) = I_{\mathrm{DS}} &+ \frac{\delta I_{\mathrm{ds}}}{\delta V_{\mathrm{gs}}} v_{\mathrm{gs}} + \frac{1}{2}\frac{\delta^2 I_{\mathrm{ds}}}{\delta V_{\mathrm{gs}}^2} v_{\mathrm{gs}}^2 + \frac{1}{6}\frac{\delta^3 I_{\mathrm{ds}}}{\delta V_{\mathrm{gs}}^2 3} v_{\mathrm{gs}}^3 + \frac{\delta I_{\mathrm{ds}}}{\delta V_{\mathrm{ds}}} v_{\mathrm{ds}} + \frac{1}{2}\frac{\delta^2 I_{\mathrm{ds}}}{\delta V_{\mathrm{ds}}^2} v_{\mathrm{ds}}^2 \\ &+ \frac{1}{6}\frac{\delta^3 I_{\mathrm{ds}}}{\delta V_{\mathrm{ds}}^3} v_{\mathrm{ds}}^3 + \frac{\delta^2 I_{\mathrm{ds}}}{\delta V_{\mathrm{gs}} \delta V_{\mathrm{ds}}} v_{\mathrm{gs}} \cdot v_{\mathrm{ds}} + \frac{1}{2}\frac{\delta^3 I_{\mathrm{ds}}}{\delta V_{\mathrm{gs}}^2 \delta V_{\mathrm{ds}}} v_{\mathrm{gs}}^2 \cdot v_{\mathrm{ds}} + \frac{1}{2}\frac{\delta^3 I_{\mathrm{ds}}}{\delta V_{\mathrm{gs}} \delta V_{\mathrm{ds}}^2} v_{\mathrm{gs}} \cdot v_{\mathrm{ds}}^2. \end{aligned} \quad (6.1)$$

Once the operating point is subtracted, the remaining small-signal expression can be written as

$$\begin{aligned} i_{\mathrm{ds}}\left(v_{\mathrm{gs}}, v_{\mathrm{ds}}\right) = g_{\mathrm{m1}} v_{\mathrm{gs}} + g_{\mathrm{m2}} v_{\mathrm{gs}}^2 + g_{\mathrm{m3}} v_{\mathrm{gs}}^3 + g_{\mathrm{ds1}} v_{\mathrm{ds}} + g_{\mathrm{ds2}} v_{\mathrm{ds}}^2 + g_{\mathrm{ds3}} v_{\mathrm{ds}}^3 \\ + g_{\mathrm{mds}} v_{\mathrm{gs}} v_{\mathrm{ds}} + g_{\mathrm{m2ds}} v_{\mathrm{gs}}^2 v_{\mathrm{ds}} + g_{\mathrm{mds2}} v_{\mathrm{gs}} v_{\mathrm{ds}}^2, \end{aligned} \quad (6.2)$$

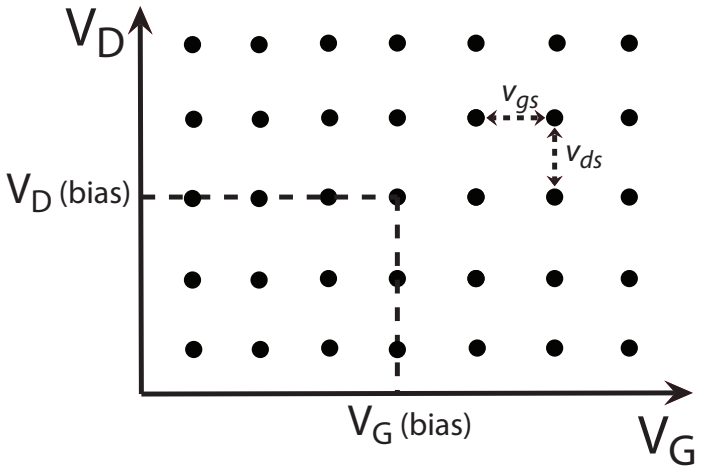


(a)



(b)

Figure 6.1: Weakly nonlinear (a) HEMT and (b) HBT models. The elements with cross-arrows indicate that the Taylor series expansion is defined.



(a)

Figure 6.2: Example DC bias grid for coefficient extraction of the nonlinear model.

where g_{m}, g_{ds} and g_{mds} are the coefficients that represent the nonlinearities associated with the transconductance, output conductance, and cross-coupling, respectively.

To extract the coefficients, s-parameter measurements were performed on a DC bias grid [166]. The dynamic grid is shown in Figure 6.2 and covers seven points on the V_{GS} domain and five points on the V_{DS} domain. Depending on the ambient temperature, the grid spacing is in the range of 8-10 mV on the V_{GS} axis and 50 mV on the V_{DS} axis. The $g_{\mathrm{m,HEMT}}$ and g_{ds} values obtained at each grid point are used to solve the unknown higher order coefficients.

Since $g_{\mathrm{m,HEMT}}$ and g_{ds} are extracted from RF measurements, Equation (6.2) can be written such that it describes the nonlinearity of $g_{\mathrm{m,HEMT}}$ and g_{ds} rather than that of i_{ds}. Differentiating i_{ds} with respect to v_{gs} and v_{ds} yields

$$g_{\mathrm{m}}\left(v_{\mathrm{gs}}, v_{\mathrm{ds}}\right)=\frac{\delta i_{\mathrm{ds}}}{\delta v_{\mathrm{gs}}}=g_{\mathrm{m1}}+2 g_{\mathrm{m2}} v_{\mathrm{gs}}+3 g_{\mathrm{m3}} v_{\mathrm{gs}}^{2}+g_{\mathrm{mds}} v_{\mathrm{ds}}+2 g_{\mathrm{m2ds}} v_{\mathrm{gs}} v_{\mathrm{ds}}+g_{\mathrm{mds2}} v_{\mathrm{ds}}^{2} \tag{6.3}$$

$$g_{\mathrm{ds}}\left(v_{\mathrm{gs}}, v_{\mathrm{ds}}\right)=\frac{\delta i_{\mathrm{ds}}}{\delta v_{\mathrm{ds}}}=g_{\mathrm{ds1}}+2 g_{\mathrm{ds2}} v_{\mathrm{ds}}+3 g_{\mathrm{ds3}} v_{\mathrm{ds}}^{2}+g_{\mathrm{mds}} v_{\mathrm{gs}}+2 g_{\mathrm{mds2}} v_{\mathrm{gs}} v_{\mathrm{ds}}+g_{\mathrm{m2ds}} v_{\mathrm{ds}}^{2}. \tag{6.4}$$

Once Equations (6.3) and (6.4) are applied to all points of the grid, they can be written in the matrix form as follows [166]

$$\mathbf{Y}=\mathbf{B} \cdot \mathbf{C}, \tag{6.5}$$

where $\mathbf{Y}$ consists of the known $g_{\mathrm{m,HEMT}}$ and g_{ds} values at all grid points and is written as

$$\mathbf{Y}=\left[g_{\mathrm{m}}(1)\ g_{\mathrm{ds}}(1)\ g_{\mathrm{m}}(2)\ g_{\mathrm{ds}}(2)\ \ldots\ g_{\mathrm{m}}(35)\ g_{\mathrm{ds}}(35)\right]^{T}, \tag{6.6}$$

$\mathbf{B}$ represents the known differential DC voltages of the grid points, where $v_{\mathrm{gsi}}=V_{\mathrm{GS,i}}-V_{\mathrm{GS,center}}$ and $v_{\mathrm{dsi}}=V_{\mathrm{DS,i}}-V_{\mathrm{DS,center}}$, and $\mathbf{C}$ is the unknown coefficients matrix given as

$$\mathbf{C}=\left[g_{\mathrm{m1}}\ g_{\mathrm{m2}}\ g_{\mathrm{m3}}\ g_{\mathrm{ds1}}\ g_{\mathrm{ds2}}\ g_{\mathrm{ds3}}\ g_{\mathrm{mds}}\ g_{\mathrm{m2ds}}\ g_{\mathrm{mds2}}\ \right]^{T}. \tag{6.7}$$

The explicit form of Equation (6.5) is given as

$$\begin{bmatrix} g_{\mathrm{m}}(1) \\ g_{\mathrm{ds}}(2) \\ \vdots \\ g_{\mathrm{m}}(35) \\ g_{\mathrm{ds}}(35) \end{bmatrix} = \begin{bmatrix} 1 & 2v_{\mathrm{gs1}} & 3v_{\mathrm{gs1}}^{2} & 0 & 0 & 0 & v_{\mathrm{ds1}} & 2v_{\mathrm{gs1}}v_{\mathrm{ds1}} & v_{\mathrm{ds1}}^{2} \\ 0 & 0 & 0 & 1 & 2v_{\mathrm{ds1}} & 3v_{\mathrm{ds1}}^{2} & v_{\mathrm{gs1}} & 2v_{\mathrm{ds1}}^{2} & v_{\mathrm{gs1}}v_{\mathrm{ds1}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 2v_{\mathrm{gs35}} & 3v_{\mathrm{gs35}}^{2} & 0 & 0 & 0 & v_{\mathrm{ds35}} & 2v_{\mathrm{gs35}}v_{\mathrm{ds35}} & v_{\mathrm{ds35}}^{2} \\ 0 & 0 & 0 & 1 & 2v_{\mathrm{ds35}} & 3v_{\mathrm{ds35}}^{2} & v_{\mathrm{gs35}} & 2v_{\mathrm{ds35}}^{2} & v_{\mathrm{gs35}}v_{\mathrm{ds35}} \end{bmatrix} \begin{bmatrix} g_{\mathrm{m1}} \\ g_{\mathrm{m2}} \\ \vdots \\ g_{\mathrm{m2ds}} \\ g_{\mathrm{mds2}} \end{bmatrix}. \tag{6.8}$$

To solve Equation (6.5) which is over-determined, the matrix form of ordinary least squares method is used and the unknown coefficient matrix C is obtained as [167]

$$\mathbf{C} = \left(\left(\mathbf{B}^T \cdot \mathbf{B} \right)^{-1} \cdot \mathbf{B}^T \right) \cdot \mathbf{Y}. \tag{6.9}$$

This approach is also applied to the extraction of capacitance nonlinearities as the entire small-signal model at each point of the grid can be determined from the s-parameter measurements. The Taylor series expansion of the AC charge associated with the gate-source capacitance can be expressed as

$$q_{\mathrm{gs}}\left(v_{\mathrm{gs}}\right) = c_{\mathrm{gs1}} v_{\mathrm{gs}} + c_{\mathrm{gs2}} v_{\mathrm{gs}}^2 + c_{\mathrm{gs3}} v_{\mathrm{gs}}^3. \tag{6.10}$$

Similarly, the Taylor series expansion of the AC charge associated with the gate-drain capacitance can be expressed as

$$q_{\mathrm{gd}}\left(v_{\mathrm{gd}}\right) = c_{\mathrm{gd1}} v_{\mathrm{gd}} + c_{\mathrm{gd2}} v_{\mathrm{gd}}^2 + c_{\mathrm{gd3}} v_{\mathrm{gd}}^3. \tag{6.11}$$

The grid based non-linear model extraction technique is also applied to the HBT. For completeness, the Taylor series expansion of the AC base current is provided as

$$i_{\mathrm{be}}\left(v_{\mathrm{be}}\right) = g_{\mathrm{be1}} v_{\mathrm{be}} + g_{\mathrm{be2}} v_{\mathrm{be}}^2 + g_{\mathrm{be3}} v_{\mathrm{be}}^3. \tag{6.12}$$

For the HBT extraction, the key differences is the use of more grid points (11×5) and a smaller grid spacing in the V_{BE} domain (1 mV) at cryogenic temperatures in order to achieve a reasonable resolution in the J_{C} domain. A detailed explanation of the nonlinear model measurement and extraction flow is provided in Appendix A.

6.3 Model to Measurement Agreement

The accuracy of the models was verified following the extraction. The OIP3 of the models were simulated and compared to on-wafer measurement of the device test structures when terminated in 50 Ω. The results at 7 K and 298 K are shown in Figures 6.3 and 6.4 for the HEMT and HBT, respectively. Similar results are achieved at intermediate temperatures. Good agreement is observed for both devices over a wide range of bias points and frequencies.

6.4 Modeling Results

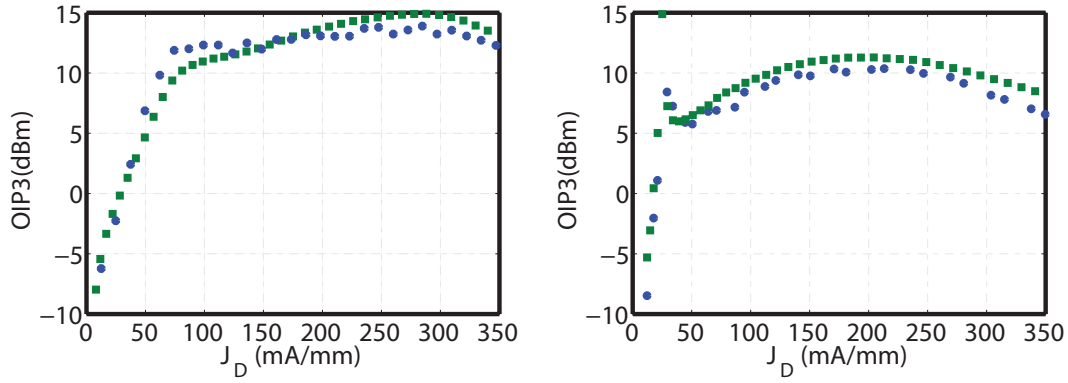
The temperature and bias dependence of the extracted non-linear model parameters are discussed in this section. The results are compared to the theoretical expectations that were provided in Chapters 2 and 3.

6.4.1 Transconductance

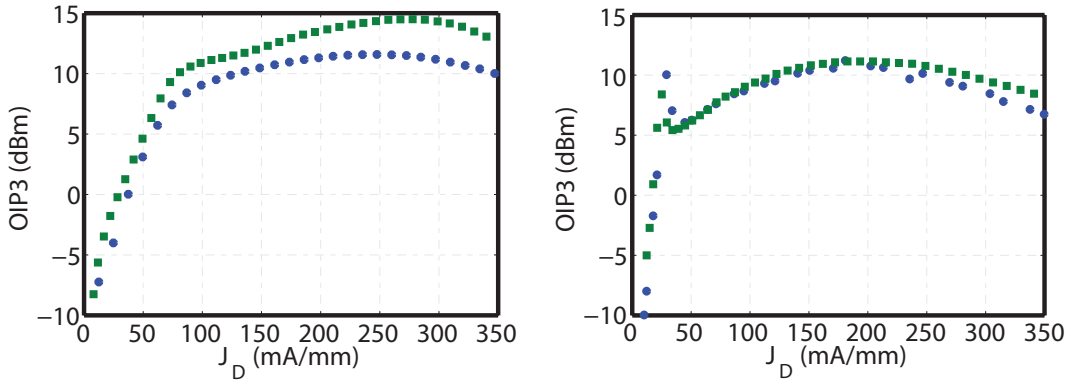
The transconductance nonlinearity is expected to be the key contributor to intermodulation. Furthermore, as the charge-control capacitance and input conductance nonlinearity are expected to have a similar trend to that of the transconductance, it is appropriate to discuss the transconductance nonlinearity first.

6.4.1.1 HEMT Transconductance

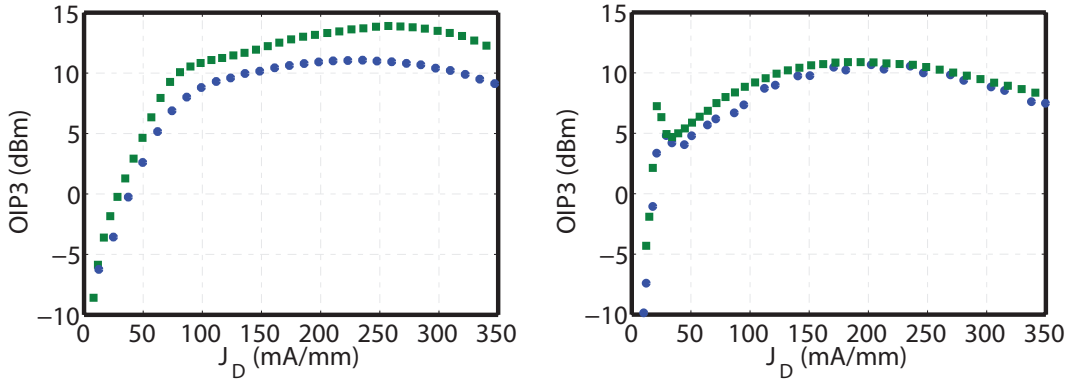
The second and third order coefficients that represent the nonlinearity of $g_{\mathrm{m,HEMT}}$ are shown in Figure 6.5. Up to 50 mA/mm, the device exhibits BJT like behavior at 200 K and 298 K, as $g_{\mathrm{m2,HEMT}}$ increases with bias and $g_{\mathrm{m3,HEMT}}$ is positive. As described in Chapter 2, this behavior is expected in the sub-threshold regime, where the channel carrier concentration is low. As the device is cooled down to 77 K or 7 K, the HBT-like characteristic at low bias disappears, which can be attributed to the improved quality of pinch-off.



(a)

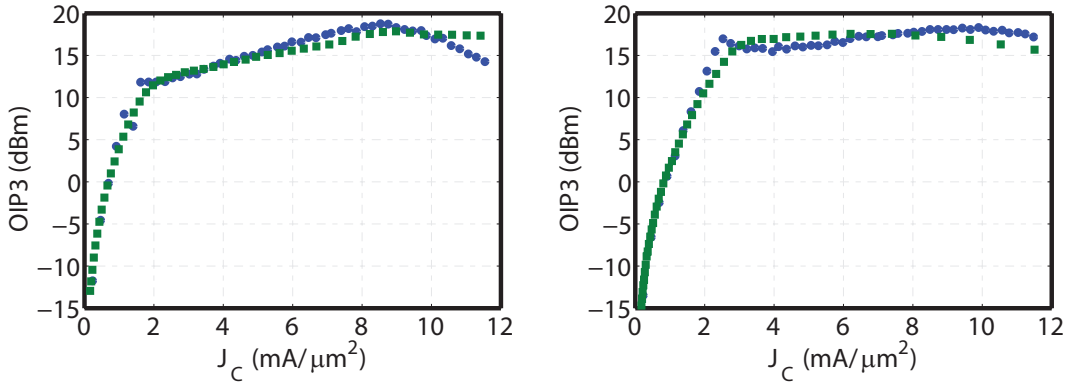


(b)

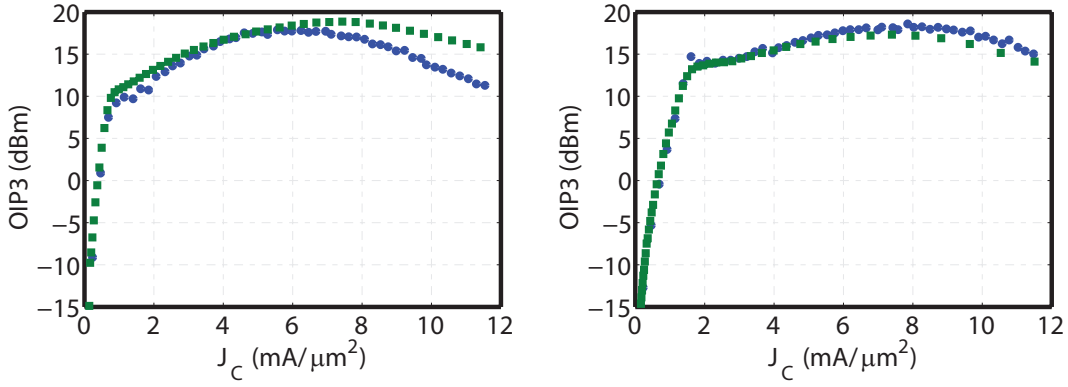


(c)

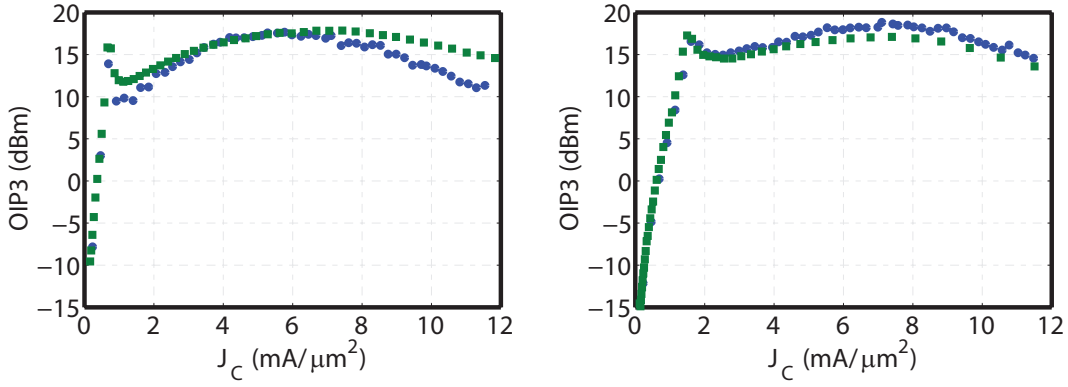
Figure 6.3: Model to measurement OIP3 comparison of the $40\times0.1\mu\text{m}^2$ HEMT at (a) 1 GHz, (b) 10 GHz and (c) 20 GHz. Square data points represent the model and circle data points represent the measurement. The left column is 7 K and the right column is 298 K.



(a)

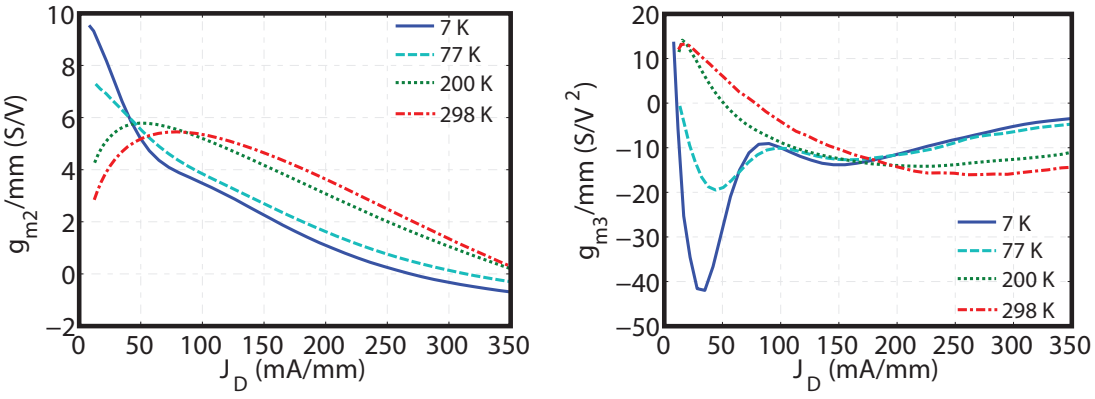


(b)



(c)

Figure 6.4: Model to measurement OIP3 comparison of the $18\times0.12\mu\text{m}^2$ HBT at (a) 1 GHz, (b) 10 GHz and (c) 20 GHz. Square data points represent the model and circle data points represent the measurement. The left column is 7 K and the right column is 298 K.

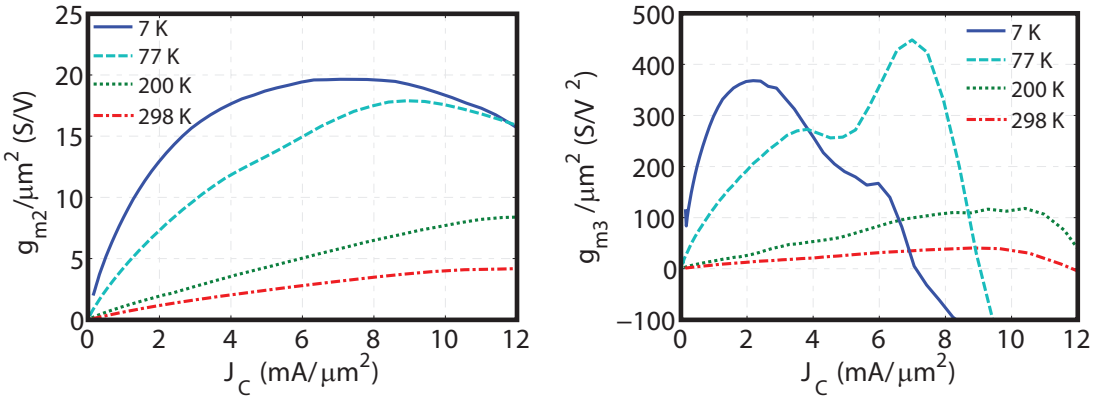


(a) (b)

Figure 6.5: (a) Second and (b) third order Taylor series expansion coefficients of the HEMT transconductance.

The overall decrease in $g_{\mathrm{m2,HEMT}}$ with increasing bias is predicted by the theoretical discussion that was presented in Chapter 2. This decrease is due to the fact that second derivative of $g_{\mathrm{m,HEMT}}$ is inversely proportional to the overdrive voltage (V_{GS}-V_{TN}). Beyond 75 mA/mm, $g_{\mathrm{m2,HEMT}}$ decreases with cooling which can be attributed to the decrease in the critical electric-field that corresponds to the velocity saturation (E_{crit}).

The bias dependence of $g_{\mathrm{m3,HEMT}}$ is also quite interesting. At all temperatures, the absolute value of $g_{\mathrm{m3,HEMT}}$ reaches a peak at a current density below 50 mA/mm. This absolute value appears to increase with cooling. This temperature trend is valid until 175 mA/mm and is explained by the decrease in E_{crit} with cooling. Beyond 175 mA/mm, the absolute value of $g_{\mathrm{m3,HEMT}}$ decreases with cooling. The transition point, at which the temperature trend is reversed, is predicted by the theoretical transconductance expression of short-channel HEMT, which was presented in Chapter 2.



(a) (b)

Figure 6.6: (a) Second and (b) third order Taylor series expansion coefficients of the HBT transconductance.

6.4.1.2 HBT Transconductance

Second and third order coefficients that are associated with the nonlinearity of $g_{\mathrm{m,HBT}}$ are shown in figure 6.6. A peak is observed for both $g_{\mathrm{m2,HBT}}$ and $g_{\mathrm{m3,HBT}}$ as a function of current density. The roll-off occurs at a lower bias point for $g_{\mathrm{m3,HBT}}$ than it does for $g_{\mathrm{m2,HBT}}$ and it shifts to a lower current density as the device is cooled down.

An ideal HBT is expected to have transconductance coefficients that continuously increase with bias. As a bias independent ideality factor (n_{c}) can not result in $g_{\mathrm{m2,HBT}}$ and $g_{\mathrm{m3,HBT}}$ to be inversely proportional to the current density, the roll-off in $g_{\mathrm{m2,HBT}}$ and $g_{\mathrm{m3,HBT}}$ indicates a significantly nonlinear n_{c}. The experimental ideality factor values reported in [26] show an exponential dependence of n_{c} on the current density which becomes more significant at low temperatures. This trend explains the roll-off in the higher order $g_{\mathrm{m,HBT}}$ coefficients and the shift in their peak with cooling.

Table 6.1: Selected Bias Points for the Constant g_{m} Biasing

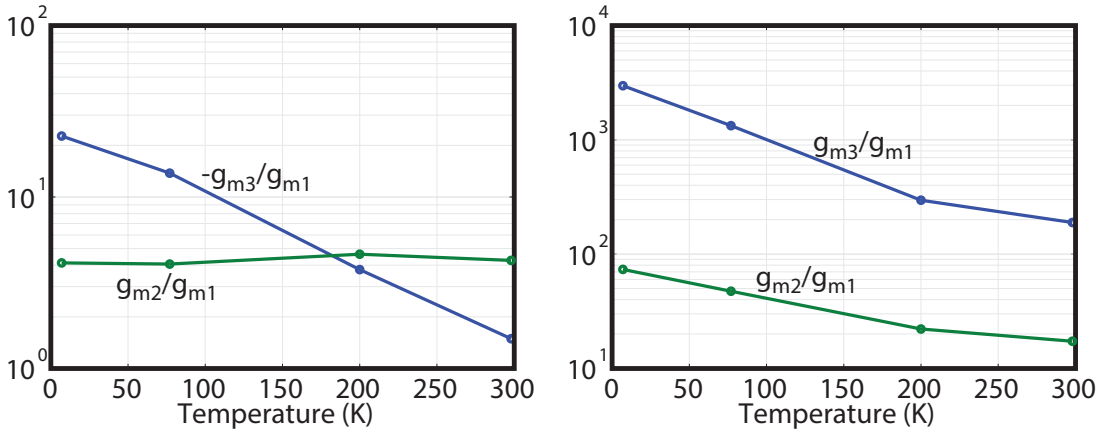
HEMT			HBT	
J_{D} mA/mm	g_{m} S/mm	Temperature K	J_{C} mA/μm^2	g_{m} mS/μm^2
50	1.27	7	0.56	81
57.5	1.26	77	0.88	83
70	1.23	200	1.94	85
87.5	1.27	298	2.55	82

6.4.1.3 Temperature Dependence of the Transconductance Nonlinearity Under Constant g_{m} Bias

So far, the discussion of transconductance nonlinearity covered a wide range of bias points. To obtain a better perspective on its temperature dependence, the discussion can be narrowed down to a single biasing scheme. The bias points can be selected such that a constant transconductance is achieved at all temperatures. This scenario is suitable and often applied for the operation of cryogenic low-noise amplifiers in order to maintain the specified gain at different temperatures.

The bias points that result in a constant transconductance across the temperature range are shown in Table 6.1. At 7 K, the selected current densities are in the bias region where the devices achieve their minimum noise in the 1–20 GHz frequency range. For these bias points, the normalized second ($g_{\mathrm{m2}}/g_{\mathrm{m}}$) and third order ($g_{\mathrm{m3}}/g_{\mathrm{m}}$) Taylor series expansion coefficients are shown in Figure 6.7.

The normalized $g_{\mathrm{m2,HEMT}}$ is fairly independent of temperature, as the bias points correspond to low overdrive levels where the normalized $g_{m2,HEMT}$ is insensitive to E_{crit}. The normalized $g_{m3,HEMT}$ however increases substantially with cooling, as E_{crit} is reduced. For the HBT, both the normalized g_{m2} and g_{m3} increase with cooling as the thermal voltage drops. However, the increase in the normalized $g_{\mathrm{m3,HBT}}$ is only a factor of 15 with cooling from 298 K to 7 K whereas the thermal voltage drops by a factor of 43. This discrepancy is due to the increasing ideality factor and its bias



(a) (b)

Figure 6.7: Normalized second and third order Taylor series expansion coefficients of the (a) HEMT and (b) HBT transconductance.

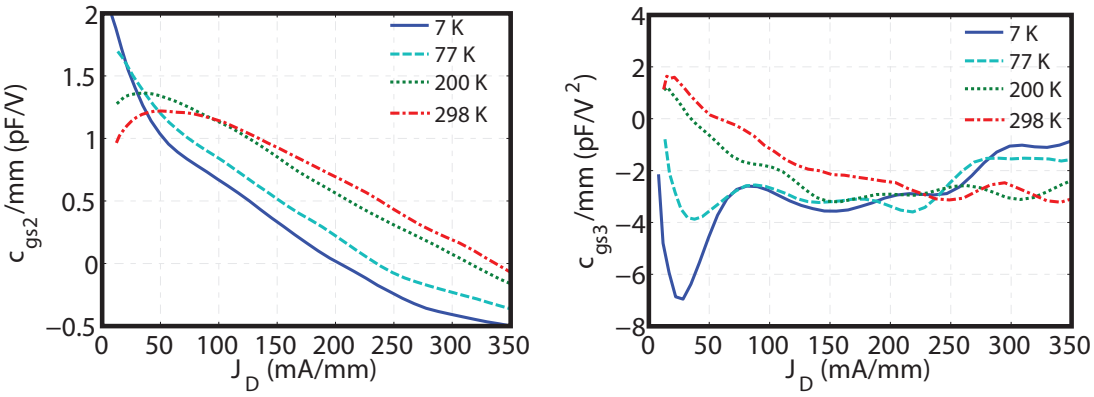
dependence. Nevertheless, at any temperature, the transconductance of the HBT is more nonlinear than that of the HEMT.

6.4.2 Charge-Control Capacitance

With the knowledge of how the transconductance nonlinearity changes with bias and temperature, the nonlinearity of the charge-control capacitance can be discussed, as it is expected to have a similar trend.

6.4.2.1 HEMT Gate-Source Capacitance

Higher order coefficients that represent nonlinearity of the HEMT gate-source capacitance are shown in Figure 6.8. The results are remarkably similar to that of the HEMT transconductance. As the gate-source capacitance relates to the transconductance through the channel transit time, this similarity indicates that the channel transit time is only weakly nonlinear.



(a) (b)

Figure 6.8: (a) Second and (b) third order Taylor series expansion coefficients of the HEMT gate-source capacitance.

6.4.2.2 HBT Base-Emitter Capacitance

Higher order coefficients that are associated with the nonlinearity of HBT base-emitter capacitance are shown in Figure 6.9. Both coefficients increase upon cooling. However, unlike the higher order HBT transconductance coefficients, no roll-off is observed beyond a certain point. Instead, the slope of the coefficients increases in the bias range where the nonlinearity of the ideality factor is expected to have a significant influence. As the base-emitter diffusion capacitance is related to the transconductance through a delay term, this discrepancy between the trends of base-emitter capacitance and the transconductance nonlinearity indicates that the base-emitter delay increases significantly as the current density reaches high injection regime. An interesting point is that the deviation in the slopes of c_{be2} and c_{be3} occurs at lower current densities as the device is cooled down.

6.4.2.3 Temperature Dependence of the Charge-Control Capacitance Nonlinearity Under Constant g_{m} Bias

As it was done under the investigation of the transconductance nonlinearity, it is reasonable to consider temperature dependence of the charge-control capacitance

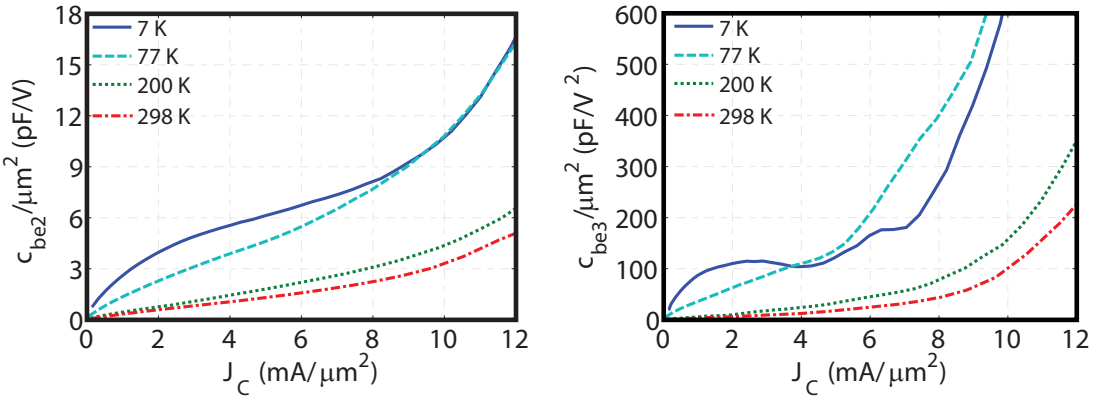
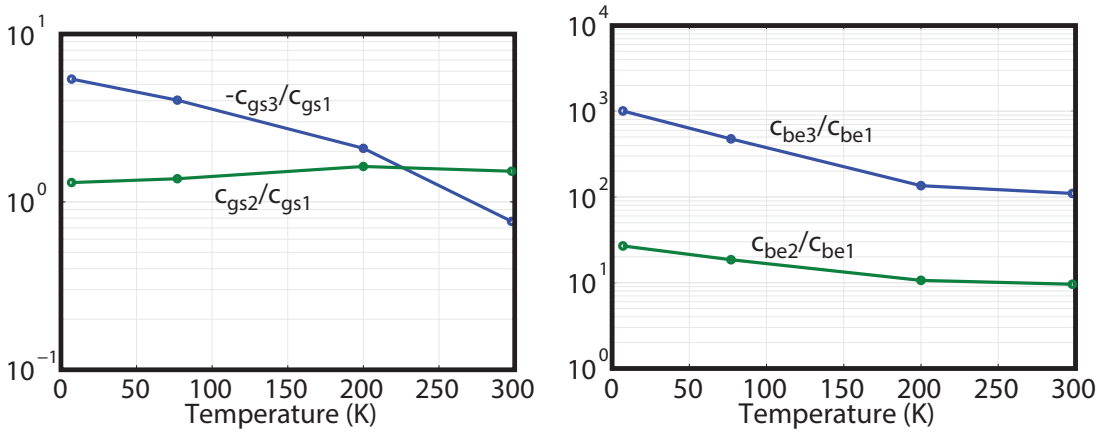


Figure 6.9: (a) Second and (b) third order Taylor series expansion coefficients of the HBT base-emitter capacitance.

nonlinearity at fixed bias points. The bias points that are reported in Table 6.1 are considered. For these bias points, the normalized second (c_{gs2}/c_{gs}, c_{be2}/c_{be}) and third order (c_{gs3}/c_{gs}, c_{be3}/c_{be}) Taylor series expansion coefficients are shown in Figure 6.10.

The trends of the normalized charge-control capacitance nonlinearity look strikingly similar to that of the transconductance. For the HEMT, the normalized c_{gs2} is fairly independent of temperature, whereas the normalized c_{gs3} increases by almost an order of magnitude with cooling from 300 K to 7 K. For the HBT, the normalized values of c_{be2} and c_{be3} increase by factors of 3 and 10, respectively, with cooling from 300 K to 7 K. At all temperatures, the charge-control capacitance is less nonlinear than the transconductance, which is explained by the fact the charge-control capacitance has a depletion component which is weakly bias dependent. Similar to the case of the transconductance, the charge-control capacitance nonlinearity of an HBT is greater than that of a HEMT at all temperatures.



(a) (b)

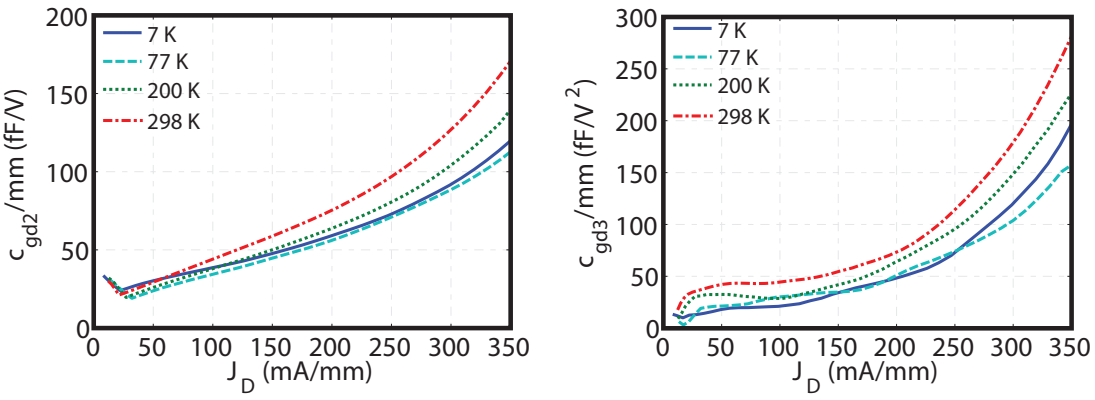
Figure 6.10: Normalized second and third order Taylor series expansion coefficients of the (a) HEMT and (b) HBT charge-control capacitance.

6.4.3 Feedback Capacitance

The feedback capacitance is a depletion capacitance and, as long as it is adequately reverse biased, its nonlinearity is expected to be less than that of the charge-control capacitance. However, due to the Miller effect, the feedback capacitance can have a significant influence on the overall linearity performance of the device.

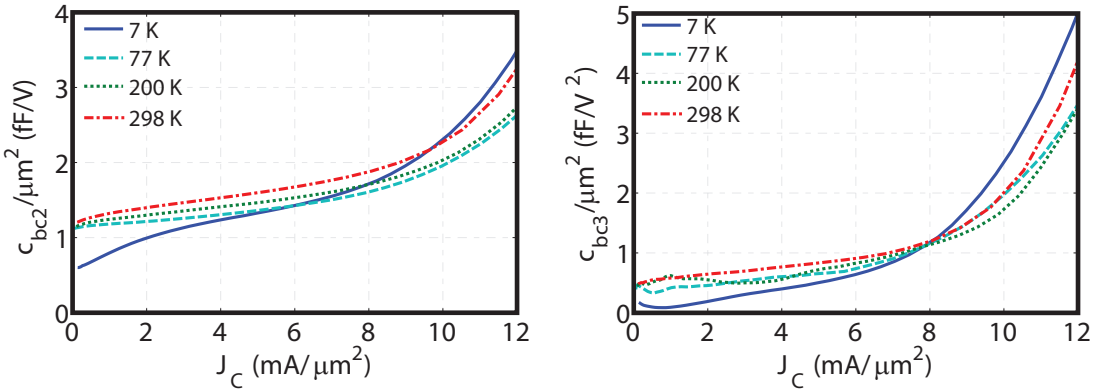
The higher order Taylor series expansion coefficients that represent nonlinearity of the feedback capacitance are shown in Figures 6.11 and 6.12 for the HEMT and HBT, respectively. The nonlinearity of the feedback capacitance increases with current density, as the reverse bias voltage across the capacitance is reduced. This reduction is due to the fact that the input DC voltage and the voltage drop across the series collector resistance are increased with the bias current.

Overall, the nonlinearity of the feedback capacitance decreases with cooling despite the fact that the reverse bias voltage drops. This indicates that the increase in the junction built-in potential with cooling is greater than the drop in the reverse bias voltage. To gain a better perspective, the normalized higher order feedback capacitance coefficients are considered, which are shown in Figure 6.13 for the bias points



(a) (b)

Figure 6.11: (a) Second and (b) third order Taylor series expansion coefficients of the HEMT gate-drain capacitance.



(a) (b)

Figure 6.12: (a) Second and (b) third order Taylor series expansion coefficients of the HBT base-collector capacitance.

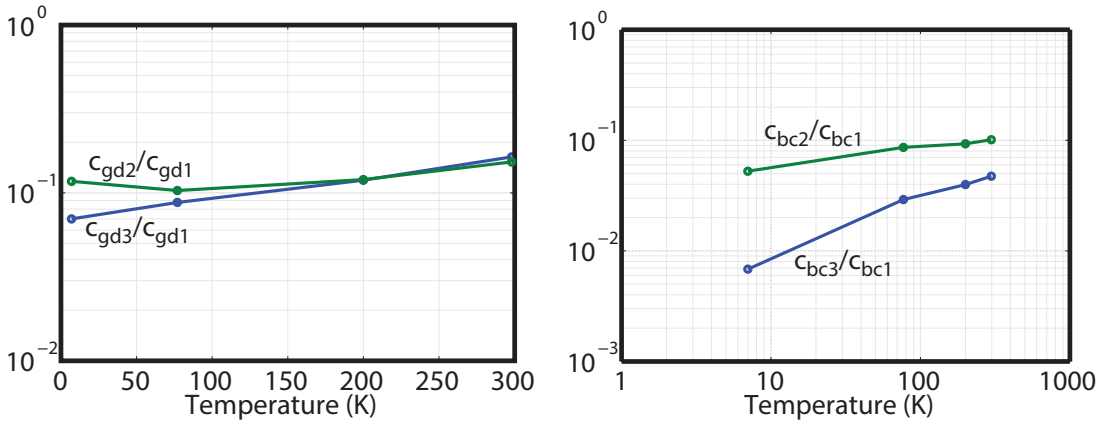
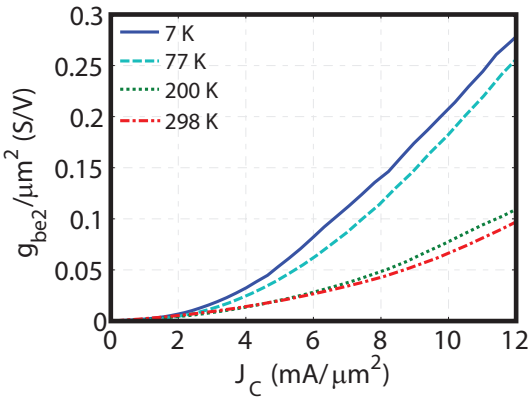


Figure 6.13: Normalized second and third order Taylor series expansion coefficients of the (a) HEMT and (b) HBT feedback capacitance.

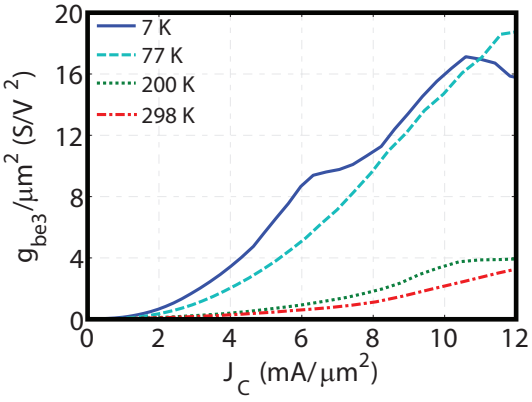
reported in Table 6.1. For both devices, the higher order coefficients decrease more with cooling than the linear coefficient, which indicates that the feedback capacitance becomes more linear at cryogenic temperatures.

6.4.4 HBT Input Conductance

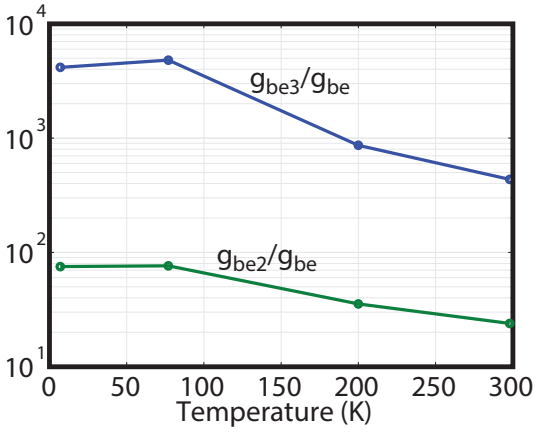
The nonlinearity of the HBT input conductance can be critical at low frequencies. The second (g_{be2}) and third (g_{be3}) order coefficients that represent nonlinearity of the HBT input conductance and their values normalized to the linear coefficient are shown in Figure 6.14. The absolute values of the coefficients increase with cooling. This increase appears to become significant beyond 2 mA/μm^2. For the bias points reported in Table 6.1, the normalized g_{be} coefficients increase with cooling in a trend similar to that of the transconductance. Since the linear input conductance decreases significantly with cooling, the corner frequency, below which the input conductance nonlinearity is critical, is expected to be lower at cryogenic temperatures.



(a)

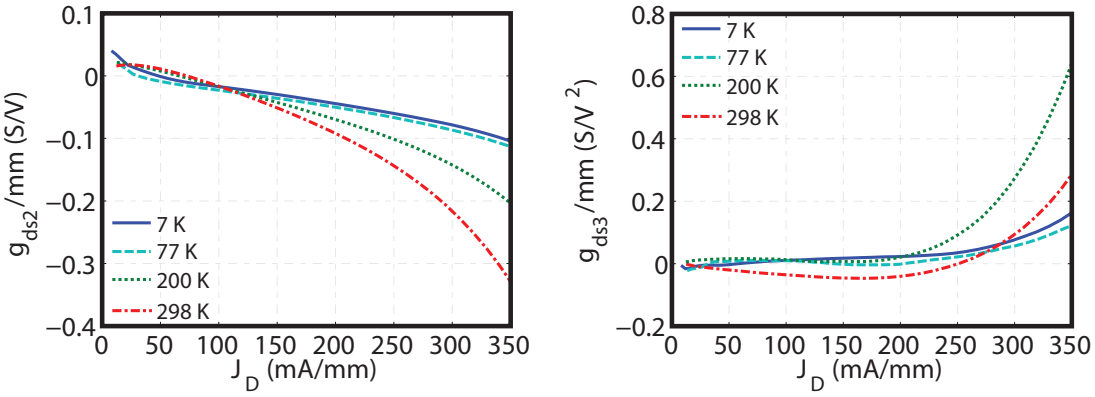


(b)



(c)

Figure 6.14: (a) Second and (b) third order Taylor series expansion coefficients of the HBT input conductance and their (c) normalized values.



(a) (b)

Figure 6.15: (a) Second and (a) third order Taylor series expansion coefficients of the HEMT output conductance.

6.4.5 HEMT Output Conductance

The nonlinearity of the HEMT output conductance is investigated. The second (g_{ds2}) and third (g_{ds3}) order coefficients that represent nonlinearity of the output conductance are shown in Figure 6.15. For the bias range where HEMTs are typically operated at, the output conductance is weakly nonlinear. As the bias current is increased through the gate voltage, the device approaches triode regime and the output conductance nonlinearity increases. This increase is less pronounced at cryogenic temperatures.

The cross-terms (g_{mds}, g_{m2ds}, g_{mds2}) that describe the coupling between the gate and drain modulation of HEMT current were also observed. While it is hard to interpret physical meanings of these terms individually, g_{mds} is weakly temperature dependent, g_{m2ds} increases with cooling and g_{mds2} decreases with cooling for the constant g_{m} biasing that was reported in Table 6.1. Thus, it can be concluded that nonlinearity due to the cross-coupling between the gate and drain of HEMT is expected to be weakly temperature dependent.

6.5 Summary

The nonlinear modeling approach and the extraction results have been discussed in this chapter. Following an explanation of the available methods, the extended small-signal modeling approach, which is based on the definition of Taylor series expansion coefficients of the strong nonlinearities, was described in detail. Model to measurement agreement was provided through comparison of the OIP3 results. The extraction results were found to agree well with the theoretical expectations that were presented in Chapters 2 and 3. Overall, for a practical constant transconductance biasing scheme, the nonlinearities that are modulated by the input were found to increase with cooling whereas the nonlinearities that are excited by the output voltage swings were found to be decreasing or weakly temperature dependent. The effect of intrinsic nonlinearities on the overall intermodulation and dynamic range performance as a function of impedance terminations will be studied in the next chapter.

CHAPTER 7

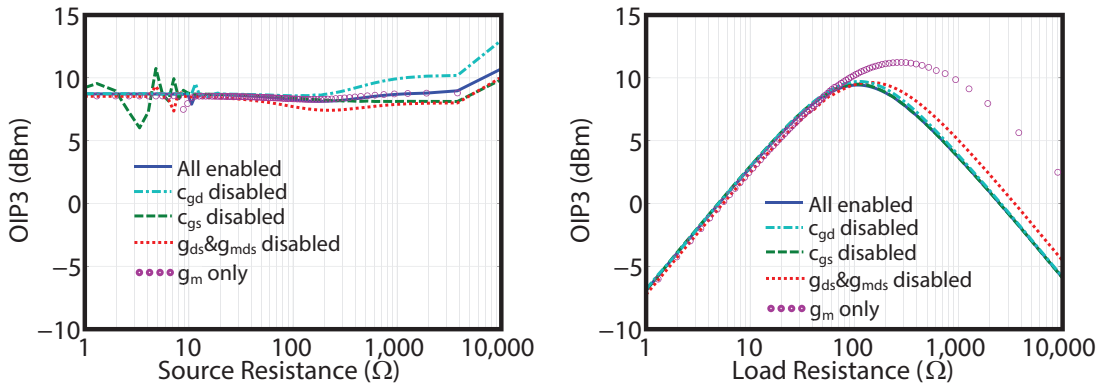
LINEARITY AND DYNAMIC RANGE PERFORMANCE

In this chapter, linearity and dynamic range performance of InP HEMTs and SiGe HBTs is studied as a function of their impedance terminations, using the nonlinear models that were presented earlier. First, the contribution of individual nonlinearities to the overall intermodulation distortion (IMD) of the devices is investigated. Second, the concept of IMD cancellation through optimum out-of-band frequency termination is discussed. Based on this phenomena, linearity performance of the transistors in a narrowband amplifier application is studied. Finally, the linearity and dynamic range performance in the case of a wideband amplifier is evaluated and the devices are compared. Changes in the linearity and dynamic range performance as a function of temperature are interpreted.

7.1 Influence of Individual Nonlinearities

To understand the impact of individual nonlinearities on the transistor intermodulation performance, an analysis was carried out, for which each of the higher order terms associated with the nonlinearities were disabled and changes in the IP3 performance were observed. This was done as a function of source and load resistances. For each case, the source and load resistances were swept one at a time while the other port was terminated in 50 Ω. This analysis was carried out at room temperature, where the input and output nonlinearities are more balanced.

The results are shown in Figure 7.1 for the HEMT. As a function of source resistance, the lack of capacitance and output conductance nonlinearities cause negligible

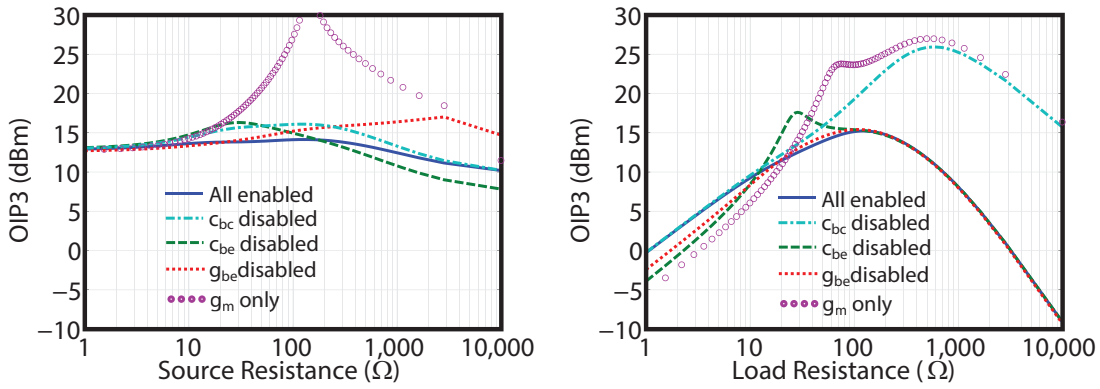


(a) (b)

Figure 7.1: Effect of idividual nonlinearities on the HEMT IP3 as a function of (a) source and (b) load resistances at 10 GHz. J_D=88 mA/mm, W_G=40 μm.

difference in the IP3 performance. However, in terms of the load resistance, the lack of capacitance and output conductance nonlinearities result in a significant IP3 offset beyond 100 Ω. This result is interesting, since the IP3 offset is much smaller when the nonlinearities are turned-off only one at a time, which points out to a complex interaction between the individual nonlinearities.

The analysis results appear in Figure 7.2 for the HBT. For source resistance (R_S) values less than 10 Ω, the transconductance nonlinearity alone is sufficient to describe the overall device performance. However, beyond R_S=10 Ω, the lack of capacitance and input conductance nonlinearities altogether result in a very significant IP3 difference. At high R_S values, the nonlinearity of input conductance becomes critical for the IP3. In terms of the load resistance, the nonlinearity of base-emitter conductance and capacitance has a considerable influence up to 50 Ω. Beyond R_L=50 Ω, the nonlinearity of base-collector capacitance becomes very significant for the IP3 performance. Hence, for the accurate prediction of IP3, all of the nonlinearities that are considered in this work are essential.



(a) (b)

Figure 7.2: Effect of idividual nonlinearities on the HBT IP3 as a function of (a) source and (b) load resistances at 10 GHz. J_C=2.55 mA/μm^2, L_E=36 μm.

7.2 Intermodulation Cancellation

In the case of a two-tone input signal, the individual nonlinearities generate intermodulation and harmonic products with arbitrary amplitude and phase. The concept of IMD cancellation takes advantage of the fact that phases of the IMD products generated by different nonlinearities in the circuit can be aligned such that amplitudes of the IMD products are subtracted from each other. This approach effectively reduces the overall IMD.

The linearization of an amplifier can be achieved in many different ways. Some of the techniques that are used for the linearization are predistortion [168], linear feedback [169], feedforward [170], optimum out-of-band termination [171] and derivative superposition [172]. Out of all these techniques, optimization of the out-of-band terminations is the only approach that does not interfere with the minimum achievable noise, which is the most critical parameter for the design of high sensitivity receivers. Thus, this approach will be investigated for the linearity and dynamic range optimization.

With the assumption that only the transconductance is nonlinear, a Volterra series analysis was performed for a common emitter BJT amplifier in [173] and a closed-

form expression was reported for the third order intercept point (IP3). As it turns out, the IP3 is inversely proportional to a difference expression, which is a function of the termination impedances at the difference ($\Delta\omega$=$\omega_1 - \omega_2$) and second harmonic frequencies(2ω, $\omega_1 + \omega_2$). This difference expression is given as

$$\frac{1}{IP3} \propto \epsilon\left(\Delta\omega, 2\omega\right) = g_{\mathrm{m3}} - \frac{2g_{\mathrm{m2}}^2}{3}\left[\frac{2}{g_{\mathrm{m}} + g\left(\Delta\omega\right)} + \frac{1}{g_{\mathrm{m}} + g\left(2\omega\right)}\right], \tag{7.1}$$

where

$$g\left(\omega\right) = \frac{1 + j\omega C_{\mathrm{CB}}\left[Z_1\left(\omega\right) + Z_3\left(\omega\right)\right] + j\omega C_{\mathrm{BE}}\left[Z_1\left(\omega\right) + Z_{\mathrm{x}}\left(\omega\right)\right]}{\left(1/\beta + j\omega\tau\right)\left[Z_1\left(\omega\right) + Z_{\mathrm{x}}\left(\omega\right)\right] + Z_{\mathrm{x}}\left(\omega\right)}, \tag{7.2}$$

where $Z_1\left(\omega\right)$, $Z_2\left(\omega\right)$, and $Z_3\left(\omega\right)$ are the impedances seen by the intrinsic base, emitter, and collector nodes[1] and

$$Z_{\mathrm{x}}\left(\omega\right) = Z_2\left(\omega\right) + j\omega C_{\mathrm{CB}}\left[Z_1\left(\omega\right)Z_2\left(\omega\right) + Z_1\left(\omega\right)Z_3\left(\omega\right) + Z_2\left(\omega\right)Z_3\left(\omega\right)\right]. \tag{7.3}$$

Thus, if $\epsilon\left(\Delta\omega, 2\omega\right)$ can be set to zero through optimization of the out-of-band source and load terminations, IMD cancellation is achieved. In order for the cancellation to be frequency independent, $g\left(\omega\right)$ should only have a real part. Equation (7.1) is under-determined and the values of out-of-band source and load terminations that result in $\epsilon\left(\Delta\omega, 2\omega\right) = 0$ are not unique. In order for the solutions to be physical, only the solutions with real parts greater than zero are of interest. In this case, the cancellation can only be achieved if g_{m2} and g_{m3} have the same sign [174].

7.3 Linearity of a Narrowband Amplifier

The theoretical analysis that was presented in the previous section provide a perspective on how the IMD cancellation can be achieved. However, it has one significant

[1]This impedance term includes the corresponding series base, emitter, collector resistances and the substrate network.

drawback, which is the fact that it was developed with the assumption that only the transocnductance is nonlinear. However, as demonstrated in Section 7.1, all nonlinearities that are discussed in Chapter 6 are critical for the IP3 performance. For the complete model, an analytical model would be too complicated to solve. Hence, an alternative approach is pursued.

In order to determine the maximum OIP3 ($OIP3_{\mathrm{MAX}}$) as a function of frequency and temperature, a numerical optimization procedure was carried out. For this, the input and output difference ($Z_{\mathrm{S}}(\Delta\omega)$, $Z_{\mathrm{L}}(\Delta\omega)$) and second harmonic ($Z_{\mathrm{S}}(2\omega)$, $Z_{\mathrm{L}}(2\omega)$), and output fundamental ($Z_{\mathrm{L}}(\omega)$) frequency terminations were optimized. The IP3 was found to be insensitive to the third harmonic frequency terminations. The input fundamental frequency termination was set to the optimum noise impedance (Z_{OPT}), which is a practical scenario regarding the design of low-noise amplifiers. Stability was ensured at all termination frequencies. Further details regarding the implementation of stability criteria and optimization environment are provided in Appendix B. The optimization was performed at center frequencies of 1, 10, and 20 GHz. A 10 MHz tone spacing was considered for the optimization. As the tone spacing is increased, the IP3 is observed to roll-off slowly as the coupling between the input and output difference frequency terminations becomes stronger.

The maximum OIP3 and corresponding IIP3 values obtained following the optimization process are reported in Figures 7.3 and 7.4, respectively. The results are reported for the bias points that provide a constant transconductance across the temperature range, as discussed in Chapter 6. The HBT provides better performance, which is explained by the fact the IM3 cancellation is achieved. As bipolar devices have higher order g_{m} coefficients with identical signs due to the exponential I-V relationship, the optimization of the out-of-band terminations enables high linearity performance through IMD cancellation. For the bias points considered for HEMT, $g_{\mathrm{m3,HEMT}}$ is negative while $g_{\mathrm{m2,HEMT}}$ is positive. Thus, the cancellation can not be

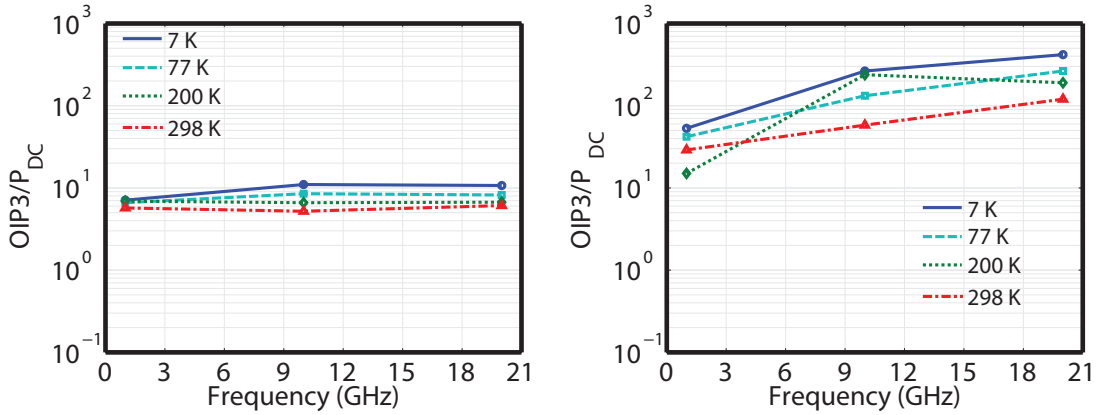


Figure 7.3: Maximum narrowband OIP3 of (a) HEMT and (b) HBT following the optimization.

achieved. Recalling from Chapter 6, $g_{\mathrm{m3,HEMT}}$ is positive for weak overdrive operation at 298 K and 200 K. Indeed, it was verified the cancellation and high IP3 can be achieved in this region. However, the weak overdrive bias regime limits the device performance in terms of noise and gain. A significant frequency slope is observed in maximum IP3 of the HBT, which indicates that the IMD cancellation is more feasible at high frequencies.

Overall, the maximum IP3 of both devices improve with cooling, despite the fact that the transconductance and charge-control capacitance nonlinearities increase. This trend implies that the output (feedback capacitance and output conductance) nonlinearity, which decreases with cooling, compensates for the increasing nonlinearities [2]. At the low frequency end, the improvement in maximum IP3 with cooling is relatively limited, compared to the cases of higher frequencies. This indicates that the output capacitance nonlinearity compensates for the input modulated nonlinearities more effectively at higher frequencies.

[2]This was verified by replacing the 7 K model parameters of interest with that of the 298 K model.

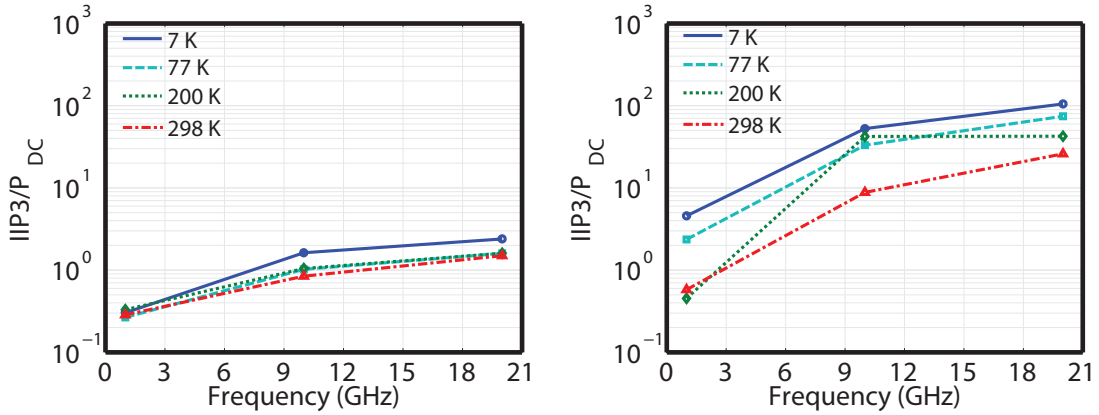


Figure 7.4: IIP3 corresponding to the maximum OIP3 of (a) HEMT and (b) HBT following the optimization.

To gain further insight on how the individual frequency terminations influence the maximum OIP3, source and load pull simulations were performed for the frequency terminations that were optimized. The contours are shown in Figures 7.5-7.9. Several important observations can be made from these figures.

The maximum OIP3 of HEMT is insensitive to $Z_{\mathrm{S}}(\Delta\omega)$. The maximum OIP3 of the HBT becomes less sensitive to $Z_{\mathrm{S}}(\Delta\omega)$ with cooling. This trend can be explained by the decrease in the linear input conductance (g_{be})[3], according to the cancellation theory that was presented in Section 7.2. In the limit of $\Delta\omega = 0$, the feedback path from the intrinsic emitter voltage (created by the output current across the emitter resistance) to the $Z_{\mathrm{S}}(\Delta\omega)$ becomes more resistive with cooling, which reduces the amount $\Delta\omega$ feedback. Thus, $OIP3_{\mathrm{MAX,HBT}}$ is desensitized to the variations in $Z_{\mathrm{S}}(\Delta\omega)$.

The maximum OIP3 of both transistors becomes less sensitive to $Z_{\mathrm{L}}(\Delta\omega)$ with cooling. This is explained as follows. The IP3 dependence on $Z_{\mathrm{L}}(\Delta\omega)$ arises from the output nonlinearity. The second order terms that are associated with the out-

[3]This was verified by scaling the input conductance.

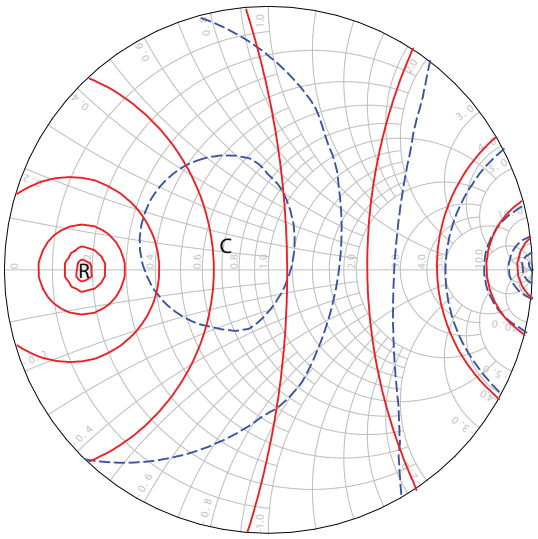
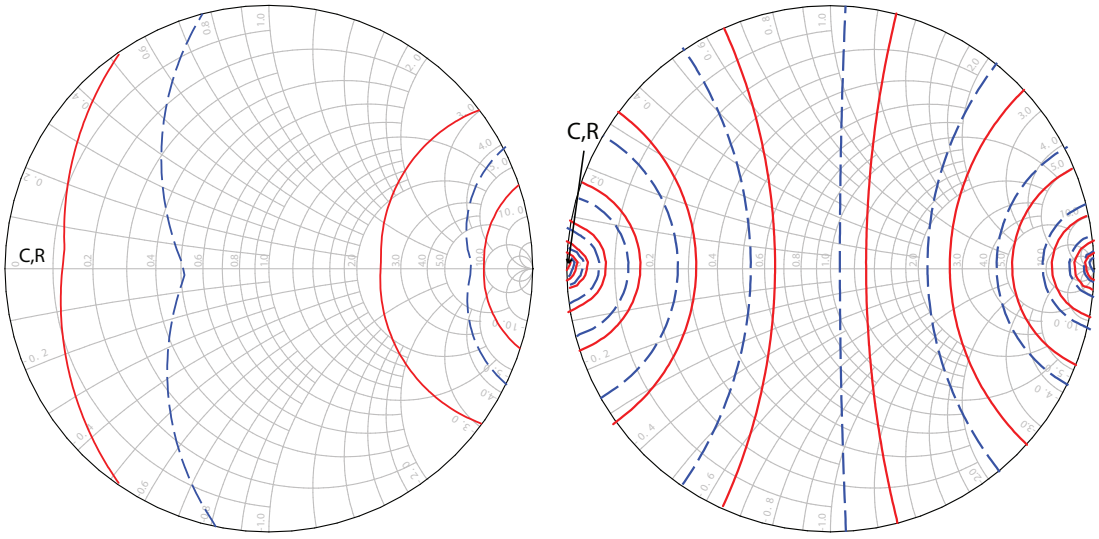
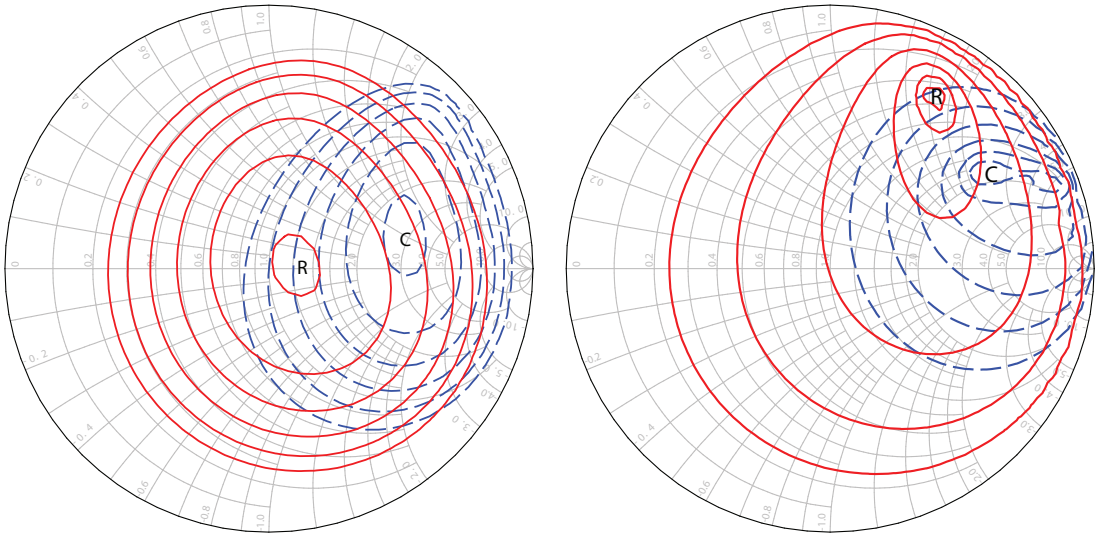


Figure 7.5: Input difference frequency OIP3 contours of the HBT at 10 GHz. Peak contour is 39 dBm, step size is 3 dB. Solid line represents 298 K and dashed line represents 7 K. Peak contours are denoted with R (298 K) and C (7 K).



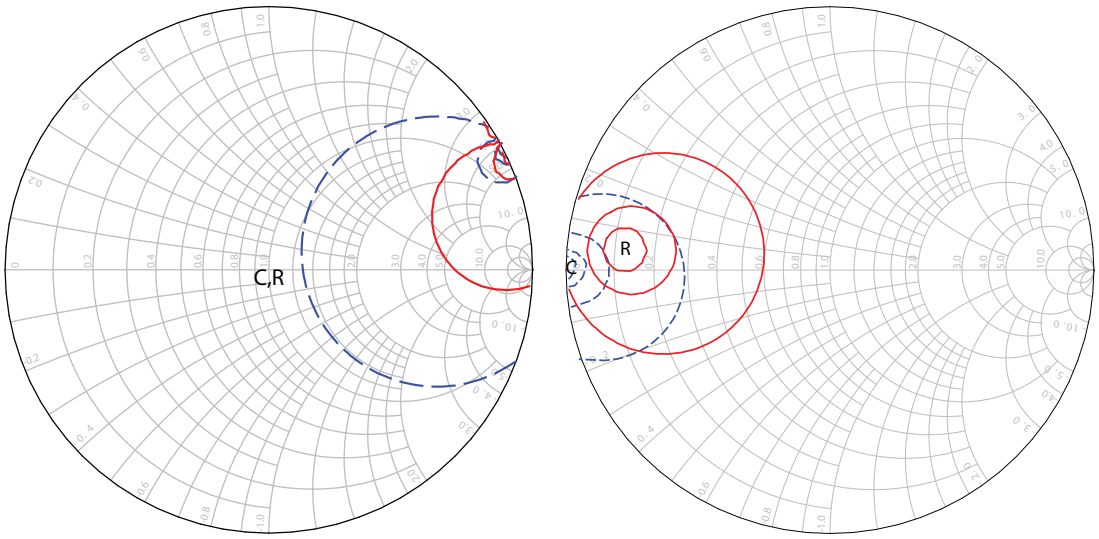
(a) (b)

Figure 7.6: Output difference frequency OIP3 contours of the (a) HEMT and (b) HBT at 10 GHz. Peak contour is 11 dBm and 36 dBm, step size is 1 dB and 3 dB for HEMT and HBT, respectively. Solid line represents 298 K and dashed line represents 7 K. Peak contours are denoted with R (298 K) and C (7 K).



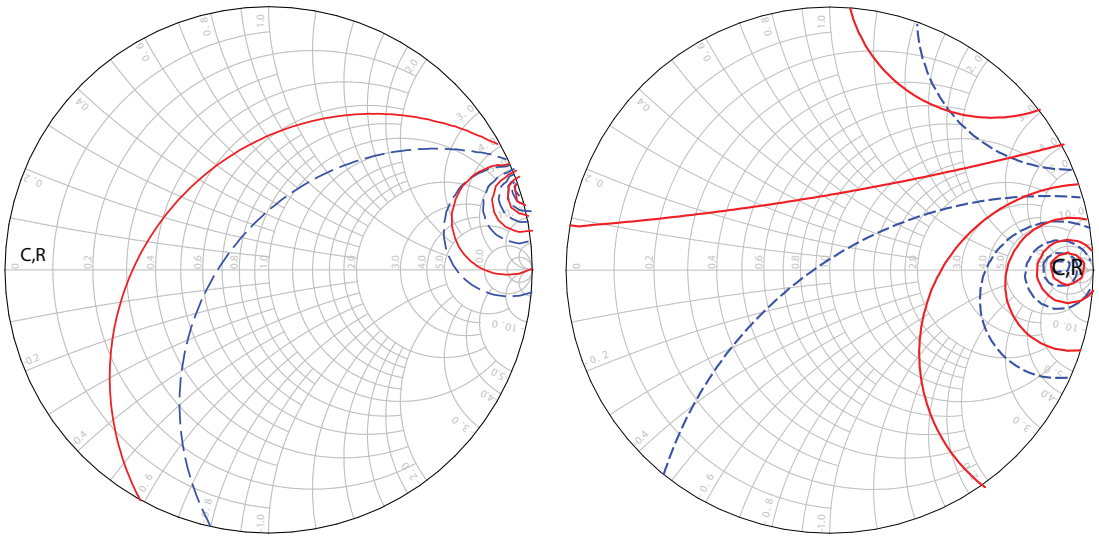
(a) (b)

Figure 7.7: Output fundamental frequency OIP3 contours of the (a) HEMT and (b) HBT at 10 GHz. Peak contour is 11 dBm and 25 dBm, step size is 1 dB and 3 dB for HEMT and HBT, respectively. Solid line represents 298 K and dashed line represents 7 K. Peak contours are denoted with R (298 K) and C (7 K).



(a) (b)

Figure 7.8: Input second harmonic frequency OIP3 contours of the (a) HEMT and (b) HBT at 10 GHz. Peak contour is 11 dBm and 25 dBm, step size is 1 dB and 3 dB for HEMT and HBT, respectively. Solid line represents 298 K and dashed line represents 7 K. Peak contours are denoted with R (298 K) and C (7 K).



(a) (b)

Figure 7.9: Output second harmonic frequency OIP3 contours of the (a) HEMT and (b) HBT at 10 GHz. Peak contour is 11 dBm and 25 dBm, step size is 1 dB and 3 dB for HEMT and HBT, respectively. Solid line represents 298 K and dashed line represents 7 K. Peak contours are denoted with R (298 K) and C (7 K).

put nonlinearity creates third-order IMD tones from the $\Delta\omega$ and ω products at the output node. As the feedback capacitance becomes more linear with cooling due to the increase in the junction built-in potential, the maximum OIP3 varies less with $Z_{\mathrm{L}}(\Delta\omega)$.

For both devices, $OIP3_{\mathrm{MAX}}$ shifts to a higher $Z_{\mathrm{L}}(\omega)$ with cooling. This trend is explained by temperature dependent changes in the input and output nonlinearities. The output fundamental frequency termination controls the voltage swings, which determines the effect of input and output nonlinearities on the overall intermodulation performance [175, 176]. A high $Z_{\mathrm{L}}(\omega)$ results in large output voltage swings, which excites the output nonlinearity, causing linearity to be output limited. A low $Z_{\mathrm{L}}(\omega)$ causes linearity to be input limited, due to the output current generated by the transconductance. Therefore, the overall performance is expected to be maximized at an impedance point where the trade-off between the input and output nonlinearities is optimum. As the input nonlinearity increases and the output nonlinearity decreases

with cooling, $OIP3_{\mathrm{MAX}}$ shifts to a higher $Z_{\mathrm{L}}(\omega)$, where the input nonlinearity is less and the output nonlinearity is more effective on the OIP3. Therefore, the linearity performance is maintained with cooling.

Dependence of $OIP3_{\mathrm{MAX,HEMT}}$ on $Z(2\omega)$ is insignificant at all temperatures, as the regions where $OIP3_{\mathrm{MAX,HEMT}}$ is achieved cover wide areas on the smith chart. The maximum OIP3 of HBT becomes slightly more sensitive to $Z(2\omega)$ with cooling. However, a direct explanation of this trend can not be provided, due to the strong coupling between $Z_{\mathrm{S}}(2\omega)$ and $Z_{\mathrm{L}}(2\omega)$ through the feedback capacitance.

7.4 Linearity and Dynamic Range of a Wideband Amplifier

Following the discussion of narrowband linearity, the optimization process can be utilized to study the intermodulation and dynamic range performance of a wideband amplifier. For this case, a bandwidth of 1–10 GHz was considered. The fundamental frequency terminations were defined with 1 GHz steps in 1–10 GHz frequency range. Similar to the narrowband application, the input fundamental frequency terminations were set to the optimum noise impedance (Z_{OPT}), whereas the output fundamental terminations were optimized for achieving the maximum OIP3. While the input and output difference frequency terminations were also included in the optimization, only open and short circuit configurations were considered for the termination of frequencies greater than 10 GHz.

For an amplifier whose bandwidth is greater than 2:1, the second-order intermodulation (IM2) is as critical as the IM3, since the second-order distortion products appear in-band. To account for this, the IM2 was constrained to be not greater than the IM3 at the noise floor. This can be described with the following function:

$$\left(IM3\left(P_{\mathrm{IN}}\right) - NF\right)/3 - \left(IM2\left(P_{\mathrm{IN}}\right) - NF\right)/2 \geq 0, \tag{7.4}$$

where P_{IN} is the applied input power and NF is the noise floor for which 1 GHz bandwidth and equivalent ambient temperature are assumed.

For an arbitrary two tone excitation, the IM2 frequency is not necessarily a factor of two greater than that of the input. For input signals that are widely spaced apart, the IM2 frequency can be less than or in-between the input excitation frequencies[4]. However, these different combinations of IM2 mixing frequencies could not be entirely integrated into the optimization process. Therefore, only the IM2 products at the second harmonic frequency was considered, for the two input tones that are spaced 10 MHz apart.

The wideband optimization results are shown in Figure 7.10 for different bias points. In general, the linearity is limited in the 1-5 GHz range, where the second harmonic frequency terminations coincide with the other fundamental frequency terminations. For the HEMT, weak overdrive operation (35 mA/mm) at room temperature is a striking example of this trade-off. In this case, the IM3 cancellation is possible, since $g_{\mathrm{m3,HEMT}}$ and $g_{\mathrm{m2,HEMT}}$ have the same sign. However, due to the impedance termination restrictions, the IM3 cancellation can not be realized in the 1-5 GHz range.

At nominal current densities, the wideband operation still results in an IP3 penalty compared to the narrowband case. For the HEMT, this penalty gets worse with cooling and a current density higher than 50 mA/mm is required at 7 K to improve $OIP3_{\mathrm{MAX}}$ in the 1 - 5 GHz range. For the HBT, the IP3 penalty reduces with cooling. Overall, the HBT provides a better performance than the HEMT at any temperature, as an HBT can achieve IMD cancellation for a wide range of bias points.

Following the linearity analysis, the spurious-free dynamic range (SFDR) performance is investigated for devices that are sized to yield R_{OPT}=50 Ω at 10 GHz. This

[4]In a 1-10 GHz amplifier, 6 and 8 GHz input signals can create an IM2 product at 2 GHz, whereas 3 and 9 GHz signals can create an IM2 product at 6 GHz.

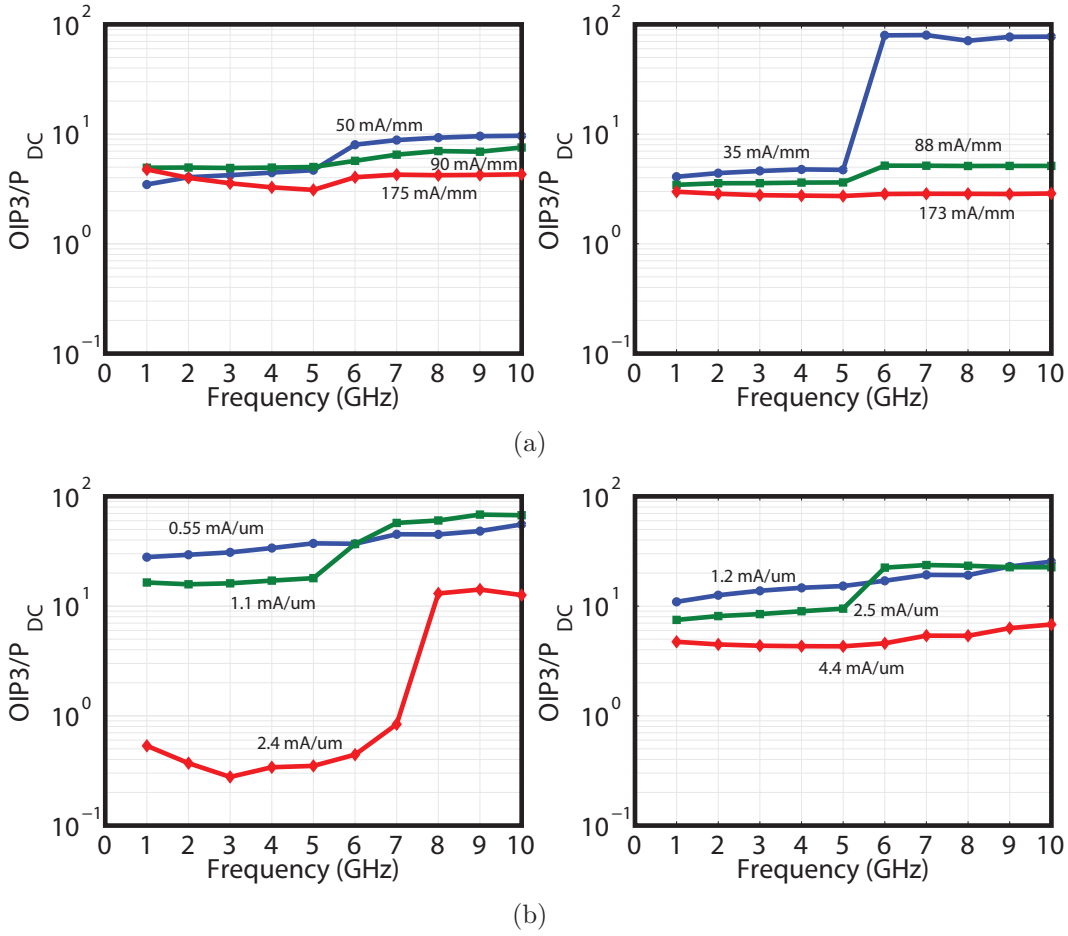


Figure 7.10: Wideband normalized OIP3 of the (a) HEMT and (b) HBT. Left column represents 7 K and right column represents 298 K.

is a common approach in the design of low-noise amplifiers. For the intrinsic HEMT, R_{OPT} is given as [81]

$$R_{\mathrm{OPT,HEMT}} = \sqrt{\left(\frac{f_{\mathrm{t}}}{f}\right)^2 \frac{r_{\mathrm{gs}}}{g_{\mathrm{ds}}} \frac{T_{\mathrm{A}}}{T_{\mathrm{D}}} + r_{\mathrm{gs}}^2}. \tag{7.5}$$

which indicates that $R_{\mathrm{OPT,HEMT}}$ decreases with ambient temperature (T_{A}) assuming that the non-quasi static gate-source resistance (r_{gs}) and output conductance (g_{ds}) do not change significantly with cooling. For HBTs, R_{OPT} is given as [26]

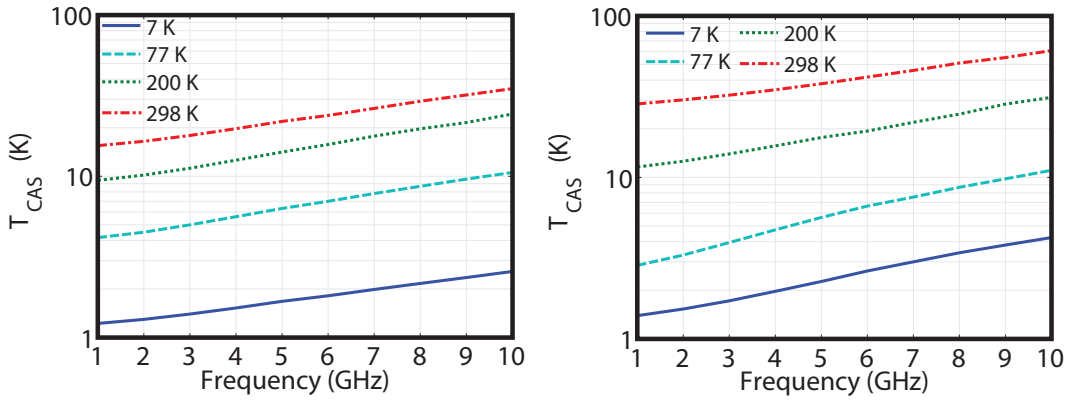
$$R_{\mathrm{OPT,HBT}} = \frac{\beta_{\mathrm{DC}}}{g_{\mathrm{m}}\left(1+\beta_{\mathrm{DC}}\left(f/f_{\mathrm{t}}\right)^2\right)} \sqrt{\frac{1}{\beta_{\mathrm{DC}}} + 2\left(\frac{1}{\beta_{\mathrm{DC}}} + \left(\frac{f}{f_{\mathrm{t}}}\right)^2\right) \frac{g_{\mathrm{m}}\left(R_{\mathrm{B}}+R_{\mathrm{E}}\right)}{n_{\mathrm{c}}}}. \tag{7.6}$$

Hence, $R_{\mathrm{OPT,HBT}}$ is also expected to decrease with cooling as β_{DC} and f_{t} increase. For both devices, the decrease in R_{OPT} with cooling indicates that smaller devices are required at cryogenic temperatures.

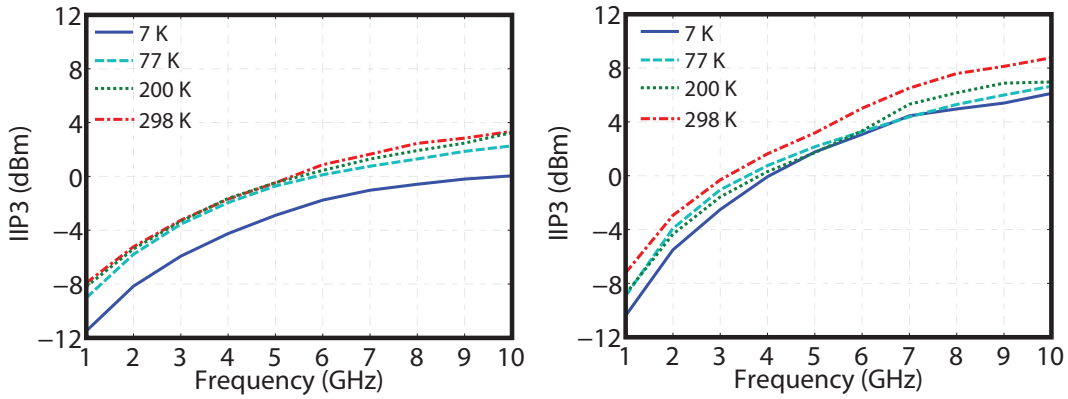
The device sizes and the corresponding power consumption values are reported in Table 7.1 for the bias points that are considered for the SFDR analysis. For the SFDR calculations, the cascaded noise temperature is considered, which takes the impact of gain on noise performance into account as

$$T_{\mathrm{CAS}} = \frac{T_{\mathrm{E}}}{1 - \dfrac{1}{G_{\mathrm{ASSOC}}}}. \tag{7.7}$$

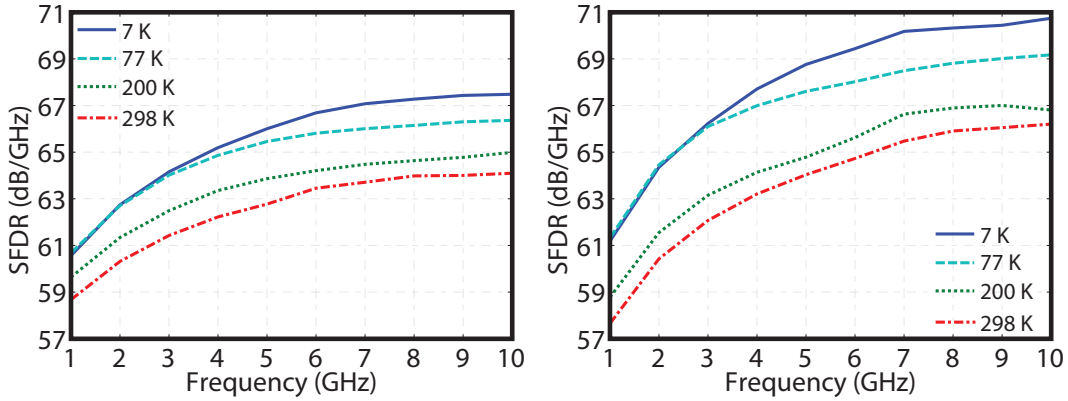
The noise, IIP3 and SFDR results obtained following the optimization process are shown in Figure 7.11. Overall, the HEMT provides a better T_{CAS}, especially towards the end of the frequency range. For both devices, the IIP3 corresponding to the maximum OIP3 ($IIP3_{\mathrm{MAX}}$) slightly degrades with cooling, which is partially due to fact that smaller devices are required at cryogenic temperatures. However, this



(a)



(b)



(c)

Figure 7.11: Optimized dynamic range results. (a) T_{CAS}, (b) IIP3 and (c) SFDR. Left column represents the HEMT and right column represents the HBT.

Table 7.1: Selected Bias Points for the SFDR Analysis

HEMT				HBT		
J_D mA/mm	W_G μm	P_{DC} mW	Temperature K	J_C mA/μm^2	L_E μm	P_{DC} mW
50	130	3.9	7	0.56	13	0.9
57.5	260	9	77	0.79	33	3.1
70	293	12.3	200	1.02	32	3.9
87.5	310	16.3	298	1.2	34	4.7

degradation is not significant, considering the temperature dependence of intrinsic nonlinearites. Since the degradation in $IIP3_{MAX}$ is not greater than the improvement in the noise performance, the SFDR improves with cooling.

Overall, both devices provide similar SFDR performance. At the high frequency end, the HBT is better due to the frequency slope of $IP3_{MAX,HBT}$, which was also observed in the case of narrowband linearity. An important point is the DC power consumption. At any temperature, the HEMT requires 3-4 times more DC power than the HBT, while providing a relatively similar SFDR performance. Thus, a practical conclusion is that HBTs provide better SFDR performance than HEMTs for a given DC power consumption.

7.5 Summary

The temperature dependence of the linearity and the dynamic range performance as a function of impedance terminations has been investigated in this chapter. The conditions that are required for the intermodulation cancellation were discussed first. The linearity and dynamic range optimization results were presented for the narrowband and broadband matching scenarios. For a given DC power consumption, the HBT was found to provide better performance due to the achievability of the intermodulation cancellation. Application of the weakly nonlinear models developed in

this work to characterize linearity of cryogenic LNAs will be demonstrated in the next chapter.

CHAPTER 8

AMPLIFIER CHARACTERIZATION

Noise and linearity measurements of the wideband amplifiers are reported in this chapter. The goal is to demonstrate that the modeling approach developed in this work advances state-of-the-art in cryogenic device modeling, which can be used in the design process of high dynamic range cryogenic amplifiers. Since the designs are not a part of this work, this chapter will only focus on the measurement results.

8.1 4-12 GHz InP HEMT MMIC LNA

A 4-12 GHz InP HEMT cryogenic LNA was characterized [177]. The MMIC amplifier was fabricated with the 0.1 μm InP HEMT process investigated in this work. The amplifier consists of three common-source stages with identical gate-width of 200 μm. The interstage and output matching networks are on-chip whereas the input matching network is realized externally. The schematic and package photo of the amplifier are shown in Figures 8.1 and 8.2, respectively. The amplifier consumes 23 mW and 36 mW DC power at 17 K and 298 K ambient temperatures, respectively. Two modules were borrowed from California Institute of Technology for characterization. The Microwave Office files required for simulations were also provided by the same source.

The cryogenic noise measurement was performed using the cold attenuator method, which enables accurate measurement of extremely low noise levels [178]. For this, a 20 dB attenuator was placed in front of the amplifier in the cryostat. Both the amplifier and attenuator were then cooled down to the base temperature. Knowing

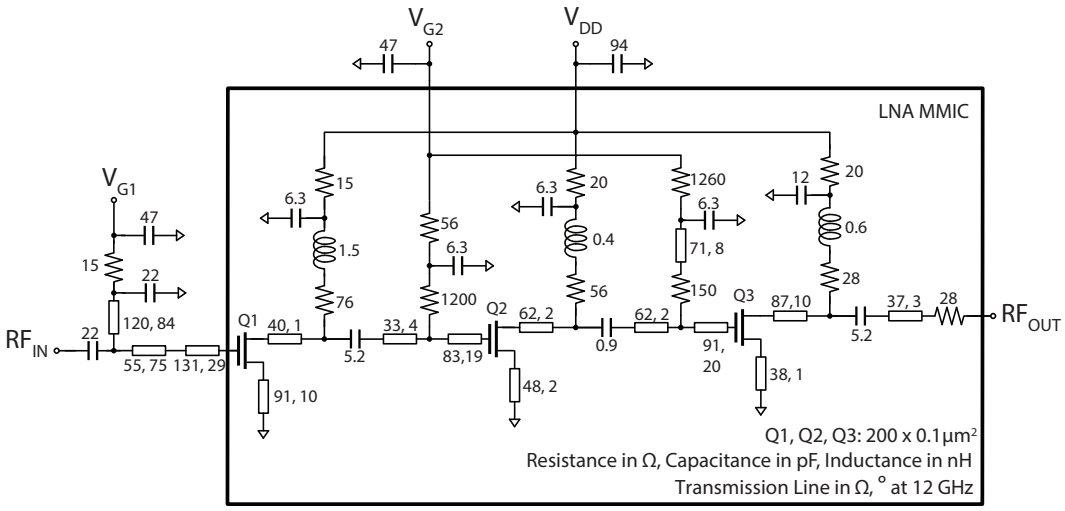


Figure 8.1: Schematic of the 4–12 GHz InP HEMT amplifier.

the noise contribution of the attenuator, the noise temperature of the amplifier was determined through Y-factor measurements using a Keysight N9030A PXA signal analyzer with noise figure measurement capability [179]. A block diagram of the amplifier measurement setup is provided in Figure 8.3.

The noise and gain measurement results of the HEMT amplifier are shown along with simulation results in Figure 8.4. Very good agreement was observed between the measurement and simulation, which is critical for validating the small-signal noise model extracted in this work. The amplifier achieves 4-7 K noise temperature and 35 dB gain across the band at 17 K ambient temperature. With cooling, the noise temperature of the amplifier improves by a factor of ten. The amplifier has greater than 15 dB return loss at all temperatures.

Coaxial OIP3 measurements were performed with the network analyzer. The measurement was calibrated to the cryostat walls. (Fig. 8.3). Measurement and simulation results appear in Figure 8.5. Good agreement between the simulation and measurement was observed, with the maximum disagreement being 2 dB. The offset between the simulation and measurement results can be partially explained by the fact the gate and drain of each amplifier stage can not be biased independently (Fig. 8.1),

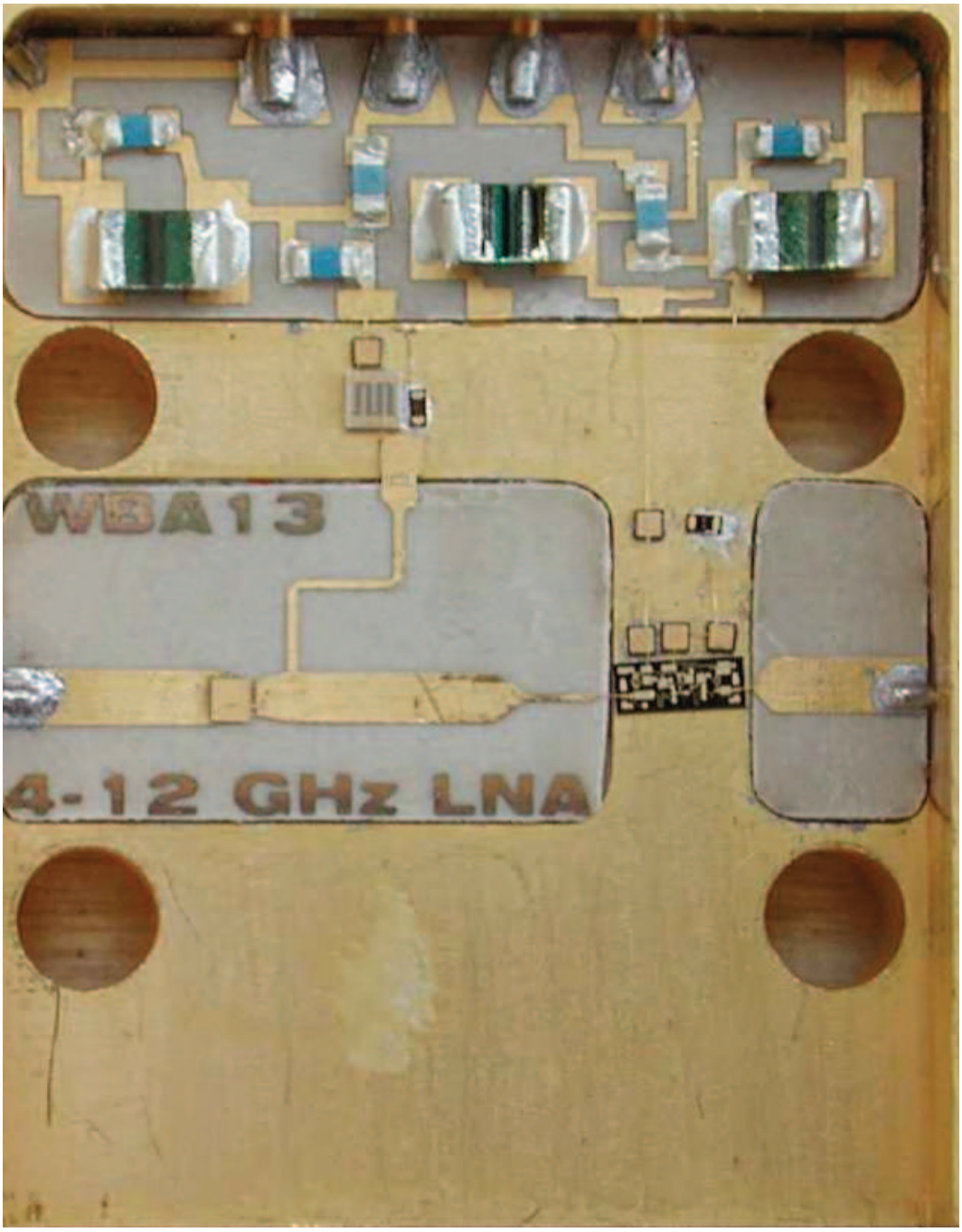


Figure 8.2: Package photo of the 4–12 GHz InP HEMT amplifier.

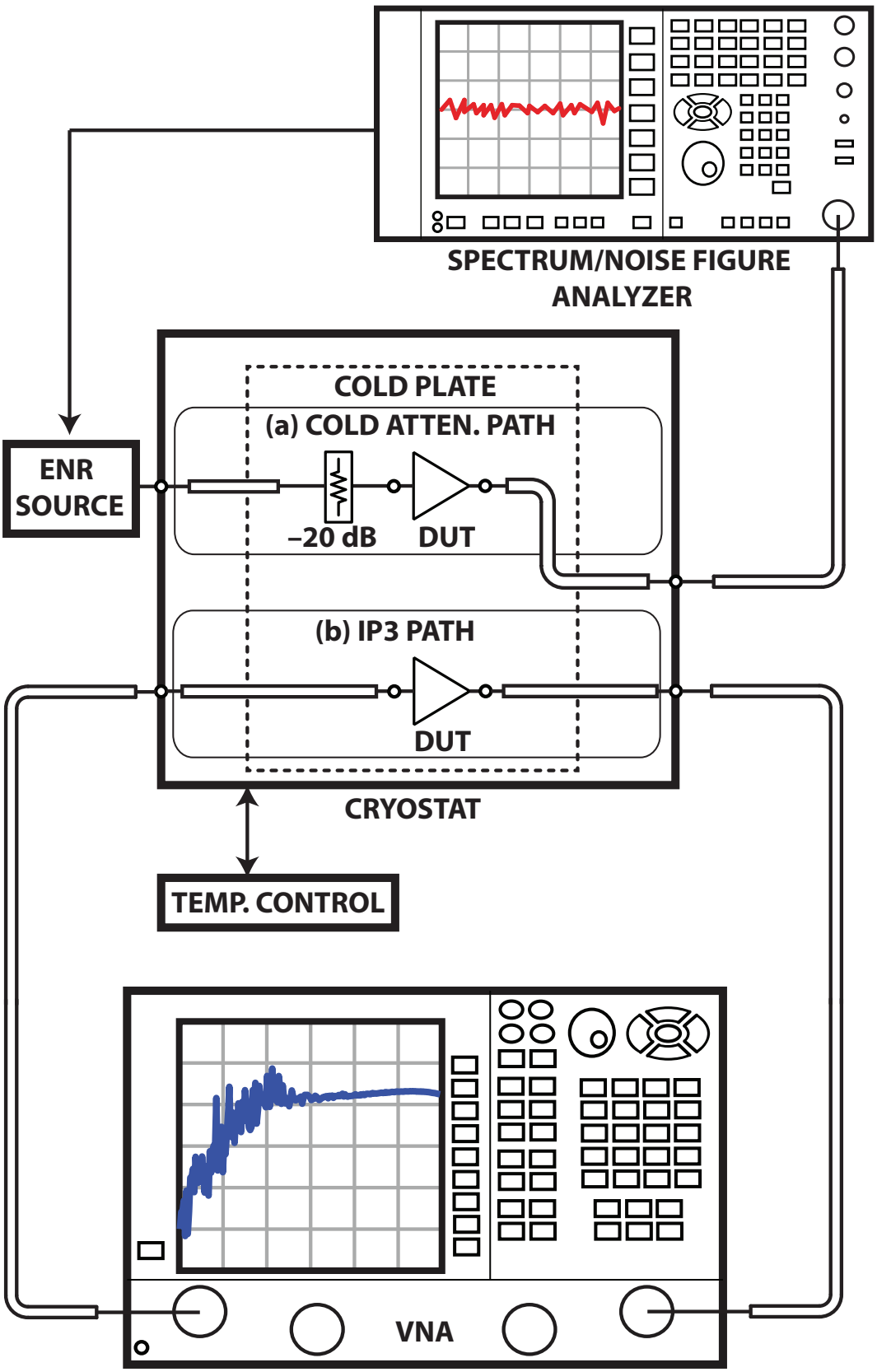
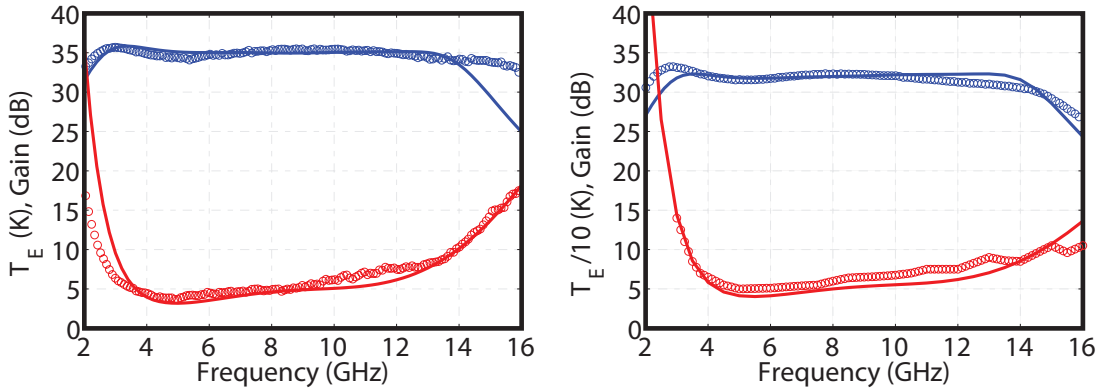
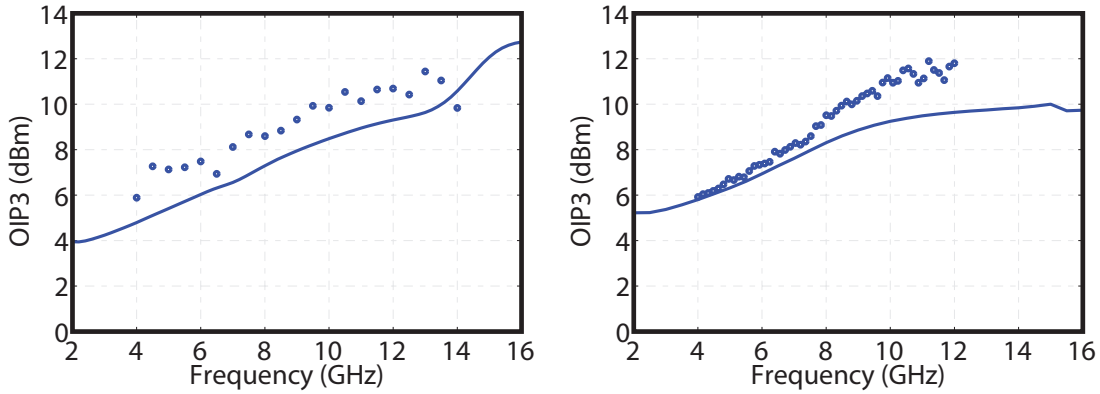


Figure 8.3: Cryogenic amplifier measurement setup.



(a) (b)

Figure 8.4: Noise temperature and gain of the InP HEMT amplifier at (a) 17 K and (b) 298 K. Solid lines represent the simulation results and markers represent the measurements. Room temperature noise is scaled down by a factor of ten.



(a) (b)

Figure 8.5: OIP3 of the InP HEMT amplifier at (a) 17 K and (b) 298 K. Solid lines represent the simulation results and markers represent the measurements.

which could not be exactly captured in the simulation, since the transistor model is only available at a fixed drain-source voltage. A frequency slope is observed in the OIP3 despite the flat gain profile. The OIP3 changes weakly with cooling. For the given noise and IP3 results, the SFDR is 47-50 dB/GHz at 17 K ambient temperature across the entire frequency range of the amplifier. A 3 dB SFDR improvement was observed with cooling.

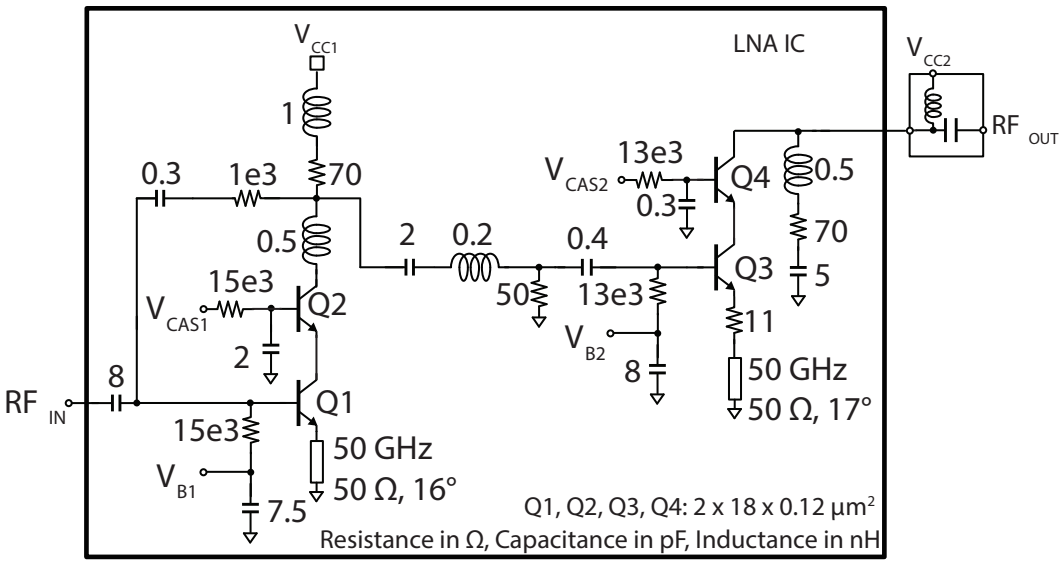


Figure 8.6: Schematic of the 1–20 GHz SiGe HBT amplifier.

8.2 1-20 GHz SiGe HBT IC LNA

A 1-20 GHz SiGe HBT cryogenic LNA was characterized. The integrated circuit amplifier was fabricated with the 0.12 μm SiGe HBT process investigated in this work. The amplifier was designed by Wei-Ting Wong, using the model developed in this work [180]. The two-stage amplifier consists of cascoded common-emitter transistors with 36 μm emitter-length. The schematic and package photo of the amplifier are shown in Figures 8.6 and 8.7, respectively. The amplifier consumes 61 mW at 17 and 77 K.

The noise of the packaged amplifier was measured with the cold attenuator method at 17 K. The measurement and simulation results appear in Figure 8.8. At low frequencies, a considerable mismatch in the noise performance was observed between the simulation and measurement. This is potentially due to incomplete characterization of the passive networks and/or localized self heating as the transistors are operated at high current densities. The gain ripple measured in-band also deviates some from simulation. This was found to be due to the packaging, as the on-wafer measurement was observed to agree better with the simulation.

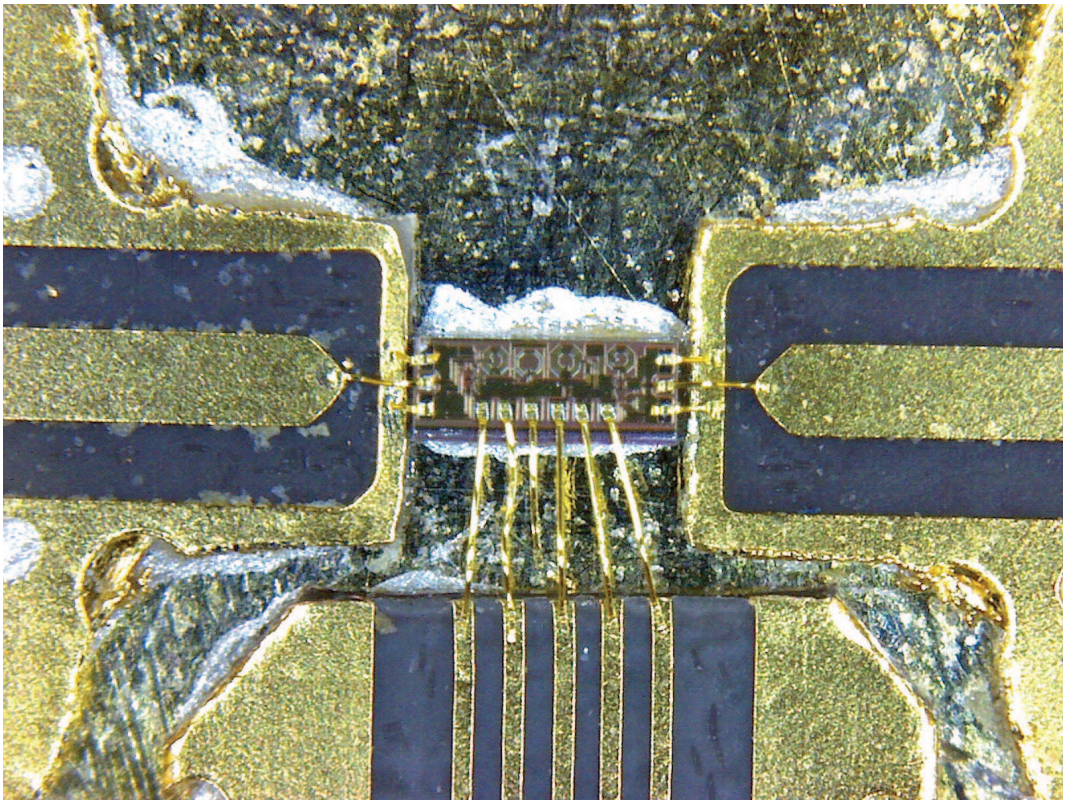
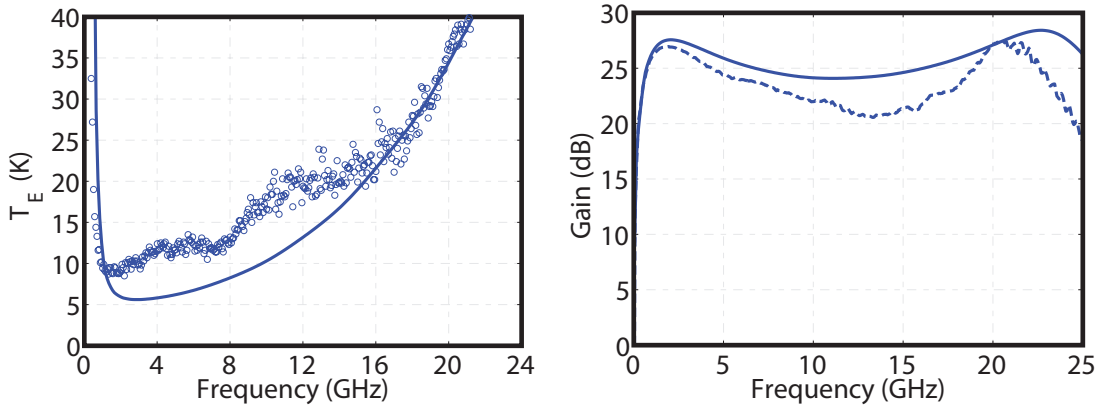
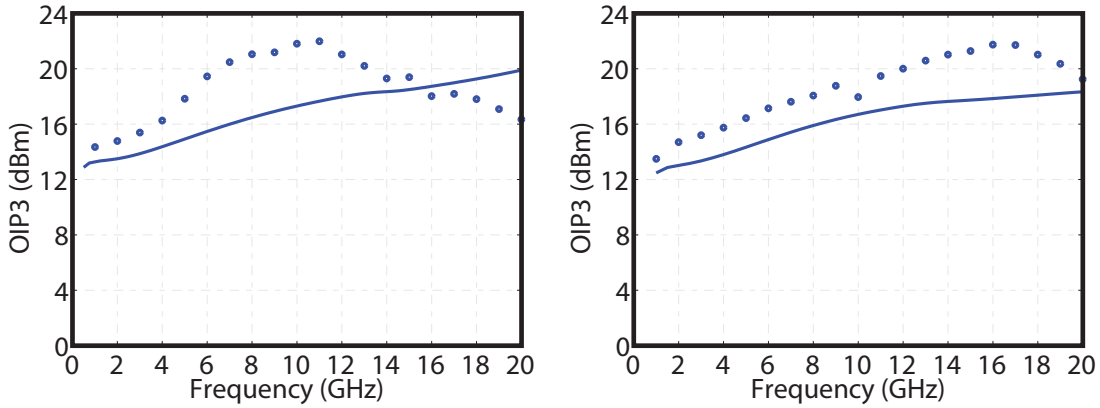


Figure 8.7: Package photo of the 1–20 GHz SiGe HBT amplifier.



(a) (b)

Figure 8.8: (a) Noise temperature and (b) gain of the SiGe HBT amplifier at 17 K. Solid lines represent the simulation results, markers and dashed lines represent the measurements.



(a) (b)

Figure 8.9: OIP3 of the SiGe HBT amplifier at (a)15 K (packaged) and (b) 77 K (on-wafer). Solid lines represent the simulation results and markers represent the measurement.

The OIP3 was measured on-wafer at 77 K and for the packaged amplifier at 15 K. The measurement and simulation results are shown in Figure 8.9. Good agreement is observed between the simulation and measurement. The packaged amplifier result has a slightly different shape than that of the simulation which is believed to be related to the gain mismatch. Overall, the amplifier achieved an SFDR of 60 dB/GHz at 17 K.

8.3 Comparison of the Wideband Cryogenic Amplifiers

Critical performance metrics of the amplifiers investigated in this work are compared to those of other cryogenic amplifiers reported in the literature. The results are provided in Table 8.1. The InP HEMT amplifier investigated in this work provides a competitive SFDR performance compared to other state-of-the-art cryogenic amplifiers reported in the literature. The 1–20 GHz SiGe HBT amplifier, which was designed using the models developed in this work, achieves 10 dB greater SFDR than any other cryogenic amplifier.

Table 8.1: Dynamic Range Of Cryogenic Low-Noise Amplifiers

Platform -	Frequency GHz	P_{DC} mW	T_E K	Gain dB	IIP3 dBm	SFDR dB/GHz
GaAs HEMT [181]	1-18	16	6.5	32	-25*	49
InP HEMT [182]	0.3-14	10	4.1	42	-42	39
InP HEMT [183]	4-16	10	3.7	43	-45	37
InP HEMT**	4-12	23	5	35	-27	49
SiGe HBT [184]	0.5-4	13	8	36	-29*	46
SiGe HBT [23]	0.1-5	20	4	31	-25	50
SiGe HBT**	1-20	60	18	23	-5	60

∗IP3 predicted from P1dB data.
∗∗ **Amplifiers characterized in this work.**

8.4 Summary

Noise and IP3 measurements of the wideband InP HEMT and SiGe HBT amplifiers have been presented in this work. Overall, a good agreement is observed in both cases, which validates the modeling approach presented earlier. Furthermore, the SiGe HBT model enabled the design of a wideband cryogenic low-noise amplifier with highest spurious-free dynamic range known to date.

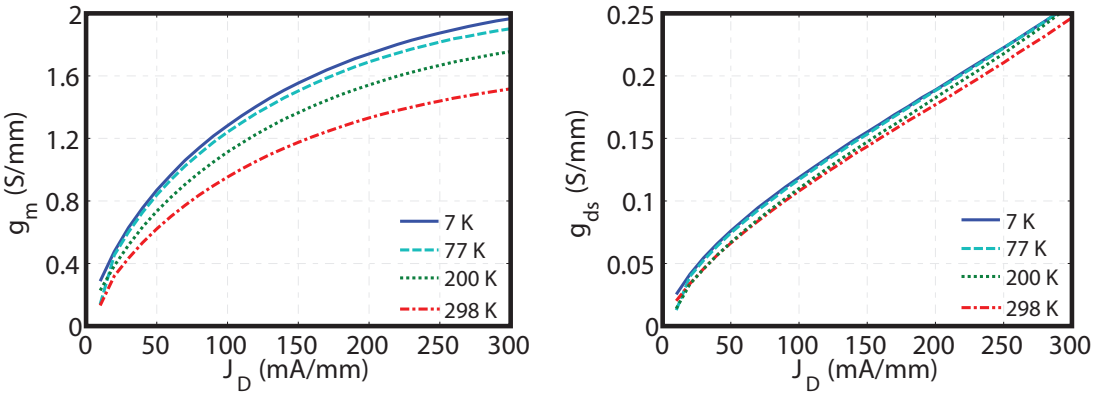
CHAPTER 9

INVESTIGATION OF ALTERNATIVE TRANSISTORS

The core of this thesis has been dedicated to understanding dynamic range performance of InP HEMTs and SiGe HBTs as a function of temperature. However, within the context of the theme, other devices such as SOI MOSFETs and InP HBTs were also considered. While the full dynamic range analysis was not performed, small-signal noise performance of these devices were studied at cryogenic temperatures. In this chapter, the key results will be summarized.

9.1 32 nm SOI MOSFET

At the onset of this work, the RF performance of short-channel MOSFETs at cryogenic temperatures had not received much attention in the literature. To investigate the performance of nanometer CMOS technology, a 32 nm gate-length SOI MOSFET process fabricated by IBM (now Global Foundries) was characterized. S-parameter measurements of 100 μm gate-width transistors along with pad-open-short deembedding structures [131] were performed in the temperature range of 7-300 K. The small-signal noise model is identical to that of an InP HEMT and the extraction techniques are quite similar. Hence, these will not be discussed. The details can be found in [138], where the results of this study are published. Here, the key changes in the small-signal model parameters and noise performance with cooling will be discussed.



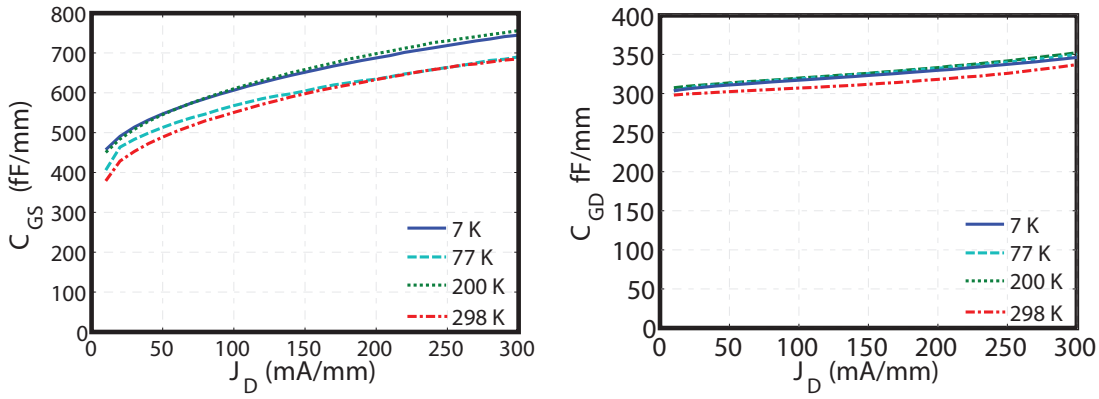
(a) (b)

Figure 9.1: (a) Transconductance and (b) output conductance of the SOI MOSFET.

9.1.1 Temperature Dependence of the Small-Signal Model Parameters

Carrier freeze-out is a significant concern for the operation of silicon devices below liquid nitrogen temperatures [185]. However, it has been known that sub-100 nm devices do not suffer from this problem and their performance improve with cooling down to 4.2 K [186], and the results of this study support this trend. The extracted transconductance ($g_{\mathrm{m,MOSFET}}$) and output conductance (g_{ds}) values are shown in Figure 9.1. At a fixed current density, the transconductance is observed to improve with cooling. Unlike HEMT, $g_{\mathrm{m,MOSFET}}$ does not have a peak value with respect to the current density. In addition to the increase in $g_{\mathrm{m,MOSFET}}$ with cooling, a decrease in extrinsic resistances are observed, which further improves the performance at cryogenic temperatures.

The output conductance only slightly changes with cooling. This result is quite interesting, since the output conductance of a HEMT was observed to increase with cooling in a similar trend to that of its transconductance. This indicates that g_{ds} of MOSFET is dominated by the channel-length modulation rather than the velocity saturation. Since a lower g_{ds} is desired for high voltage gain, the temperature dependence of MOSFET g_{ds} is more ideal than that of the HEMT.



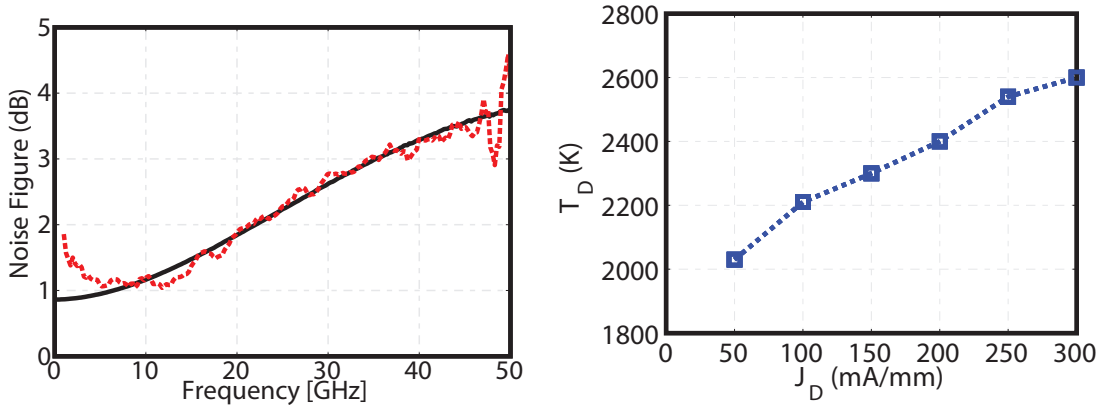
(a) (b)

Figure 9.2: (a) Gate-source and (b) gate-drain capacitance of the SOI MOSFET.

The extracted gate-source (C_{GS}) and gate-drain (C_{GD}) capacitance values of the SOI MOSFET are shown in Figure 9.2. The gate-source capacitance increases with cooling which indicates that the carrier distribution is closer to gate at cryogenic temperatures. The gate-drain capacitance is weakly temperature dependent, which indicates that the drain is strongly doped and its carrier concentration does not change with cooling. For the HEMT, C_{GD} was observed to decrease with cooling due to the increase in the Schottky junction built-in potential. Since a lower feedback capacitance is desired for a transistor, the temperature dependence of HEMT C_{GD} is more preferable than that of the MOSFET.

9.1.2 Noise Modeling

Unlike the HEMTs characterized in this work, the drain noise temperature (T_{D}) of the short-channel MOSFETs had not been previously reported in the literature. Therefore, T_{D} was extracted at room temperature through on-wafer NF50 measurements [187]. The vector-corrected on-wafer NF50 measurements were performed with the network analyzer, using the cold-source method [188]. The drain noise temperature was obtained by back-fitting noise simulation of the model to the NF50 measurement in 1–40 GHz range. Example fit and extracted values of T_{D} are shown in



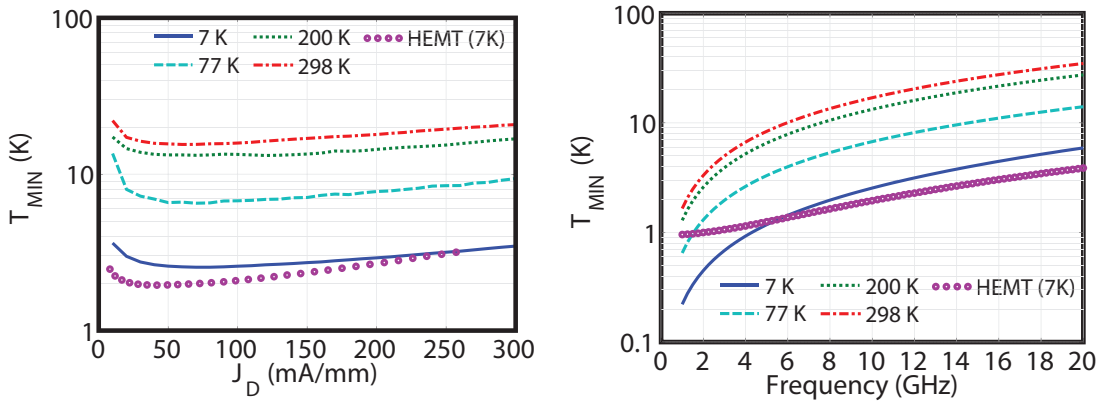
(a) (b)

Figure 9.3: (a) Modeled and measured 50 Ω noise of the SOI MOSFET and (b) extracted T_{D} values.

Figure 9.3. Good agreement between the model and measurement is observed and T_{D} increases with bias as expected.

The dependence of T_{D} on physical temperature for the short channel MOSFET is unknown. The results of a 65 nm MOSFET reported in [189] indicate that the drain noise current does not change with cooling from room temperature to 77 K. In alignment with this trend, T_{D} was assumed to be independent of physical temperature in this work. With the knowledge of small-signal model parameters and T_{D}, the temperature dependence of the noise parameters was investigated.

The minimum achievable noise (T_{MIN}) is shown as a function of bias and frequency in Figure 9.4. A constant g_{m} biasing scheme is considered for the frequency dependence of T_{MIN}. The minimum noise temperature of HEMT investigated in this work is shown at 7 K for comparison. A steady improvement with cooling is observed for the MOSFET and the predicted performance is close to that of the HEMT at 7 K. To achieve the lowest possible T_{MIN}, the MOSFET requires a current density of 75 mA/mm, whereas the HEMT requires 40 mA/mm. At the low frequency end, the broadband noise of HEMTs are limited by the gate leakage current whereas MOSFETs do not have this problem.



(a) (b)

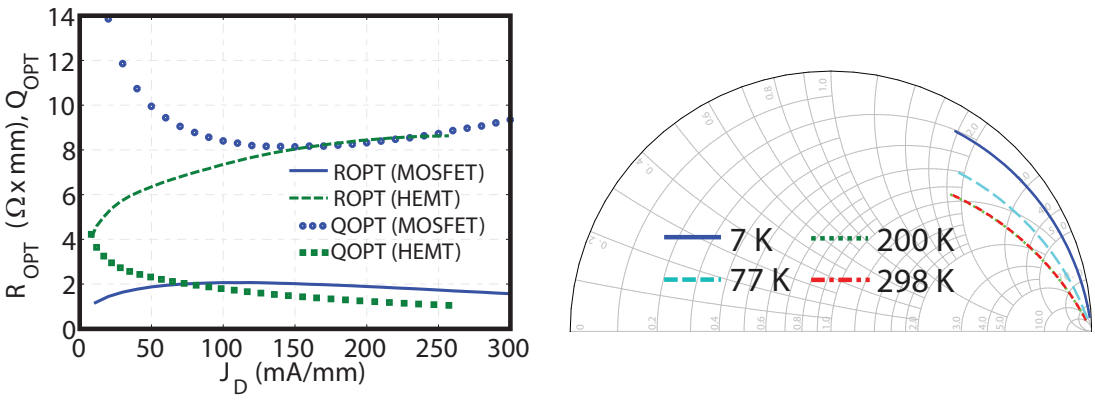
Figure 9.4: (a) T_{MIN} as a function of bias at 10 GHz. (b) T_{MIN} as a function of frequency. J_{D}=80, 90 , 110 and 150 mA/mm at 7, 77, 200 and 298 K for the MOSFET. For the HEMT, J_{D}=42 mA/mm.

The optimum noise input impedance (Z_{OPT}) was investigated next as a function of bias and frequency and the results are shown in Figure 9.5. With cooling, R_{OPT} decreases while X_{OPT} is weakly temperature dependent. While the decrease in R_{OPT} indicates that smaller devices are required at cryogenic temperatures to achieve R_{OPT}=50 Ω, which is beneficial for the DC power consumption, the increase in $Q_{\mathrm{OPT}} = X_{\mathrm{OPT}}/R_{\mathrm{OPT}}$ will complicate the design of broadband matching networks. As shown in Figure 9.5(a), the HEMT has a Q_{OPT} lower than that of the MOSFET at 7 K, which is an important practical advantage when it comes to the design of broadband amplifiers.

Finally, the sensitivity of the noise performance to deviations from the optimum generator impedance was considered. This can be expressed as

$$\frac{T_{\mathrm{E}}\left(\Gamma_{\mathrm{S}}\right) - T_{\mathrm{MIN}}}{T_{\mathrm{MIN}}} = 4N\frac{T_0}{T_{\mathrm{MIN}}}\frac{\left|\Gamma_{\mathrm{S}} - \Gamma_{\mathrm{OPT}}\right|^2}{\Re\left\{\Gamma_{\mathrm{S}}\right\}\Re\left\{\Gamma_{\mathrm{OPT}}\right\}}. \tag{9.1}$$

Hence, $4NT_0/T_{\mathrm{MIN}}$ is the normalized noise penalty, which is also known as Waitr-Pospieszalski noise parameter. For practical devices, this term is expected to be in the range of 1-2 [190]. $4NT_0/T_{\mathrm{MIN}}$ is plotted in Figure 9.6 as a function of bias



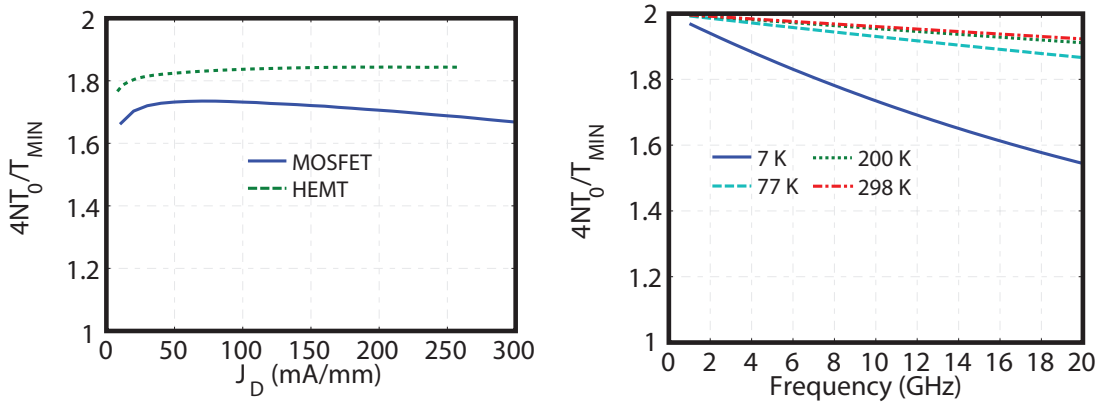
(a) (b)

Figure 9.5: (a) Z_{OPT} as a function of bias at 7 K at 10 GHz. (b) Z_{OPT} in the frequency range of 1-20 GHz. J_{D}=80, 90 , 110 and 150 mA/mm at 7, 77, 200 and 298 K.

and frequency and satisfies the aforementioned condition. NT_0/T_{MIN} appears to be weakly bias dependent and decreases upon cooling. $4NT_0/T_{\mathrm{MIN}}$ is slightly higher for the HEMT than it is for MOSFET at 7 K.

9.1.3 Discrete Amplifier

In order to measure the noise of the 32 nm MOSFET at cryogenic temperatures, a two-stage narrowband discrete amplifier was designed using the models developed in this work. Transistor test structures were diced from the fabricated reticule and placed in pockets that were cut out from a Rogers RT/duroid 6002 printed circuit board (PCB). The PCB was manufactured in-house using a LPKF ProtoLaser [191]. The transistor test structure gate and drain pads were wire-bonded to the traces on the PCB whereas the ground pads were wire-bonded to the chassis. The required inductances were realized with transmission lines on the PCB. Bondable and surface mounted resistors and capacitors were used in the design. The schematic of the amplifier is shown in Figure 9.7. The layout and a close-up photograph of the amplifier are shown in Figure 9.8. The amplifier was tuned at 6 GHz.



(a) (b)

Figure 9.6: (a) Normalized noise penalty as a function of bias at 7 K at 10 GHz. (b) Normalized noise penalty as a function of frequency. J_D=80, 90 , 110 and 150 mA/mm at 7, 77, 200 and 298 K.

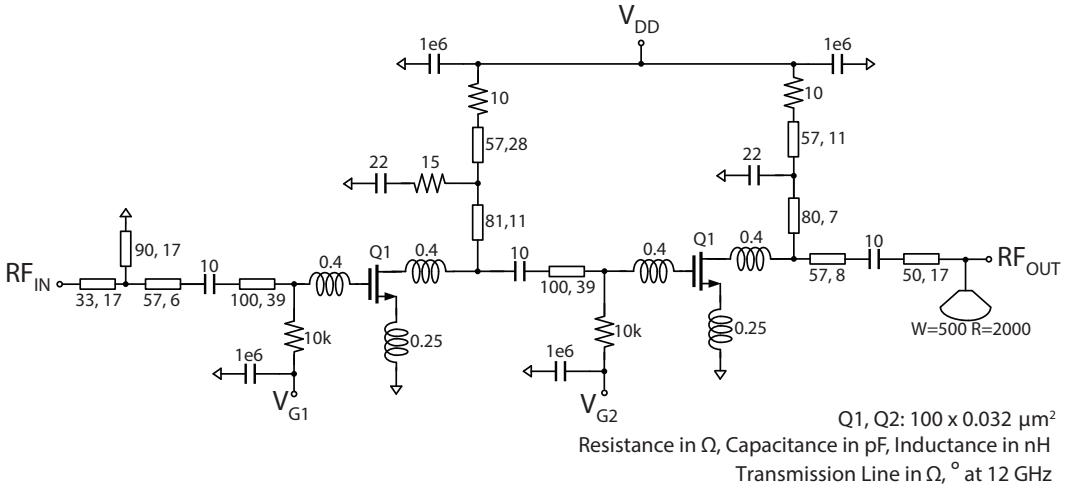
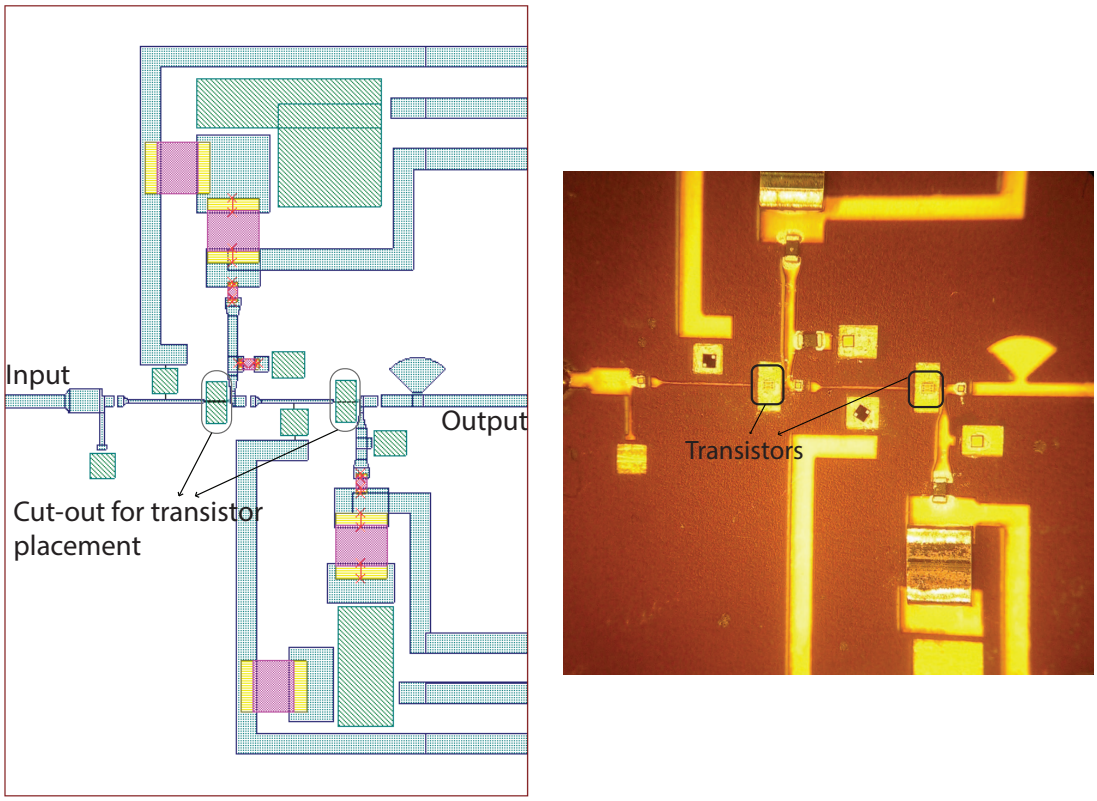
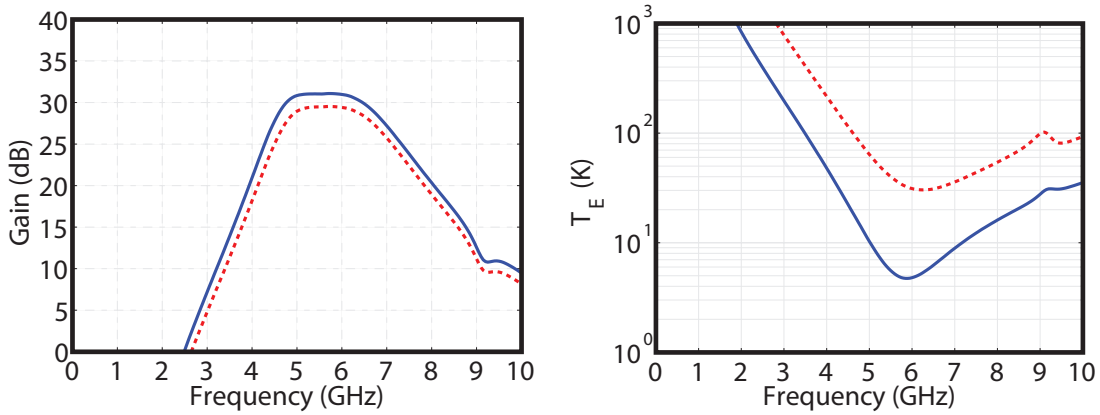


Figure 9.7: Schematic of the two-stage MOSFET amplifier.



(a) (b)

Figure 9.8: (a) Layout and (b) close-up photo of the MOSFET amplifier.



(a) (b)

Figure 9.9: Simulated (a) gain and (b) noise performance of the MOSFET amplifier. Solid lines represent 16 K and dashed lines represent 298 K.

100 μm gate-width transistors were used for both stages and the amplifier consumes 12 mW at 16 K and 24 mW at room temperature. The gain and noise simulation, and measurement results are shown in Figures 9.9 and 9.10, respectively. While an order of magnitude improvement in the noise temperature with cooling is observed, the result at 16 K is a factor of three higher than the expected 5 K noise temperature. During the debugging process, it was observed that this offset is mainly due to the sensitivity of performance to the amplifier assembly. Due to the high Q requirement of the input matching network, it is believed that the required inductance could not be accurately realized. Therefore, a reasonable next step would be to design an integrated circuit amplifier which should provide more precise results.

9.2 0.25 μm InP HBT

Compound semiconductor (CS) HBTs target applications in the terahertz frequency regime [192,193]. Recently, heterogeneous integration of these transistors with silicon CMOS wafers has been reported [194]. In this study, the room temperature noise performance of such HBTs and temperature dependence of their DC and small-signal performance have been investigated. The results are published in [195, 196].

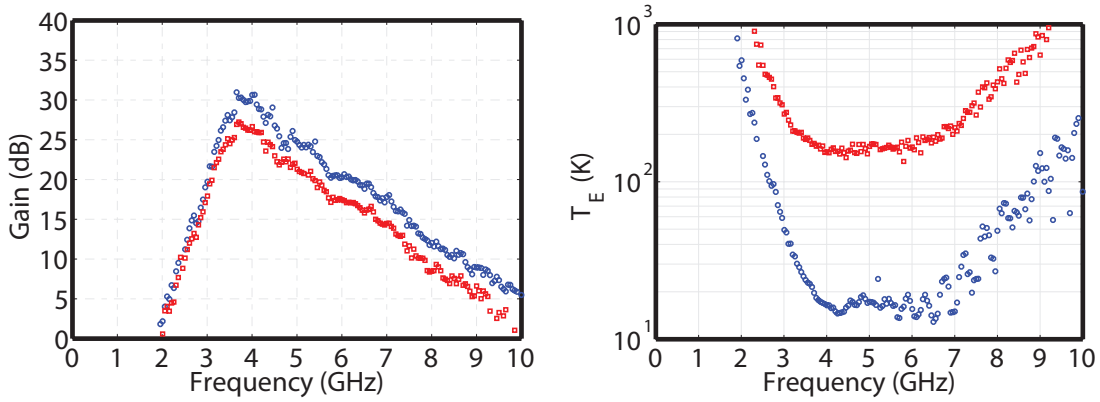


Figure 9.10: Measured (a) gain and (b) noise performance of the MOSFET amplifier. Square data points represent 298 K and circle data points represent 16 K.

9.2.1 Room Temperature Noise Performance

Small-signal noise modeling of the InP HBT was only performed at room temperature. Once the dominant source of noise at room temperature is known, it would be possible to predict the temperature dependence of the noise performance with basic DC and RF measurements at cryogenic temperatures. The small-signal noise model is shown in Figure 9.11. This model takes shot-noise correlation into account [98], which was discussed in Chapter 2.

The noise parameters plotted at different current densities appear in Figure 9.12. The minimum noise temperature is almost frequency independent in the DC-100 GHz range. This is because of the high noise floor created by the substantial base current. Shot noise correlation has a positive effect on T_{MIN} at high frequencies. However, due to the roll-off in the associated gain with frequency, the cascaded noise temperature increases at higher frequencies.

The frequency dependence of R_{OPT} decreases with current density. The quality factor of Z_{OPT} is less than unity in the DC-100 GHz frequency range for current densities greater than 2 mA/μm^2. The noise penalty associated with deviations in the input impedance from Z_{OPT} also decreases with the current density. These fea-

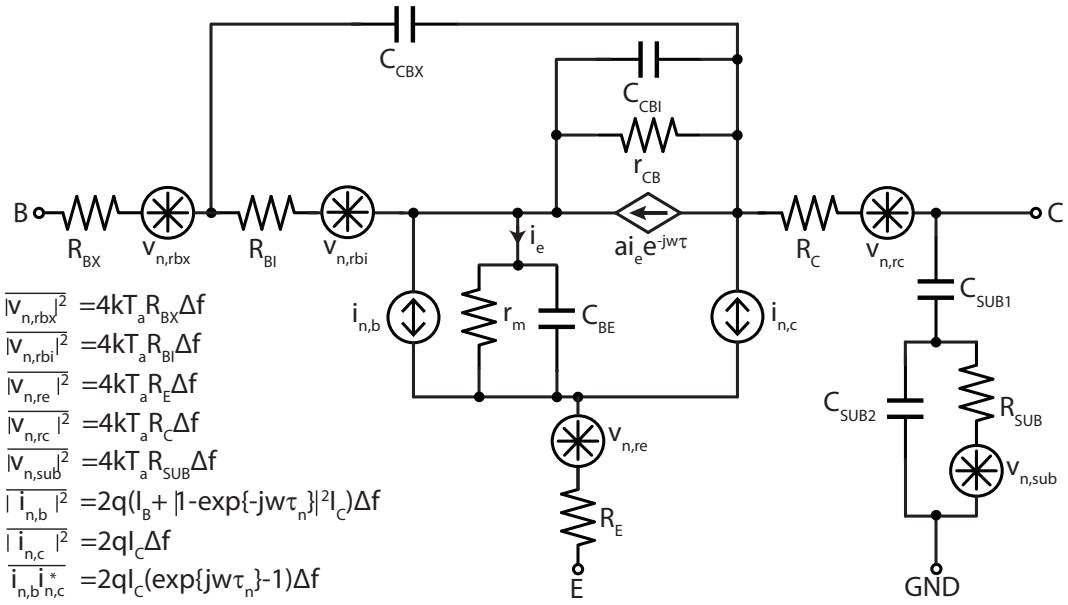


Figure 9.11: InP HBT small-signal noise model. Figure is reproduced from [195] © 2014 IEEE.

tures, along with the weakly frequency dependent T_{MIN}, would enable the design of extremely wideband amplifiers.

9.2.2 Temperature Dependence of the DC and Small-Signal Performance

DC Gummel as well as small-signal f_{t} and f_{max} measurements of the InP HBT was performed as a function of temperature. The DC current gain (β) results appear in Figure 9.13. At room temperature, the peak β is about 25 which is significantly lower than what is typically achieved with SiGe HBTs. This discrepancy is due to the fact that the base of a CS HBT is typically doped at degenerate levels in order to keep the base resistance low enough to maintain the high frequency performance [197]. However, this results in a high recombination current in the base and lowers β. As the recombination current dominates the base current, the temperature dependence of the minority carrier lifetime in the base weakens and eventually has a positive coefficient [198]. Therefore, a weakly temperature dependent β is observed in Figure 9.13.

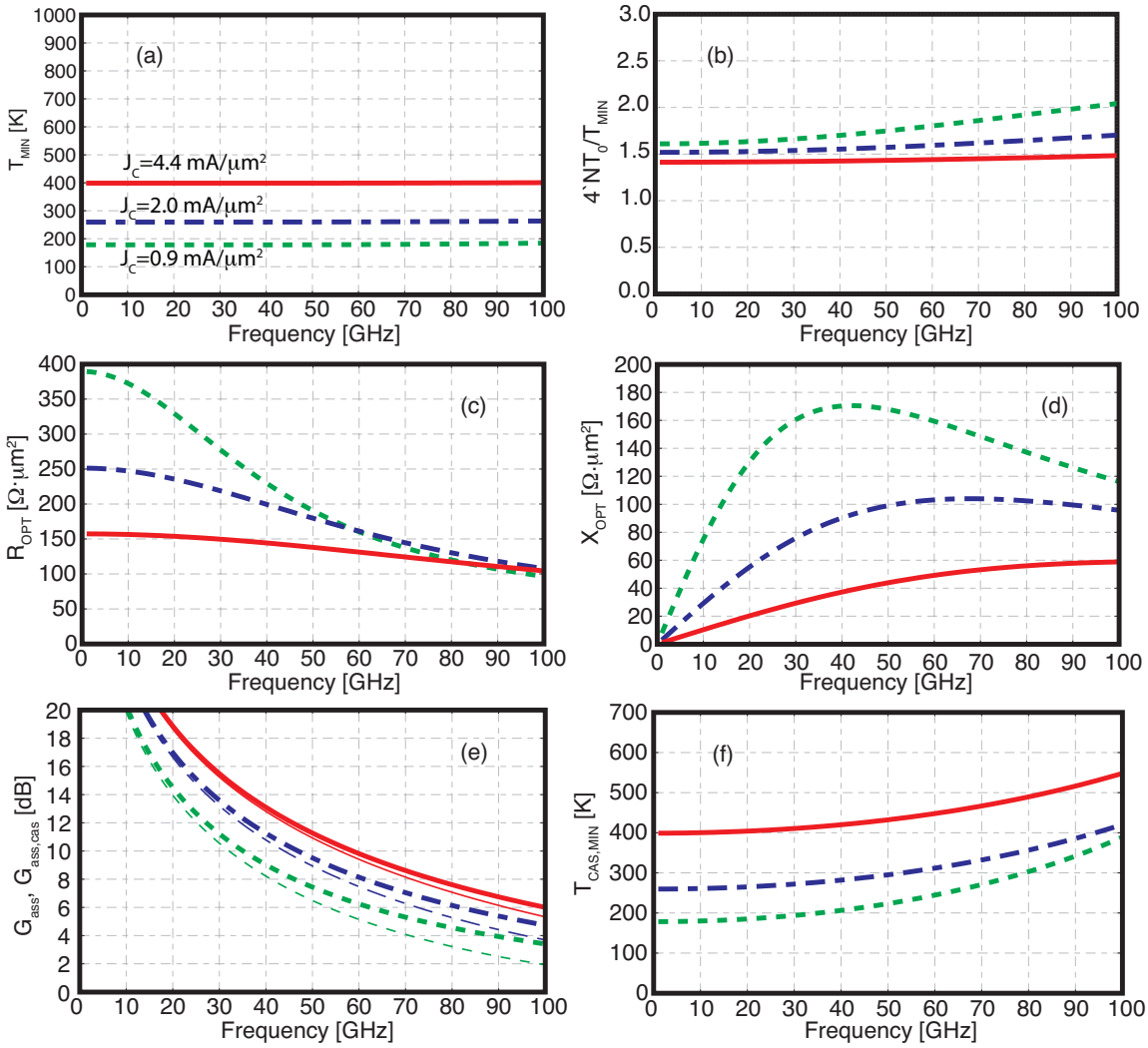


Figure 9.12: Periphery independent noise parameters of the InP HBT at three different bias points with $V_{CB} = 0\ V$. The red (solid), blue (dash-dot), and green (dash) curves represent data taken at 4.4, 2.0, and 0.9 mA/μm^2, respectively. (a) Minimum noise temperature, (b) optimum generator resistance, (c) optimum generator reactance, and (d) the Waitr-Pospieszalski noise parameter, $4NT_0/T_{MIN}$, where $N = R_n \Re\left\{Z_{OPT}^{-1}\right\}$. Also shown are (e) the associated gain for optimum noise match (thin lines) and optimum cascaded noise temperature match (thick lines) and (f) the minimum value of the cascaded noise temperature, $T_{CAS} = T_0 M = T_e/\left(1 - 1/G_{av}\right)$. Plots are reproduced from [195] © 2014 IEEE.

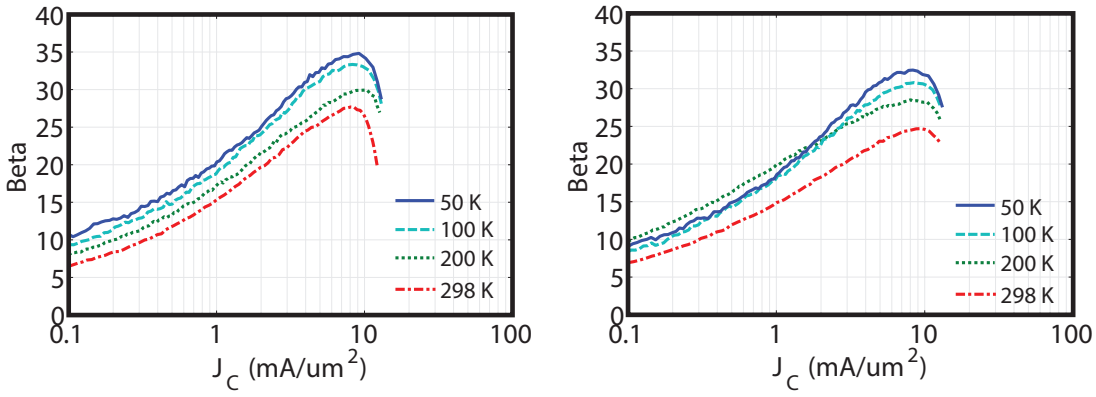
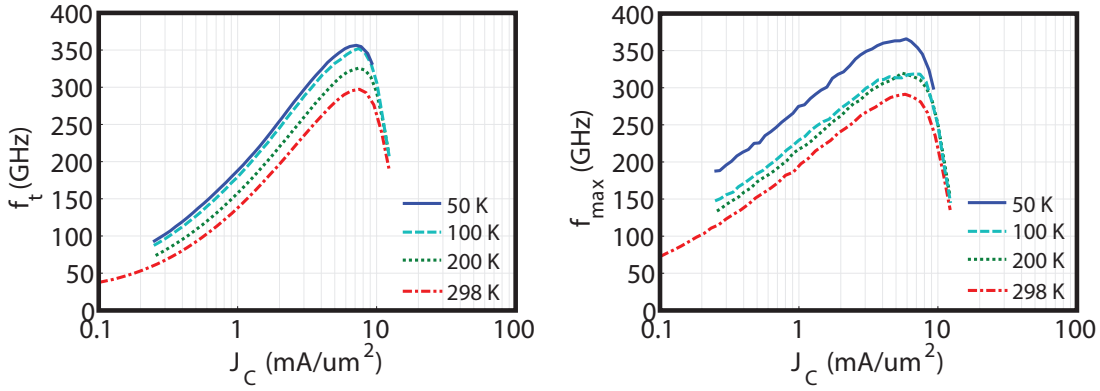


Figure 9.13: Measured DC current gain of the InP HBT. (a) $3\times0.25\mu m^2$ and (b) $4\times0.25\mu m^2$.

Aside from this, thermal runaway issues were observed for certain size devices below 50 K.

The extracted values of f_t and f_{max} for the InP HBT are shown in Figure 9.14. The peak f_t improves with cooling from 298 K to 50 K by a factor of 1.2, which is significantly lower compared to SiGe HBT investigated in this work, whose f_t improvement was a factor of 1.5. This offset potentially indicates that the collector current ideality factor of CS HBTs increases more with cooling than that of SiGe HBTs. The f_{max} improvement of the InP HBT with cooling is a factor of 1.3, which is greater than the f_t improvement. This indicates that the base-collector capacitance and series resistances decrease with cooling. The f_{max} improvement of the SiGe HBT was similar.

Unfortunately, unlike SiGe HBTs, noise performance of CS HBTs with high base doping is not expected to improve dramatically with cooling due to the high base current, which dominates the noise floor.



(a) (b)

Figure 9.14: Extracted (a) f_{t} and (b) f_{max} of the $4{\times}0.25\mu\mathrm{m}^2$ InP HBT.

9.3 Summary

In this chapter, the noise performance of the technologies that are alternative to what is commonly used for low-noise amplification at cryogenic temperatures has been investigated. It was observed that short-channel MOSFETs may offer competitive performance at cryogenic temperatures, although its application to wideband amplifiers is challenging due to the quality factor being required for the input matching network. On the other hand, InP HBTs are very suitable for ultra wideband amplifiers, although their wideband noise performance is fundamentally mediocre due to their low DC current gain.

CHAPTER 10

CONCLUSION

The temperature dependence of the linearity and dynamic range of HEMTs and HBTs has been studied in this dissertation. First, the theoretical device operation and its implications on dynamic range was presented. For HEMTs, it was observed that the velocity-electric field relationship in the channel is very critical for the small-signal, noise, and linearity performance of the device. As HBTs are minority carrier devices, it was observed that they are inherently nonlinear and their intrinsic noise is determined by the DC currents. For both devices, it was demonstrated that the nonlinearity of the charge-control capacitance is tightly coupled to that of the transconductance.

With cooling, the noise performance of both devices improves in conjunction with their transport properties. However, it was demonstrated that the devices also become more nonlinear at lower temperatures. For HEMTs, the relative change in nonlinearity is tightly coupled to the temperature dependence of the electric field at which velocity saturation occurs. For HBTs, the temperature dependence of the nonlinearity is mainly determined by the thermal voltage and ideality factor, the latter of which becomes more bias dependent at cryogenic temperatures. For both devices, the feedback capacitance nonlinearity, which is critical due to the Miller effect, decreases with cooling, as the junction built-in potential is inversely proportional to ambient temperature.

Following the theoretical and experimental analysis of the intrinsic nonlinearities, the intermodulation and dynamic range performance of the devices was investigated as

a function of termination impedances. Despite the fact that both devices become more nonlinear with cooling, the source and load terminations can be optimized to avoid degradation of intermodulation performance. This is mainly achieved through the compensation of increasing transconductance and charge-control capacitance nonlinearities with decreasing feedback capacitance and output conductance nonlinearities. Furthermore, despite the fact the intrinsic nonlinearity of HBTs is greater than that of HEMTs, they can provide better linearity and SFDR performance as optimization of the out-of-band frequency terminations enables intermodulation cancellation.

10.1 Suggestions for Future Work

In this work, it has been shown that the nonlinearity of HEMTs and HBTs can be predicted at cryogenic temperatures, by defining the Taylor series expansion coefficients of strongly bias-dependent elements in a small-signal model. While this study is the first of its kind in terms of the investigation of transistor linearity at such low temperatures, there is room for further research. The results presented here suggest that the significant temperature dependence of the critical physical properties of transistors, such as the electric field that marks the onset of velocity saturation and collector current ideality factor, influence the nonlinear behavior of devices as well as their small-signal and noise characteristics at cryogenic temperatures. Hence, further investigation of device physics at cryogenic temperatures is motivated, which can potentially enable development of physical models that cover the broad temperature range of interest. Furthermore, the key findings of this work can be reflected upon the physical design of transistors to improve their dynamic range performance at cryogenic temperatures.

Within the context of large-signal device characteristics and its modeling, the output bias voltage is critical as it determines the maximum voltage swings a device can provide. Although the importance of this matter becomes somewhat less

important when the weak nonlinearities are of interest, it would still be a good next step to investigate intermodulation and dynamic range performance as a function of output DC voltage. This might reveal additional second-order phenomena critical to the linearity performance and enable further optimization of the linearity-DC power consumption trade-off.

Finally, it would be an important step to utilize intermodulation cancellation techniques for the design of cryogenic low-noise amplifiers. Although this has been conceptually investigated in this work in terms of out-of-band frequency termination optimization, a practical implementation would be quite interesting and prove to be quite useful.

APPENDIX A

NONLINEAR MODEL MEASUREMENT AND EXTRACTION FLOW

A.1 Measurement

Communication with the source-meter and network analyzer is established through the GPIB interface. The Keithley 2600 series source-meter uses a unique programming language for which the details can be found in [199]. The network analyzer commands are processed in the universal SCPI environment [200].

The s-parameter measurements for the nonlinear model extraction starts with determination of the DC bias grid. First, for the nominal output voltage, gate and base voltages corresponding to lowest and highest currents of the grid are determined. This is done through a search function which finds the input voltage required for a certain output current. The MATLAB code is provided below.

```
function [ibase,ic,volt]=setid(k2612,VNA,target,initial,vds)
fprintf(VNA, 'OUTP OFF');
fprintf(k2612,sprintf('smua.source.func = smua.OUTPUT_DCVOLTS'));
fprintf(k2612,sprintf('smub.source.func = smub.OUTPUT_DCVOLTS'));
fprintf(k2612,sprintf('display.smua.measure.func = display.MEASURE_DCAMPS'));
fprintf(k2612,sprintf('display.smub.measure.func = display.MEASURE_DCAMPS'));
fprintf(k2612,sprintf('smua.source.output = smua.OUTPUT_ON'));
fprintf(k2612,sprintf('smub.source.output = smub.OUTPUT_ON'));


volt=initial;
```

```
fprintf(k2612,sprintf('smua.source.levelv=%d',volt));
fprintf(k2612,sprintf('smub.source.levelv=%d',vds));
fprintf(k2612,'ccurrent=smub.measure.i()');
fprintf(k2612,'print(ccurrent)');
actual=str2num(fscanf(k2612));
ic=actual;
fprintf(k2612,'bcurrent=smua.measure.i()');
fprintf(k2612,'print(bcurrent)');
ibase=str2num(fscanf(k2612));

while actual>=(target+target*0.01)| actual<=(target-target*0.01)
    if actual<=(target-target*0.01)
        volt=volt+.0004;
        fprintf(k2612,sprintf('smua.source.levelv=%d',volt));
        fprintf(k2612,'bcurrent=smua.measure.i()');
        fprintf(k2612,'print(bcurrent)');
        ibase=str2num(fscanf(k2612));
        fprintf(k2612,sprintf('smub.source.levelv=%d',vds));
        fprintf(k2612,'ccurrent=smub.measure.i()');
        fprintf(k2612,'print(ccurrent)');
        actual=str2num(fscanf(k2612));
    end
    if actual>=(target+target*0.01)
        volt=volt-.0004;
        fprintf(k2612,sprintf('smua.source.levelv=%d',volt));
        fprintf(k2612,'bcurrent=smua.measure.i()');
        fprintf(k2612,'print(bcurrent)');
```

```
        ibase=str2num(fscanf(k2612));
        fprintf(k2612,sprintf('smub.source.levelv=%d',vds));
        fprintf(k2612,'ccurrent=smub.measure.i()');
        fprintf(k2612,'print(ccurrent)');
        actual=str2num(fscanf(k2612));
    end
    if volt<=-1 | volt>=.65
        break
    end
    ic=actual;
end
    fprintf(VNA, 'OUTP ON');
```

Once the low and high end of the voltage grid are determined, the intermediate voltages are found through linear spacing of the grid points. To complete the grid, points in the output voltage domain are added with a specified step size. The s-parameter measurements are then performed starting from the lowest voltage point in both voltage domains. At each point of the grid, the s-parameter and bias data from both ports are recorded to a data structure. MATLAB code for the individual s-parameter measurement is provided below.

```
function [ hemt ] = activesweepvoltageG( k2612,VNA,Vgs,Vds)
    clrdevice(k2612)
    clrdevice(VNA)
    fprintf(k2612,sprintf('smua.source.func = smua.OUTPUT_DCVOLTS'));
    fprintf(k2612,sprintf('smub.source.func = smub.OUTPUT_DCVOLTS'));
    fprintf(k2612,sprintf('display.smua.measure.func = display.MEASURE_DCAMPS'));
    fprintf(k2612,sprintf('display.smub.measure.func = display.MEASURE_DCAMPS'));
    fprintf(k2612,sprintf('smua.source.levelv=0'));
```

```
fprintf(k2612,sprintf('smub.source.levelv=0'));
fprintf(k2612,sprintf('smua.source.output = smua.OUTPUT_ON'));
fprintf(k2612,sprintf('smub.source.output = smub.OUTPUT_ON'));
for i=1:length(Vgs)
    fprintf(k2612,sprintf('smua.source.levelv=%d',Vgs(i)));
    for k=1:length(Vds)
        fprintf(VNA, 'OUTP OFF');
        fprintf(k2612,sprintf('smub.source.levelv=%d',Vds(k)));
        fprintf(k2612,'bcurrent=smua.measure.i()');
        fprintf(k2612,'print(bcurrent)');
        Ig(k)=str2num(fscanf(k2612));
        clrdevice(k2612);
        fprintf(k2612,'ccurrent=smub.measure.i()');
        fprintf(k2612,'print(ccurrent)');
        Id(k)=str2num(fscanf(k2612));
        clrdevice(k2612);
        pause(0.01);
        fprintf(VNA, 'OUTP ON');
        sp=s2pread(VNA,1);
        clrdevice(k2612);
        hemt.data(:,:,k,i)=sp;
    end
    hemt.Ig(:,i)=Ig;
    hemt.Vds(:,i)=Vds;
    hemt.Id(:,i)=Id;
    hemt.Vgs(i)=Vgs(i);
end
```

```
    fprintf(k2612,sprintf('smub.source.output = smub.OUTPUT_OFF'));
    fprintf(k2612,sprintf('smua.source.output = smua.OUTPUT_OFF'));



end
```

MATLAB code for the complete measurement set is provided below.

```
function [ SS ] = gridhemt( k2612,VNA,min,max,rate,initial,vdsm,vdsstep)

    [igate,id1,v1]=setid(k2612,VNA,min,initial,vdsm);
    initial=0.4;
    [igate,id2,v2]=setid(k2612,VNA,max,initial,vdsm);

    vgs=linspace(v1,v2,rate);

    clrdevice(k2612)
    pause(0.25)

    for k=1:length(vgs);
        for i=1:5
            vds=vdsm+(i-3)*vdsstep;

            [data] = activesweepvoltageG( k2612,VNA,vgs(k),vds);
            SS(k).Vgs(i)=data.Vgs;
            SS(k).Vds(i)=data.Vds;
            SS(k).Ig(i)=data.Ig;
            SS(k).Id(i)=data.Id;
            SS(k).data(:,:,i)=data.data;
```

```
        end
end
```

A.2 Extraction

Once the measurement is complete, the data processing is performed in MATLAB. Grid size for an individual bias point is provided as a function input along with parasitic resistances of the DC paths and frequency range of extraction. At low and high ends of the measurement suite input voltage domain, the grid size is automatically reduced to keep the grid symmetric around the center bias point. The model is extracted for the center points along the input voltage domain. Following the extraction, the complete parameter set is written to a text file in DSCR format [201] and imported into Microwave Office for the harmonic balance simulations. MATLAB code used for the extraction is provided below.

```
function [ALLC]= modelip3v32(hemt,range,R11,R12,R22,coveri,s)

    v=length(hemt.SS(1).Id);
    h=length(hemt.SS);

    ALLC=zeros(h-2,32);

    [Rg,Rs,Rd]= Exres2 (hemt.coldsweep,100:300,6,1);

    for i=1:length(hemt.SS(1).Id)
        for k=1:length(hemt.SS)
            [Gm(i,k),Rds(i,k),Rgs(i,k),Rgd(i,k),Cgs(i,k),Cgd(i,k),Cds(i,k),
            Tao(i,k)]= model2 (hemt,i,k,range,Rg,Rs,Rd,0);
```

```
    end

    Gm(i,:)=smooth(Gm(i,:),s);
    Rds(i,:)=smooth(Rds(i,:),s);
    Rgs(i,:)=smooth(Rgs(i,:),s);
    Rgd(i,:)=smooth(Rgd(i,:),s);
    Cgs(i,:)=smooth(Cgs(i,:),s);
    Cgd(i,:)=smooth(Cgd(i,:),s);
    Cds(i,:)=smooth(Cds(i,:),s);

end

for k=2:(length(hemt.SS)-1);
    i=1;
     if k<=coveri
        covera=k-1;
    end
    if abs(k-length(hemt.SS))<=coveri
        covera=abs(k-length(hemt.SS));
    end
    if k>coveri && abs(k-length(hemt.SS))>coveri
        covera=coveri;
    end

    pts=(covera*2+1)*length(hemt.SS(k).Ig);

    rgs=zeros(1,pts);
```

```
rgd=zeros(1,pts);

cds=zeros(1,pts);

cgs=zeros(1,pts);

cgd=zeros(1,pts);

gds=zeros(1,pts);

rds=zeros(1,pts);

gm=zeros(1,pts);

tao=zeros(1,pts);


Id=zeros(1,pts);

Ig=zeros(1,pts);

Vgs=zeros(1,pts);

Vgd=zeros(1,pts);

Vds=zeros(1,pts);

Vg=zeros(1,pts);

Vd=zeros(1,pts);

Vs=zeros(1,pts);


Vgsdiff=zeros(1,pts);

Vgddiff=zeros(1,pts);

Vdsdiff=zeros(1,pts);


C=ones(pts,9);

Y=ones(1,2*pts);

gsc1=ones(pts,6);

dsc1=ones(pts,6);

gdc=ones(pts,3);
```

```
for j=-covera:covera
    for l=1:length(hemt.SS(k).Ig)

        cds(i)=Cds(l,k+j);
        cgd(i)=Cgd(l,k+j);
        cgs(i)=Cgs(l,k+j);
        rgd(i)=Rgd(l,k+j);
        rgs(i)=Rgs(l,k+j);
        rds(i)=Rds(l,k+j);
        gm(i)=Gm(l,k+j);
        tao(i)=Tao(l,k+j);

        Id(i)=hemt.SS(k+j).Id(l);
        Ig(i)=hemt.SS(k+j).Ig(l);

        Vgs(i)=hemt.SS(k+j).Vgs(l);
        Vds(i)=hemt.SS(k+j).Vds(l);

        Vd(i)=Vds(i)-Id(i)*(R22-R12+Rd);
        Vg(i)=Vgs(i)-Ig(i)*(R11-R12+Rg);
        Vs(i)=(Rs+R12)*(Ig(i)+Id(i));

        Vgs(i)=Vg(i)-Vs(i);
        Vds(i)=Vd(i)-Vs(i);
        Vgd(i)=Vg(i)-Vd(i);
```

```
        gds(i)=1/rds(i);
        i=i+1;
    end
end

 for i=1:pts
    Vgsdiff(i)=Vgs(i)-Vgs(round(pts/2));
    Vdsdiff(i)=Vds(i)-Vds(round(pts/2));
    Vgddiff(i)=Vgd(i)-Vgd(round(pts/2));
 end
 for i=1:pts
      gsc1(i,1)=1;
      gsc1(i,2)=2*Vgsdiff(i);
      gsc1(i,3)=3*Vgsdiff(i)^2;
      gsc1(i,4)=Vdsdiff(i);
      gsc1(i,5)=2*Vdsdiff(i)*Vgsdiff(i);
      gsc1(i,6)=Vdsdiff(i)^2;

      dsc1(i,1)=1;
      dsc1(i,2)=2*Vdsdiff(i);
      dsc1(i,3)=3*Vdsdiff(i)^2;
      dsc1(i,4)=Vgsdiff(i);
      dsc1(i,5)=Vgsdiff(i)^2;
      dsc1(i,6)=2*Vgsdiff(i)*Vdsdiff(i);

      gdc(i,1)=1;
      gdc(i,2)=2*Vgddiff(i);
```

```
        gdc(i,3)=3*Vgddiff(i)^2;

        gsc=gsc1;
        dsc=dsc1;

        C(2*i-1,1)=gsc(i,1);
        C(2*i-1,2)=gsc(i,2);
        C(2*i-1,3)=gsc(i,3);
        C(2*i-1,4:6)=0;
        C(2*i-1,7)=gsc(i,4);
        C(2*i-1,8)=gsc(i,5);
        C(2*i-1,9)=gsc(i,6);

        C(2*i,1:3)=0;
        C(2*i,4)=dsc(i,1);
        C(2*i,5)=dsc(i,2);
        C(2*i,6)=dsc(i,3);
        C(2*i,7)=dsc(i,4);
        C(2*i,8)=dsc(i,5);
        C(2*i,9)=dsc(i,6);

        Y(2*i-1)=gm(i);
        Y(2*i)=gds(i);
    end

    COEFFS=((inv(transpose(C)*C))*transpose(C))*transpose(Y);
```

```
      cgdC=((inv(transpose(gdc)*gdc))*transpose(gdc))*transpose(cgd);
      cgsC=((inv(transpose(gsc)*gsc))*transpose(gsc))*transpose(cgs);
      cdsC=((inv(transpose(dsc)*dsc))*transpose(dsc))*transpose(cds);

      center=round(pts/2);

      ALLC(k-1,:)=[hemt.SS(k).Ig(round(v/2)) hemt.SS(k).Id(round(v/2)) Rg Rs
      Rd rgs(center) rgd(center) transpose(cgsC) transpose(cgdC)
      transpose(COEFFS) tao(center) transpose(cdsC)];
end
writedscr(ALLC)
```

APPENDIX B

STABILITY CRITERIA AND OPTIMIZATION IMPLEMENTATION

During the linearity and dynamic range optimization process, stable termination of the transistors was ensured. Conditional stability is pursued in this work as it imposes a limitation to the selection of impedance values but not to the intrinsic device operation. The requirement for conditional stability is that the real part of total port impedance must be greater than zero [202]. The total port impedance can be expressed as

$$Z_{\mathrm{IN,TOTAL}} = Z_{\mathrm{S}} + Z_{\mathrm{IN}} \tag{B.1}$$

and

$$Z_{\mathrm{OUT,TOTAL}} = Z_{\mathrm{L}} + Z_{\mathrm{OUT}}, \tag{B.2}$$

where Z_{S}, Z_{L} are the value of source and load terminations and Z_{IN}, Z_{OUT} describe the impedance seen towards the corresponding transistor terminal. Z_{IN} and Z_{OUT} can be converted from the reflection coefficients as

$$Z_{\mathrm{IN}} = 50\frac{1+\Gamma_{\mathrm{IN}}}{1-\Gamma_{\mathrm{IN}}} \tag{B.3}$$

and

$$Z_{\mathrm{OUT}} = 50\frac{1+\Gamma_{\mathrm{OUT}}}{1-\Gamma_{\mathrm{OUT}}}. \tag{B.4}$$

These reflection coefficients are obtained from the s-parameters as

$$|\Gamma_{\mathrm{IN}}| = \left| S_{11} + \frac{S_{12}S_{21}\Gamma_{\mathrm{L}}}{1-S_{22}\Gamma_{\mathrm{L}}} \right| \tag{B.5}$$

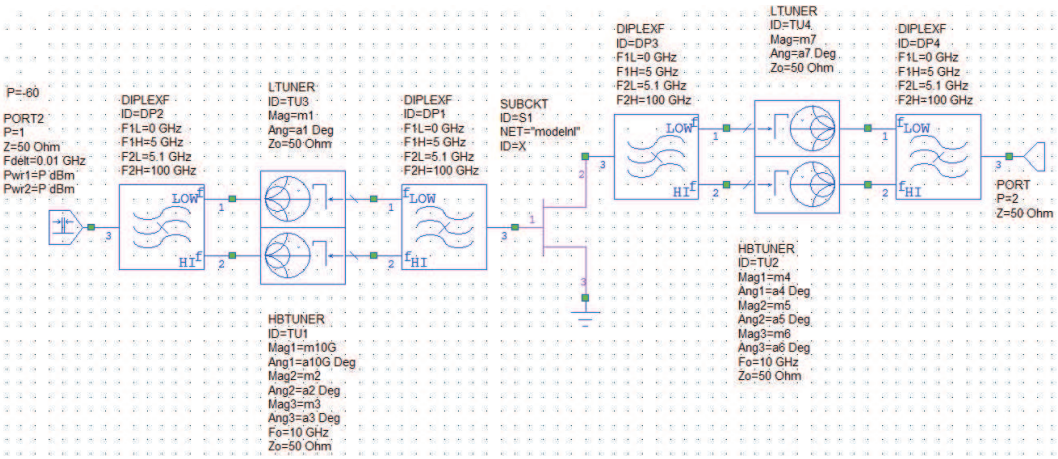


Figure B.1: Screen capture of the Microwave Office narrowband linearity optimization schematic.

and

$$|\Gamma_{\mathrm{OUT}}| = \left| S_{22} + \frac{S_{12}S_{21}\Gamma_{\mathrm{S}}}{1 - S_{11}\Gamma_{\mathrm{S}}} \right|, \tag{B.6}$$

where Γ_{S} and Γ_{L} are the reflections from source and load terminations, respectively. Hence, these conditions are implemented for the fundamental, difference and second harmonic frequency terminations in the optimization process.

Prior to the optimization, Z_{OPT} and s-parameters of the transistor for a given bias point and fundamental frequency are determined. Input fundamental frequency terminations are then set to Z_{OPT} and the s-parameters are used for stability calculations in each iteration. Pointer-Robust and Simplex algorithms are used for the optimization [203]. The narrowband optimization schematic is shown in figure B.1. To enable source/load pull simulations for difference and second harmonic frequency terminations, two different tuners had to be used. These tuners are separated by diplexers. The load pull script is provided within Microwave Office and allows the number of points to be determined. Covering the entire smith chart, 1500 points were typically used for the source/load pull simulations. For the wideband optimization, a third type of tuner is used which enables the impedance definition of arbitrary number of frequencies.

BIBLIOGRAPHY

[1] S. Weinreb, "Low-noise cooled GASFET amplifiers," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 28, pp. 1041–1054, Oct 1980.

[2] M. Pospieszalski and E. Wollack, "Ultra-low-noise, InP field effect transistor radio astronomy receivers: state-of-the-art," in *Microwaves, Radar and Wireless Communications. 2000. MIKON-2000. 13th International Conference on*, vol. 3, pp. 23–32 vol.3, 2000.

[3] C. Groppi, C. Walker, C. Kulesa, D. Golish, J. Kloosterman, S. Weinreb, G. Jones, J. Bardin, H. Mani, T. Kuiper, *et al.*, "Supercam: A 64 pixel heterodyne array receiver for the 350 GHz atmospheric window," in *20th International Symposium on Space Terahertz Technology*, pp. 90–96, 2009.

[4] W. Shan, J. Yang, S. Shi, Q. Yao, Y. Zuo, Z. Lin, S. Chen, X. Zhang, W. Duan, A. Cao, S. Li, Z. Li, J. Liu, and J. Zhong, "Development of superconducting spectroscopic array receiver: A multibeam 2SB SIS receiver for millimeter-wave radio astronomy," *Terahertz Science and Technology, IEEE Transactions on*, vol. PP, pp. 1 –12, november 2012.

[5] C. Walker, C. Groppi, C. d'Aubigny, C. Kulesa, A. Hedden, D. Prober, I. Siddiqi, J. Kooi, G. Chen, and A. Lichtenberger, "Integrated heterodyne array receivers for submillimeter astronomy," in *Astronomical Telescopes and Instrumentation*, pp. 349–354, International Society for Optics and Photonics, 2003.

[6] A. Wallraff, D. I. Schuster, A. Blais, L. Frunzio, R.-S. Huang, J. Majer, S. Kumar, S. M. Girvin, and R. J. Schoelkopf, "Strong coupling of a single photon to a superconducting qubit using circuit quantum electrodynamics," *Nature*, vol. 431, pp. 162–166, Sept. 2004.

[7] C. Rigetti, J. M. Gambetta, S. Poletto, B. L. T. Plourde, J. M. Chow, A. D. Córcoles, J. A. Smolin, S. T. Merkel, J. R. Rozen, G. A. Keefe, M. B. Rothwell, M. B. Ketchen, and M. Steffen, "Superconducting qubit in a waveguide cavity with a coherence time approaching 0.1 ms," *Phys. Rev. B*, vol. 86, p. 100506, Sep 2012.

[8] M. Jung, M. Schroer, K. Petersson, and J. Petta, "Radio frequency charge sensing in InAs nanowire double quantum dots," *Applied Physics Letters*, vol. 100, no. 25, pp. 253508–253508, 2012.

[9] M. Shaw, J. Bueno, P. Day, C. Bradford, and P. Echternach, "Quantum capacitance detector: A pair-breaking radiation detector based on the single cooper-pair box," *Physical Review B*, vol. 79, no. 14, p. 144511, 2009.

[10] D. Gupta, D. E. Kirichenko, V. V. Dotsenko, R. Miller, S. Sarwana, A. Talalaevskii, J. Delmas, R. J. Webber, S. Govorkov, A. F. Kirichenko, I. V. Vernik, and J. Tang, "Modular, multi-function digital-RF receiver systems," *IEEE Transactions on Applied Superconductivity*, vol. 21, pp. 883–890, June 2011.

[11] J. van Heteren, N. Fenzi, T. James, and L. Bourne, "Thin film high temperature superconductor RF coil and cryogenic preamplifier for low field MRI," in *Nuclear Science Symposium and Medical Imaging Conference, 1993., 1993 IEEE Conference Record.*, pp. 1708–1712 vol.3, Oct 1993.

[12] A. Knack, J. Mazierska, and H. Piel, "Superconducting technology: Development of a CDMA base station cryogenic front end receiver," in *Microwave Conference, 2007. APMC 2007. Asia-Pacific*, pp. 1–4, Dec 2007.

[13] M. Pospieszalski, "Extremely low-noise amplification with cryogenic FETs and HFETs: 1970-2004," *Microwave Magazine, IEEE*, vol. 6, pp. 62–75, Sept 2005.

[14] S. S. Iyer, G. L. Patton, S. S. Delage, S. Tiwari, and J. M. C. Stork, "Silicon-germanium base heterojunction bipolar transistors by molecular beam epitaxy," in *1987 International Electron Devices Meeting*, vol. 33, pp. 874–876, 1987.

[15] H. Kroemer, "Theory of a wide-gap emitter for transistors," *Proceedings of the IRE*, vol. 45, pp. 1535–1537, Nov 1957.

[16] J.Cressler and G. Niu, *Silicon Germanium Heterojunction Bipolar Transistors*. Artech House, Mar. 2003.

[17] "DOTSEVEN: Towards 0.7 Terahertz Silicon Germanium Heterojunction Bipolar Technology," consortium, European Consortium. `http://www.dotseven.eu/`.

[18] B. Heinemann et al., "SiGe HBT with ft/fmax of 505 GHz/720 GHz," in *Proc. IEEE IEDM, in press*, Dec 2016.

[19] W. Deal, X. B. Mei, K. M. K. H. Leong, V. Radisic, S. Sarkozy, and R. Lai, "THz monolithic integrated circuits using InP high electron mobility transistors," *IEEE Transactions on Terahertz Science and Technology*, vol. 1, pp. 25–32, Sept 2011.

[20] A. Leuther, A. Tessmann, P. Doria, M. Ohlrogge, M. Seelmann-Eggebert, H. Maler, M. Schlechtweg, and O. Ambacher, "20 nm metamorphic HEMT technology for terahertz monolithic integrated circuits," in *European Microwave Integrated Circuit Conference (EuMIC), 2014 9th*, pp. 84–87, Oct 2014.

[21] J. Bardin and S. Weinreb, "Experimental cryogenic modeling and noise of SiGe HBTs," in *Microwave Symposium Digest, 2008 IEEE MTT-S International*, pp. 459–462, June 2008.

[22] S. Weinreb, J. C. Bardin, and H. Mani, "Design of cryogenic SiGe low-noise amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 55, pp. 2306–2312, Nov 2007.

[23] J. Bardin and S. Weinreb, "A 0.1–5 GHz cryogenic SiGe MMIC LNA," *Microwave and Wireless Components Letters, IEEE*, vol. 19, pp. 407 –409, june 2009.

[24] S. Weinreb, J. Bardin, H. Mani, and G. Jones, "Matched wideband low-noise amplifiers for radio astronomy," *Review of Scientific Instruments*, vol. 80, pp. 044702 –044702–5, apr 2009.

[25] J. Schleeh, H. Rodilla, N. Wadefalk, P. Nilsson, and J. Grahn, "Cryogenic noise performance of InGaAs/InAlAs HEMTs grown on InP and GaAs substrate," *Solid-State Electronics*, vol. 91, pp. 74 – 77, 2014.

[26] J. C. Bardin, *Silicon-germanium heterojunction bipolar transistors for extremely low-noise applications*. Phd thesis, California Institute of Technology, 2009.

[27] J. Schleeh, G. Alestig, J. Halonen, A. Malmros, B. Nilsson, P. Nilsson, J. Starski, N. Wadefalk, H. Zirath, and J. Grahn, "Ultralow-power cryogenic InP HEMT with minimum noise temperature of 1 K at 6 GHz," *Electron Device Letters, IEEE*, vol. 33, pp. 664–666, May 2012.

[28] J. Pandian, L. Baker, G. Cortes, P. Goldsmith, A. Deshpande, R. Ganesan, J. Hagen, L. Locke, N. Wadefalk, and S. Weinreb, "Low-noise 6-8 GHz receiver," *Microwave Magazine, IEEE*, vol. 7, pp. 74–84, Dec 2006.

[29] B. A. Abelan, M. Seelmann-Eggebert, D. Bruch, A. Leuther, H. Massler, B. Baldischweiler, M. Schlechtweg, J. D. Gallego-Puyol, I. Lopez-Fernandez, C. Diez-Gonzalez, I. Malo-Gomez, E. Villa, and E. Artal, "4-12- and 25-34-GHz cryogenic mHEMT MMIC low-noise amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, pp. 4080–4088, Dec 2012.

[30] T. Thrivikraman, J. Yuan, J. Bardin, H. Mani, S. Phillips, W.-M. L. Kuo, J. Cressler, and S. Weinreb, "SiGe HBT x-band LNAs for ultra-low-noise cryogenic receivers," *Microwave and Wireless Components Letters, IEEE*, vol. 18, pp. 476–478, July 2008.

[31] M. Pospieszalski, W. Lakatosh, L. Nguyen, M. Lui, T. Lin, M. Le, M. Thompson, and M. Delaney, "Q- and e-band cryogenically-coolable amplifiers using AlInAs/GaInAs/InP HEMTs," in *Microwave Symposium Digest, 1995., IEEE MTT-S International*, pp. 1121–1124 vol.3, May 1995.

[32] M. Varonen, R. Reeves, P. Kangaslahti, L. Samoska, A. Akgiray, K. Cleary, R. Gawande, A. Fung, T. Gaier, S. Weinreb, A. Readhead, C. Lawrence, S. Sarkozy, and R. Lai, "A 75-116 GHz LNA with 23-K noise temperature at 108 GHz," in *Microwave Symposium Digest (IMS), 2013 IEEE MTT-S International*, pp. 1–3, June 2013.

[33] W. T. Wong, P. Ravindran, S.-W. Chang, and J. C. Bardin, "A SiGe Ka-band cryogenic low-noise amplifier," in *2016 IEEE MTT-S International Microwave Symposium (IMS)*, pp. 1–3, May 2016.

[34] S. Montazeri, P. K. Grimes, C. Y. E. Tong, and J. C. Bardin, "A wide-band high-gain compact sis receiver utilizing a 300-uw SiGe IF LNA," *IEEE Transactions on Applied Superconductivity, In Press*.

[35] N. Jarosik, "Measurements of the low-frequency-gain fluctuations of a 30-GHz high-electron-mobility-transistor cryogenic amplifier," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 44, pp. 193–197, Feb 1996.

[36] E. W. Bryerton, M. Morgan, and M. W. Pospieszalski, "Ultra low noise cryogenic amplifiers for radio astronomy," in *2013 IEEE Radio and Wireless Symposium*, pp. 358–360, Jan 2013.

[37] S. Montazeri, W. T. Wong, A. H. Coskun, and J. C. Bardin, "Ultra-low-power cryogenic SiGe low-noise amplifiers: Theory and demonstration," *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, pp. 178–187, Jan 2016.

[38] C. k. Luo, P. S. Gudem, and J. F. Buckwalter, "A 0.4-6 GHz 17 dBm B1dB 36 dBm IIP3 channel-selecting low-noise amplifier for SAW-less 3G/4G FDD diversity receivers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, pp. 1110–1121, April 2016.

[39] S. Hameed, N. Sinha, M. Rachid, and S. Pamarti, "A programmable receiver front-end achieving ¿17dBm IIP3 at ¡1.25 x BW frequency offset," in *2016 IEEE International Solid-State Circuits Conference (ISSCC)*, pp. 446–447, Jan 2016.

[40] Z. Ru, E. A. M. Klumperink, G. J. M. Wienk, and B. Nauta, "A software-defined radio receiver architecture robust to out-of-band interference," in *2009 IEEE International Solid-State Circuits Conference - Digest of Technical Papers*, pp. 230–231,231a, 2009.

[41] M. Kitsunezuka, H. Kodama, N. Oshima, K. Kunihiro, T. Maeda, and M. Fukaishi, "A 30 MHz-2.4 GHz CMOS receiver with integrated RF filter and dynamic-range-scalable energy detector for cognitive radio systems," *IEEE Journal of Solid-State Circuits*, vol. 47, pp. 1084–1093, May 2012.

[42] M.-W. Hsieh, Y.-M. Hsin, K.-H. Liang, Y.-J. Chan, and D. Tang, "RF power characteristics of SiGe HBTs at cryogenic temperatures," *Electron Devices, IEEE Transactions on*, vol. 53, pp. 1452–1458, June 2006.

[43] A. Cardoso, A. Omprakash, P. Chakraborty, N. Karaulac, D. Fleischhauer, A. Ildefonso, S. Zeinolabedinzadeh, M. Oakley, T. Bantu, N. Lourenco, and J. Cressler, "On the cryogenic RF linearity of SiGe HBTs in a fourth-generation 90-nm SiGe BiCMOS technology," *Electron Devices, IEEE Transactions on*, vol. 62, pp. 1127–1135, April 2015.

[44] A. H.Akigray, *New Technologies Driving Decade-Bandwidth Radio Astronomy: Quad-Ridged Flared Horn and Compound-Semiconductor LNAs*. Phd thesis, California Institute of Technology, 2013.

[45] "Radiometers," course note, National Radio Astronomy Observatory. `http://www.cv.nrao.edu/course/astr534/Radiometers.html`.

[46] S. J. Wijnholds and A. J. van der Veen, "Fundamental imaging limits of radio telescope arrays," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, pp. 613–623, Oct 2008.

[47] P. K. Day, H. G. LeDuc, B. A. Mazin, A. Vayonakis, and J. Zmuidzinas, "A broadband superconducting detector suitable for use in large arrays," *Nature*, vol. 425, pp. 817–821, Sept. 2003.

[48] R. Duan, S. McHugh, B. Serfass, B. A. Mazin, A. Merrill, S. R. Golwala, T. P. Downes, N. G. Czakon, P. K. Day, J. Gao, J. Glenn, M. I. Hollister, H. G. Leduc, P. R. Maloney, O. Noroozian, H. T. Nguyen, J. Sayers, J. A. Schlaerth, S. Siegel, J. E. Vaillancourt, A. Vayonakis, P. R. Wilson, and J. Zmuidzinas, "An open-source readout for MKIDs," 2010.

[49] "4 - 8 GHz Low-Noise Amplifier ," datasheet, Low Noise Factory. `http://www.lownoisefactory.com/products/cryogenic/lnc-4-8-ghz/`.

[50] "16 - 28 GHz Low-Noise Amplifier ," datasheet, Low Noise Factory. `http://www.lownoisefactory.com/products/cryogenic/16-28-ghz-cryogenic-waveguide/`.

[51] "0.5 - 3 GHz Low-Noise Amplifier," datasheet, Arizona State University. `http://thz.asu.edu/sites/default/files/ASU_3GHz_CRYOLNA_DataSheet_092015_0.pdf`.

[52] S. Yngvesson., *Microwave Semiconductor Devices*. Kluwer Academic Publishers, 1991.

[53] P. Roblin and H. Rohdin., *High-Speed Heterostructure Devices*. Cambridge University Press, New York, 2002.

[54] Y. Ando and T. Itoh, "Analysis of charge control in pseudomorphic two-dimensional electron gas field-effect transistors," *IEEE Transactions on Electron Devices*, vol. 35, pp. 2295–2301, Dec 1988.

[55] K. Lee, M. S. Shur, T. J. Drummond, and H. Morkoc, "Current - voltage and capacitance - voltage characteristics of modulation-doped field-effect transistors," *IEEE Transactions on Electron Devices*, vol. 30, pp. 207–212, Mar 1983.

[56] D. Delagebeaudeuf and N. T. Linh, "Metal-(n) AlGaAs-GaAs two-dimensional electron gas FET," *IEEE Transactions on Electron Devices*, vol. 29, pp. 955–960, Jun 1982.

[57] F. Stern and W. E. Howard, "Properties of semiconductor surface inversion layers in the electric quantum limit," *Phys. Rev.*, vol. 163, pp. 816–835, Nov 1967.

[58] T. J. Drummond, W. T. Masselink, and H. Morkoc, "Modulation-doped GaAs/(Al,Ga)As heterojunction field-effect transistors: MODFETs," *Proceedings of the IEEE*, vol. 74, pp. 773–822, June 1986.

[59] K. Lee, M. Shur, T. J. Drummond, and H. Morkoc, "Electron density of the twodimensional electron gas in modulation doped layers," *Journal of Applied Physics*, vol. 54, no. 4, pp. 2093–2096, 1983.

[60] J. Wood and C. G. Morton, "An analysis of the effective position of the two-dimensional electron gas in the channel of MODFET epitaxial layer structures," *IEEE Transactions on Electron Devices*, vol. 45, pp. 1622–1624, Jul 1998.

[61] K. Lee, M. S. Shur, T. J. Drummond, and H. Morkoc, "Current - voltage and capacitance - voltage characteristics of modulation-doped field-effect transistors," *IEEE Transactions on Electron Devices*, vol. 30, pp. 207–212, Mar 1983.

[62] C. Canali, G. Majni, R. Minder, and G. Ottaviani, "Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature," *IEEE Transactions on Electron Devices*, vol. 22, pp. 1045–1047, Nov 1975.

[63] S. L. Teitel and J. W. Wilkins, "Ballistic transport and velocity overshoot in semiconductors: Part i - uniform field effects," *IEEE Transactions on Electron Devices*, vol. 30, pp. 150–153, Feb 1983.

[64] M. Tomizawa, K. Yokoyama, and A. Yoshii, "Hot-electron velocity characteristics at AlGaAs/GaAs heterostructures," *IEEE Electron Device Letters*, vol. 5, pp. 464–465, Nov 1984.

[65] W. Masselink, N. Braslau, D. LaTulipe, W. Wang, and S. Wright, "Electron velocity at high electric fields in AlGaAs/GaAs modulation-doped heterostructures," *Solid-State Electronics*, vol. 31, no. 3, pp. 337 – 340, 1988.

[66] G. Salmer, J. Zimmermann, and R. Fauquembergue, "Modeling of MODFETs," *IEEE Transactions on Microwave Theory and Techniques*, vol. 36, pp. 1124–1140, Jul 1988.

[67] H. Rodilla, J. Schleeh, P. . Nilsson, N. Wadefalk, J. Mateos, and J. Grahn, "Cryogenic performance of low-noise InP HEMTs: A monte carlo study," *IEEE Transactions on Electron Devices*, vol. 60, pp. 1625–1631, May 2013.

[68] S. Adachi., *Physical Properties of III-V Semiconductor Compunds*. John Wiley & Sons, May. 1992.

[69] J. L. Thobel, L. Baudry, A. Cappy, P. Bourel, and R. Fauquembergue, "Electron transport properties of strained InxGa1-xAs," *Applied Physics Letters*, vol. 56, no. 4, pp. 346–348, 1990.

[70] U. K. Mishra, A. S. Brown, M. J. Delaney, P. T. Greiling, and C. F. Krumm, "AlInAs-GaInAs HEMT for microwave and millimeter-wave applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 37, pp. 1279–1285, Sep 1989.

[71] H. Rohdin, "Reverse modeling of e/d logic submicrometer MODFETs and prediction of maximum extrinsic MODFET current gain cutoff frequency," *IEEE Transactions on Electron Devices*, vol. 37, pp. 920–934, Apr 1990.

[72] P. Roblin, H. Rohdin, C. J. Hung, and S. W. Chiu, "Capacitance-voltage analysis and current modeling of pulse-doped MODFETs," *IEEE Transactions on Electron Devices*, vol. 36, pp. 2394–2404, Nov 1989.

[73] K. Park and K. D. Kwack, "A model for the current - voltage characteristics of MODFET's," *IEEE Transactions on Electron Devices*, vol. 33, pp. 673–676, May 1986.

[74] M. J. Moloney, F. Ponse, and H. Morkoc, "Gate capacitance-voltage characteristic of MODFET's: Its effect on transconductance," *IEEE Transactions on Electron Devices*, vol. 32, pp. 1675–1684, Sep 1985.

[75] E. H. Rhoderick, "Metal-semiconductor contacts," *IEE Proceedings I - Solid-State and Electron Devices*, vol. 129, pp. 1–, February 1982.

[76] Y. Tsividis, *Operation and Modeling of the MOS Transistor*. Oxford University Press, New York, 1999.

[77] S. M. Sze, *Physics of Semiconductor Devices*. John Wiley & Sons, May 1981.

[78] P. Heymann and H. Prinzler, "Improved noise model for MESFETs and HEMTs in lower gigahertz frequency range," *Electronics Letters*, vol. 28, pp. 611–612, March 1992.

[79] M. W. Pospieszalski, "On noise properties of transistors and amplifiers a critical review," in *Microwaves, Radar, and Wireless Communication (MIKON), 2014 20th International Conference on*, pp. 1–5, June 2014.

[80] J. P. Nougier, "Fluctuations and noise of hot carriers in semiconductor materials and devices," *IEEE Transactions on Electron Devices*, vol. 41, pp. 2034–2049, Nov 1994.

[81] M. W. Pospieszalski, "Modeling of noise parameters of MESFETs and MODFETs and their frequency and temperature dependence," *IEEE Transactions on Microwave Theory and Techniques*, vol. 37, pp. 1340–1350, Sep 1989.

[82] H. Rodilla, J. Schleeh, P.-A. Nilsson, and J. Grahn, "Cryogenic kink effect in InP pHEMTs: A pulsed measurements study," *Electron Devices, IEEE Transactions on*, vol. 62, pp. 532–537, Feb 2015.

[83] S. A. Maas, *Nonlinear Microwave and RF Circuits*. Artech House Inc., Norwood MA, 2003.

[84] J. S. Yuan, *SiGe, GaAs and InP Heterojunction Bipolar Transistors*. John Wiley & Sons, 1999.

[85] B. Streetman and S. Banerjee, *Solid State Electronic Devices, 5th edition*. Prentice Hall, 2000.

[86] D. L. Harame, D. C. Ahlgren, D. D. Coolbaugh, J. S. Dunn, G. G. Freeman, J. D. Gillis, R. A. Groves, G. N. Hendersen, R. A. Johnson, A. J. Joseph, S. Subbanna, A. M. Victor, K. M. Watson, C. S. Webster, and P. J. Zampardi, "Current status and future trends of SiGe BiCMOS technology," *IEEE Transactions on Electron Devices*, vol. 48, pp. 2575–2594, Nov 2001.

[87] H. Kroemer, "Quasi-electric and quasi-magnetic fields in nonuniform semiconductors," *RCA Review*, vol. 18, pp. 332–342, 1957.

[88] S. S. Iyer, G. L. Patton, J. M. C. Stork, B. S. Meyerson, and D. L. Harame, "Heterojunction bipolar transistors using Si-Ge alloys," *IEEE Transactions on Electron Devices*, vol. 36, pp. 2043–2064, Oct 1989.

[89] M. Chrzanowska-Jeske and R. C. Jaeger, "BILOW-simulation of low-temperature bipolar device behavior," *IEEE Transactions on Electron Devices*, vol. 36, pp. 1475–1488, Aug 1989.

[90] W. M. Webster, "On the variation of junction-transistor current-amplification factor with emitter current," *Proceedings of the IRE*, vol. 42, pp. 914–920, June 1954.

[91] P. Ashburn, *SiGe Heterojunction Bipolar Transistors*. Hoboken, NJ: Jon Wiley and Sons, 2003.

[92] C. T. Kirk, "A theory of transistor cutoff frequency (ft) falloff at high current densities," *IRE Transactions on Electron Devices*, vol. 9, pp. 164–174, March 1962.

[93] O. Hansen, "Diffusion in a short base," *Solid-State Electronics*, vol. 37, no. 9, pp. 1663 – 1669, 1994.

[94] D. Richey, A. Joseph, J. Cressler, and R. Jaeger, "Evidence for non-equilibrium base transport in Si and SiGe bipolar transistors at cryogenic temperatures," *Solid-State Electronics*, vol. 39, no. 6, pp. 785 – 789, 1996.

[95] J. R. Pierce, "Physical sources of noise," *Proceedings of the IRE*, vol. 44, pp. 601–608, May 1956.

[96] A. van der Ziel, "Theory of shot noise in junction diodes and junction transistors," *Proceedings of the IRE*, vol. 43, pp. 1639–1646, Nov 1955.

[97] G. Niu, J. D. Cressler, S. Zhang, W. E. Ansley, C. S. Webster, and D. L. Harame, "A unified approach to RF and microwave noise parameter modeling in bipolar transistors," *IEEE Transactions on Electron Devices*, vol. 48, pp. 2568–2574, Nov 2001.

[98] M. Rudolph, F. Korndorfer, P. Heymann, and W. Heinrich, "Compact large-signal shot-noise model for HBTs," *IEEE Transactions on Microwave Theory and Techniques*, vol. 56, pp. 7–14, Jan 2008.

[99] E. Ramirez-Garcia, F. Aniel, M. A. Enciso-Aguilar, and N. Zerounian, "Intrinsic transit times and noise transport time study of Si/SiGe:C heterejunction bipolar transistors," in *2012 7th European Microwave Integrated Circuit Conference*, pp. 175–178, Oct 2012.

[100] Y. Varshni, "Temperature dependence of the energy gap in semiconductors," *Physica*, vol. 34, no. 1, pp. 149 – 154, 1967.

[101] V. Palankovski, *Simulation of Heterojunction Bipolar Transistors*. Phd thesis, Technischen Universitt Wien, 2000.

[102] F. Stern and S. Das Sarma, "Electron energy levels in GaAs-ga$_{1-x}$al$_x$As heterojunctions," *Phys. Rev. B*, vol. 30, pp. 840–848, Jul 1984.

[103] J. Schleeh, *Cryogenic Ultra-Low Noise InP High Electron Mobility Transistors*. Phd thesis, Chalmers University of Technology, 2013.

[104] W. Bludau, A. Onton, and W. Heinke, "Temperature dependence of the band gap of silicon," *Journal of Applied Physics*, vol. 45, no. 4, pp. 1846–1848, 1974.

[105] W.-P. Hong and P. K. Bhattacharya, "High-field transport in InGaAs/InAlAs modulation-doped heterostructures," *IEEE Transactions on Electron Devices*, vol. 34, pp. 1491–1495, Jul 1987.

[106] C. Jacoboni, C. Canali, G. Ottaviani, and A. A. Quaranta, "A review of some charge transport properties of silicon," *Solid-State Electronics*, vol. 20, no. 2, pp. 77 – 89, 1977.

[107] M. V. Fischetti, "Monte carlo simulation of transport in technologically significant semiconductors of the diamond and zinc-blende structures. i. homogeneous transport," *IEEE Transactions on Electron Devices*, vol. 38, pp. 634–649, Mar 1991.

[108] C. C. Hu, *Modern Semiconductor Devices for Integrated Circuits*. Pearson, 2000.

[109] P. P. Altermatt, A. Schenk, B. Schmithsen, and G. Heiser, "A simulation model for the density of states and for incomplete ionization in crystalline silicon. ii. investigation of Si:As and Si:B and usage in device simulation," *Journal of Applied Physics*, vol. 100, no. 11, p. 113715, 2006.

[110] "The Cryogenic DC Behavior of Cryo3/AZ1 InP 0.1-by-80-Micrometer-Gate High Electron Mobility Transistor Devices," progress report, NASA. `http://ipnpr.jpl.nasa.gov/progress_report/42-169/169C.pdf`.

[111] M. H. Somerville, J. A. del Alamo, and W. Hoke, "Direct correlation between impact ionization and the kink effect in InAlAs/InGaAs HEMTs," *IEEE Electron Device Letters*, vol. 17, pp. 473–475, Oct 1996.

[112] W. Kruppa and J. B. Boos, "Examination of the kink effect in InAlAs/InGaAs/InP HEMTs using sinusoidal and transient excitation," *IEEE Transactions on Electron Devices*, vol. 42, pp. 1717–1723, Oct 1995.

[113] J. Frey, "Effects of intervalley scattering on noise in GaAs and InP field-effect transistors," *IEEE Transactions on Electron Devices*, vol. 23, pp. 1298–1303, Dec 1976.

[114] R. Grundbacher, R. Lai, M. Barsky, R. Tsai, T. Gaier, S. Weinreb, D. Dawson, J. J. Bautista, J. F. Davis, N. Erickson, T. Block, and A. Oki, "0.1 mum InP HEMT devices and MMICs for cryogenic low noise amplifiers from X-band to W-band," in *Indium Phosphide and Related Materials Conference, 2002. IPRM. 14th*, pp. 455–458, 2002.

[115] "Process Overview," datasheet, Northrop Grumman. `http://www.northropgrumman.com/BusinessVentures/Microelectronics/Pages/FoundryServices.aspx`.

[116] P. . Nilsson, H. Rodilla, J. Schleeh, N. Wadefalk, and J. Grahn, "Influence of gate-channel distance in low-noise InP HEMTs," in *Indium Phosphide and Related Materials (IPRM), 2013 International Conference on*, pp. 1–2, May 2013.

[117] "Recent NGST HEMT Device & MMIC Development," MMIC array receivers and spectrographs workshop, Northrop Grumman. `http://www.kiss.caltech.edu/workshops/mmic2008/presentations/lai.pdf`.

[118] Y. C. Chou, R. Grundbacher, D. Leung, R. Lai, Q. Kan, D. Eng, P. H. Liu, T. Block, and A. Oki, "Degradation mechanism and reliability improvement of InGaAs/InAlAs/InP HEMTs using new gate metal electrode technology," in *International Conference on Indium Phosphide and Related Materials, 2005*, pp. 223–226, May 2005.

[119] Y. C. Chou, D. Leung, R. Lai, R. Grundbacher, M. Barsky, Q. Kan, R. Tsai, D. Eng, M. Wojtowicz, T. Block, P. H. Liu, S. Olson, A. Oki, and D. C. Streit, "0.1 mum InGaAs/InAlAs/InP HEMT MMICs - a flight qualified technology," in *Gallium Arsenide Integrated Circuit (GaAs IC) Symposium, 2002. 24th Annual Technical Digest*, pp. 77–80, Oct 2002.

[120] P. M. Smith, "Status of InP HEMT technology for microwave receiver applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 44, pp. 2328–2333, Dec 1996.

[121] B. A. Orner, Q. Z. Liu, B. Rainey, A. Stricker, P. Geiss, P. Gray, M. Zierak, M. Gordon, D. Collins, V. Ramachandran, W. Hodge, C. Willets, A. Joseph, J. Dunn, J. S. Rieh, S. J. Jeng, E. Eld, G. Freeman, and D. Ahlgren, "A 0.13 mum BiCMOS technology featuring a 200/280 GHz (ft/fmax) SiGe HBT," in *Bipolar/BiCMOS Circuits and Technology Meeting, 2003. Proceedings of the*, pp. 203–206, Sept 2003.

[122] P. Chevalier, T. F. Meister, B. Heinemann, S. V. Huylenbroeck, W. Liebl, A. Fox, A. Sibaja-Hernandez, and A. Chantre, "Towards THz SiGe HBTs," in *2011 IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 57–65, Oct 2011.

[123] G. Freeman, D. R. Greenberg, K. Walter, and S. Subbanna, "SiGe HBT performance improvements from lateral scaling," in *Solid-State Device Research Conference, 1999. Proceeding of the 29th European*, vol. 1, pp. 724–727, Sept 1999.

[124] "CRX-4K Cryogenic Probe Station," datasheet, Lakeshore. `http://www.lakeshore.com/products/Cryogenic-Probe-Stations/Model-CRX-4K-Cryogenic-Probe-Station/pages/Overview.aspx`.

[125] "F-70 Indoor Water-Cooled Compressor," datasheet, Sumitomo Heavy Industries. `http://www.shicryogenics.com/products/compressors/f-70-indoor-water-cooled-compressor-series/`.

[126] "2-Port and 4-Port PNA-X Network Analyzer N5247A - 10 MHz to 67 GHz," datasheet, Agilent. `http://literature.cdn.keysight.com/litweb/pdf/N5247-90002.pdf?id=1994005`.

[127] "Keithley Series 2600B SourceMeter SMU Instruments," datasheet, Tektronix. `http://www.tek.com/keithley-source-measure-units/smu-2600b-series-sourcemeter`.

[128] "$|Z|$ Probe 67 GHz," datasheet, Cascade Microtech. https://www.cascademicrotech.com/files/ZProbe67_SS_0310.pdf.

[129] A. Davidson, K. Jones, and E. Strid, "LRM and LRRM calibrations with automatic determination of load inductance," in *36th ARFTG Conference Digest*, vol. 18, pp. 57–63, Nov 1990.

[130] "CSR Calibration Substrates," datasheet, Cascade Microtech. https://www.cmicro.com/products/calibration-tools/csr-cal-substrates.

[131] L. Tiemeijer, R. Havens, A. B. M. Jansman, and Y. Bouttement, "Comparison of the "pad-open-short" and "open-short-load" deembedding techniques for accurate on-wafer RF characterization of high-quality passives," *IEEE Trans. Microw. Theory Techn.*, vol. 53, no. 2, pp. 723–729, 2005.

[132] "Making Accurate Intermodulation Distortion Measurements with the PNA-X Network Analyzer, 10 MHz to 26.5 GHz," application note, Keysight Technologies. http://literature.cdn.keysight.com/litweb/pdf/5989-7265EN.pdf?id=1294558.

[133] J. Schleeh, N. Wadefalk, P. . Nilsson, J. P. Starski, and J. Grahn, "Cryogenic broadband ultra-low-noise MMIC LNAs for radio astronomy applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, pp. 871–877, Feb 2013.

[134] M. W. Pospieszalski, W. J. Lakatosh, R. Lai, K. L. Tan, D. C. Streit, P. H. Liu, R. M. Dia, and J. Velebir, "Millimeter-wave, cryogenically-coolable amplifiers using AlInAs/GaInAs/InP HEMTs," in *Microwave Symposium Digest, 1993., IEEE MTT-S International*, pp. 515–518 vol.2, June 1993.

[135] S. Weinreb, "Noise temperature estimates for a next generation very large microwave array [HEMT LNAs]," in *Microwave Symposium Digest, 1998 IEEE MTT-S International*, vol. 2, pp. 673–676 vol.2, June 1998.

[136] J. Schleeh, H. Rodilla, N. Wadefalk, P. . Nilsson, and J. Grahn, "Characterization and modeling of cryogenic ultralow-noise InP HEMTs," *IEEE Transactions on Electron Devices*, vol. 60, pp. 206–212, Jan 2013.

[137] A. Akgiray, S. Weinreb, R. LeBlanc, M. Renvoise, P. Frijlink, R. Lai, and S. Sarkozy, "Noise measurements of discrete HEMT transistors and application to wideband very low-noise amplifiers," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 61, pp. 3285–3297, Sept 2013.

[138] A. H. Coskun and J. C. Bardin, "Cryogenic small-signal and noise performance of 32nm SOI CMOS," in *2014 IEEE MTT-S International Microwave Symposium (IMS2014)*, pp. 1–4, June 2014.

[139] G. Dambrine, A. Cappy, F. Heliodore, and E. Playez, "A new method for determining the FET small-signal equivalent circuit," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 36, pp. 1151–1159, Jul 1988.

[140] K. Xia, G. Niu, D. Sheridan, and S. Sweeney, "Input non-quasi-static effect in SiGe HBTs and its impact on noise modeling," in *Proceedings of the Bipolar/BiCMOS Circuits and Technology Meeting, 2005.*, pp. 180–183, Oct 2005.

[141] K. Xia, G. Niu, D. Sheridan, and S. Sweeney, "Frequency and bias-dependent modeling of correlated base and collector current RF noise in SiGe HBTs using quasi-static equivalent circuit," *Electron Devices, IEEE Transactions on*, vol. 53, pp. 515–522, March 2006.

[142] M. Rudolph, *Introduction to Modeling HBTs*. Norwood, MA: Artech House Inc., 2006.

[143] H.-Y. Chen, K.-M. Chen, G.-W. Huang, and C.-Y. Chang, "An improved parameter extraction method of SiGe HBTs' substrate network," *IEEE Microwave and Wireless Components Letters*, vol. 16, pp. 321–323, June 2006.

[144] N. F. Mott, "On the transition to metallic conduction in semiconductors," *Canadian Journal of Physics*, vol. 34, no. 12A, pp. 1356–1368, 1956.

[145] P. J. Tasker and M. Fernandez-Barciela, "HBT small signal t and pi model extraction using a simple, robust and fully analytical procedure," in *Microwave Symposium Digest, 2002 IEEE MTT-S International*, vol. 3, pp. 2129–2132 vol.3, June 2002.

[146] F. J. Ohkawa, "Temperature dependence of electrical resistivity of metals," *Journal of the Physical Society of Japan*, vol. 44, no. 4, pp. 1105–1111, 1978.

[147] A. K. Jonscher, "Semiconductors at cryogenic temperatures," *Proceedings of the IEEE*, vol. 52, pp. 1092–1104, Oct 1964.

[148] A. Leuther, A. Tessmann, I. Kallfass, R. Losch, M. Seelmann-Eggebert, N. Wadefalk, F. Schafer, J. D. G. Puyol, M. Schlechtweg, M. Mikulla, and O. Ambacher, "Metamorphic HEMT technology for low-noise applications," in *Indium Phosphide Related Materials, 2009. IPRM '09. IEEE International Conference on*, pp. 188–191, May 2009.

[149] E. Ramirez-Garcia, N. Zerounian, P. Crozat, M. Enciso-Aguilar, P. Chevalier, A. Chantre, and F. Aniel, "SiGe heterojunction bipolar transistor issues towards high cryogenic performances," *Cryogenics*, vol. 49, no. 11, pp. 620 – 625, 2009. Workshop on Low Temperature Electronics 2008WOLTE 2008.

[150] J. te Winkel, "Extended charge-control model for bipolar transistors," *IEEE Transactions on Electron Devices*, vol. 20, pp. 389–394, Apr 1973.

[151] I. Angelov, H. Zirath, and N. Rosman, "A new empirical nonlinear model for HEMT and MESFET devices," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 40, pp. 2258–2266, Dec 1992.

[152] S. Maas, "FET models for volterra-series analysis," *Microwave Journal*, vol. 42, p. 260, May 1999.

[153] I. Angelov, L. Bengtsson, and M. Garcia, "Extensions of the chalmers nonlinear HEMT and MESFET model," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 44, pp. 1664–1674, Oct 1996.

[154] C. McAndrew, J. Seitchik, D. F. Bowers, M. Dunn, M. Foisy, I. Getreu, M. McSwain, S. Moinian, J. Parker, D. Roulston, M. Schroter, P. van Wijnen, and L. Wagner, "VBIC95, the vertical bipolar inter-company model," *Solid-State Circuits, IEEE Journal of*, vol. 31, pp. 1476–1483, Oct 1996.

[155] M. Schrter and A. Chakravorty, *Compact hierarchical modeling of bipolar transistors with HICUM*. World Scientific, Singapore, 2010.

[156] "The Mextram Bipolar Transistor Model," manual, Auburn University. `http://www.eng.auburn.edu/~niuguof/mextram/_downloads/MextramDefinition.pdf`.

[157] B. O. Woods, H. A. Mantooth, and J. D. Cressler, "SiGe HBT compact modeling for extreme temperatures," in *Semiconductor Device Research Symposium, 2007 International*, pp. 1–2, Dec 2007.

[158] "Wide Temperature Range Compact Modeling (93-393K)," progress report, Auburn University. `http://www.eng.auburn.edu/~niuguof/research/modeling.html`.

[159] R. Minasian, "Intermodulation distortion analysis of MESFET amplifiers using the volterra series representation," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 28, pp. 1–8, Jan 1980.

[160] A. E. Parker and G. Qu, "Intermodulation nulling in HEMT common source amplifiers," *IEEE Microwave and Wireless Components Letters*, vol. 11, pp. 109–111, March 2001.

[161] E. R. Srinidhi and G. Kompa, "Investigation of IMD3 in GaN HEMT based on extended volterra series analysis," in *Microwave Integrated Circuit Conference, 2007. EuMIC 2007. European*, pp. 52–55, Oct 2007.

[162] S. A. Maas, B. L. Nelson, and D. L. Tait, "Intermodulation in heterojunction bipolar transistors," *IEEE Transactions on Microwave Theory and Techniques*, vol. 40, pp. 442–448, Mar 1992.

[163] A. Samelis and D. Pavlidis, "Mechanisms determining third order intermodulation distortion in AlGaAs/GaAs heterojunction bipolar transistors," *IEEE Transactions on Microwave Theory and Techniques*, vol. 40, pp. 2374–2380, Dec 1992.

[164] G. Niu, Q. Liang, J. D. Cressler, C. S. Webster, and D. L. Harame, "RF linearity characteristics of SiGe HBTs," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, pp. 1558–1565, Sep 2001.

[165] J. C. Pedro and J. Perez, "Accurate simulation of GaAs MESFET's intermodulation distortion using a new drain-source current model," *IEEE Transactions on Microwave Theory and Techniques*, vol. 42, pp. 25–33, Jan 1994.

[166] J. Vuolevi and T. Rahkonen, "Extraction of a nonlinear AC FET model using small-signal s-parameters," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 50, pp. 1311–1315, May 2002.

[167] "OLS in Matrix Form," course note, Stanford University. `https://web.stanford.edu/~mrosenfe/soc_meth_proj3/matrix_OLS_NYU_notes.pdf`.

[168] J. Macdonald, "Nonlinear distortion reduction by complementary distortion," *IRE Transactions on Audio*, vol. AU-7, pp. 128–133, September 1959.

[169] H. S. Black, "Stabilized feed-back amplifiers," *Transactions of the American Institute of Electrical Engineers*, vol. 53, pp. 114–120, Jan 1934.

[170] N. Pothecary, *Feedforward Linear Power Amplifiers*. Artech House, Norwood, MA, 1999.

[171] V. Aparin and C. Persico, "Effect of out-of-band terminations on intermodulation distortion in common-emitter circuits," in *Microwave Symposium Digest, 1999 IEEE MTT-S International*, vol. 3, pp. 977–980 vol.3, June 1999.

[172] V. Aparin and L. E. Larson, "Modified derivative superposition method for linearizing FET low-noise amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, pp. 571–581, Feb 2005.

[173] V. Aparin, *Linearization of CDMA Receiver Front-Ends*. Phd thesis, University of California San Diego, 2005.

[174] V. Aparin and L. E. Larson, "Linearization of monolithic LNAs using low-frequency low-impedance input termination," in *Solid-State Circuits Conference, 2003. ESSCIRC '03. Proceedings of the 29th European*, pp. 137–140, Sept 2003.

[175] S. Narayanan, "Transistor distortion analysis using volterra series representation," *Bell System Technical Journal, The*, vol. 46, pp. 991–1024, May 1967.

[176] M. Iwamoto, P. Asbeck, T. Low, C. Hutchinson, J. Scott, A. Cognata, X. Qin, L. Camnitz, and D. D'Avanzo, "Linearity characteristics of GaAs HBTs and the influence of collector design," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 48, pp. 2377–2388, Dec 2000.

[177] "4 - 12 GHz Low-Noise Amplifier," datasheet, Cosmic Microwave Technology. http://radiometer.caltech.edu/datasheets/amplifiers/CITCRYO4-12ADataSheet.pdf.

[178] J. Fernandez, "A noise-temperature measurement system using a cryogenic attenuator," *TMO Progress Report*, pp. 42–135, 1998.

[179] "Noise Figure Measurement Accuracy The Y-Factor Method," application note, Keysight Technologies. http://cp.literature.agilent.com/litweb/pdf/5952-3706E.pdf.

[180] W. T. Wong, A. H. Coksun, and J. C. Bardin, "Design and characterization of a wideband high-dynamic range SiGe HBT cryogenic low noise amplifier," in *IEEE MTT-S International Microwave Symposium (IMS), submitted*, June 2017.

[181] "CIT118 Low-Noise Amplifier ," datasheet, Cosmic Microwave Technology. http://radiometer.caltech.edu/datasheets/amplifiers/CIT118DataSheet.pdf.

[182] "LNF-LNC03_14A low noise amplifier," datasheet, Low Noise Factory. http://www.lownoisefactory.com/files/5014/7385/8794/LNF-LNC03_14A.pdf.

[183] "LNF-LNC4_16B low noise amplifier," datasheet, Low Noise Factory. http://www.lownoisefactory.com/files/3114/7392/8774/LNF-LNC4_16B.pdf.

[184] "CITLF4 low-noise amplifier," datasheet, Cosmic Microwave Technology. http://radiometer.caltech.edu/datasheets/amplifiers/CITLF4.pdf.

[185] M. Aoki, S. Hanamura, T. Masuhara, and K. Yano, "Performance and hot-carrier effects of small CRYO-CMOS devices," *IEEE Transactions on Electron Devices*, vol. 34, pp. 8–18, Jan 1987.

[186] S. H. Hong, G. B. Choi, R. H. Baek, H. S. Kang, S. W. Jung, and Y. H. Jeong, "Low-temperature performance of nanoscale MOSFET for deep-space RF applications," *IEEE Electron Device Letters*, vol. 29, pp. 775–777, July 2008.

[187] G. Dambrine, H. Happy, F. Danneville, and A. Cappy, "A new method for on wafer noise measurement," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 41, no. 3, pp. 375–381, 1993.

[188] "Optimizing On-Wafer Noise Figure Measurements up to 67 GHz," application note, Keysight Technologies. http://literature.cdn.keysight.com/litweb/pdf/5991-2524EN.pdf?id=2355063.

[189] A. Siligaris, G. Pailloncy, S. Delcourt, R. Valentin, S. Lepilliet, F. Danneville, D. Gloria, and G. Dambrine, "High-frequency and noise performances of 65-nm MOSFET at liquid nitrogen temperature," *Electron Devices, IEEE Transactions on*, vol. 53, no. 8, pp. 1902–1908, 2006.

[190] L. Boglione, "Considerations on the 4 NT0/TMIN ratio and the noise correlation matrix of active and passive two-port networks," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, pp. 4145–4153, Dec 2013.

[191] "In-House PCB Prototyping," tech guide, LPKF. http://www.lpkf.com/_mediafiles/4080-techguide-in-house-pcb-prototyping-en.pdf.

[192] J. Hacker, M. Seo, A. Young, Z. Griffith, M. Urteaga, T. Reed, and M. Rodwell, "THz MMICs based on InP HBT technology," in *Proc. IEEE IMS*, pp. 1126–1129, May 2010.

[193] J. Hacker, M. Urteaga, M. Seo, A. Skalare, and R. Lin, "InP HBT amplifier MMICs operating to 0.67 THz," in *Proc. IEEE IMS*, pp. 1–3, June 2013.

[194] J. Li, Y. Royter, P. Patterson, T. Hussain, J. Duvall, M. C. Montes, D. Le, D. Hitko, M. Sokolich, D. Chow, and K. Elliott, "Heterogeneous wafer-scale integration of 250nm, 300GHz InP DHBTs with a 130nm RF-CMOS technology," in *Proc. IEEE IEDM*, pp. 1–3, Dec 2008.

[195] J. C. Bardin, A. H. Coskun, M. Ayata, Z. G. Boynton, and J. C. Li, "Broadband noise performance of heterogeneously integrated InP BiCMOS DHBTs," *IEEE Electron Device Letters*, vol. 35, pp. 998–1000, Oct 2014.

[196] J. C. Bardin, J. C. Li, A. H. Coskun, M. Ayata, and Z. G. Boynton, "Microwave noise properties of heterojunction bipolar transistors," in *2014 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM)*, pp. 17–24, Sept 2014.

[197] J. C. Li, *Design Considerations for 400 GHz InP/InGaAs Heterojunction Bipolar Transistors*. Phd thesis, University of California San Diego, 2006.

[198] G. W. t Hooft and C. van Opdorp, "Temperature dependence of interface recombination and radiative recombination in (Al, Ga)As heterostructures," *Applied Physics Letters*, vol. 42, no. 9, pp. 813–815, 1983.

[199] "Converting a Series 2400 SourceMeter SCPI Application to a Series 2600 System SourceMeter Script Application," application note, Tektronix. http://www.tek.com/sites/tek.com/files/media/document/resources/2616%20SCPI_to_TSP_AN.pdf.

[200] "SCPI Command Tree," application note, Keysight Technologies. http://na.support.keysight.com/pna/help/latest/Programming/GP-IB_Command_Finder/SCPI_Command_Tree.htm.

[201] "Data File Formats," application note, Applied Wave Resarch. https://awrcorp.com/download/faq/english/docs/Users_Guide/data_file_formats.html.

[202] D. M. Pozar, *Microwave Engineering*. John Wiley & Sons, 2005.

[203] "Microwave Office Optimization Method," application note, Applied Wave Research. https://awrcorp.com/download/faq/english/docs/simulation/sa_basics.html#optimization_methods.