UNIVERSITY OF GLASGOW

Fabrication, Characterisation and Modelling of Heterojunction Bipolar Transistors and Resonant Tunnelling Diodes based on Indium Phosphide

by

Giangiacomo Cramarossa

A thesis submitted in fulfillment for the degree of Master by Research

in the

RES. DIVISION OF ELECTRONICS AND NANOSCALE ENGINEERING

School of Engineering

July 2012

UNIVERSITY OF GLASGOW

Abstract

School of Engineering
Research Division of electronics and nanoscale engineering

Master by research

by Giangiacomo Cramarossa

The ever increasing demand for higher power devices at higher frequencies has prompted much research recently into the InGaAs/InP material system; this thesis describes two alternative devices (based on InGaAs/InP) that can be employed into THz–oscillator circuits: the hetero-junction bipolar transistor and the resonant tunnelling diode.
The challenges of designing with those devices includes several aspects that span from the simple fabrication scaling effort to the physics of the device and its intrinsic properties. For these reasons a deep understanding of the devices is an essential prerequisite for future applications such as oscillator design.
In this work, HBTs and RTDs were fabricated with standard photolithographic techniques: the minimum features reached were 3x3 μm^2 for metal deposition and 2 μm for gaps between objects. In spite of their good DC–to–RF conversion rate ($\approx 36\%$), HBTs main disadvantage is connected to their fabrication: their performances (f_T and f_{max}) strongly depends on the dimensions of contacts and distances between terminals. Medium–scaled devices (emitter area $\approx 10x10\,\mu$m^2) reached an f_{max} of 5 GHz
RTDs, on the other side, have a lower DC–to–RF figure (theoretical $\approx 20\%$, practical $\approx 2\%$) but their performance and mainly determined by their peculiar layer structure more than their dimensions. Fabricated RTDs showed a PVCR of 110 KA/cm^2 with a PVCR of 1.5.

These figures are quite promising for THz–oscillator production: recent work produced GHz–sources (17.5 GHz with an output power of -21.83 dBm) with similar devices.

Contents

List of Figures

List of Tables

To me, my fiancée Rosangela and my whole Family
and to all those who in their life had the chance to throw the towel
and did't...

Acknowledgements

First and foremost, I would like to thank my supervisor, Dr. Edward Wasige, for having made all this possible. His continuous suggestions and advices during these years constituted an immeasurable value to me.

Thanks to group–mates and office–mates: their friendship helped me in the tough moments.
My appreciation to the Technical Staff of the School of Engineering and the James Watt Nanofabrication Centre: without their professional help none of activities running every day would see success.
A special thought to my italian friends who made more bearable the distance from home.

Thanks to my family: their unconditional support has been the essential and indispensable element to all my achievements.
Finally my fiancée Rosangela: she is the starting point of my thoughts, she is the purpose of my actions. Every day I try to be a better person than the day before: I really hope she's proud of me.

Chapter 1

Introduction

1.1 The terahertz frequency range

The terahertz (THz) frequency range[1] of the electromagnetic spectrum, located between 0.1 and 10 THz (fig. 1.1), can be described as the most scientifically useful and least explored region of the spectrum. Scientists have been exploiting THz waves for a broad variety of purposes: the main fields where those waves find application are sensing, imaging and communications.

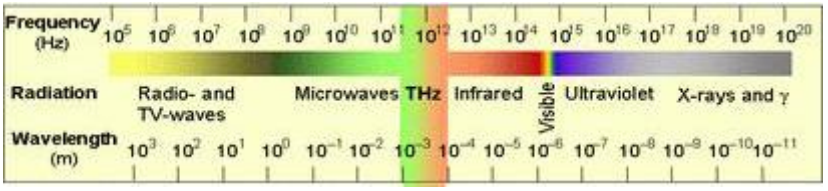


FIGURE 1.1: Spectrum of electromagnetic radiation: frequencies spanning from $0.1 \cdot 10^{12}$ to $10 \cdot 10^{12}$ Hz correspond to the terahertz frequency range.

Historically, spectroscopy has been amongst the first applications of THz technology [1]: for instance, in the development of basic fingerprints of simple molecules

[1]Some Authors refer to the waves in the THz range as **T–rays**, as it happens for the X–rays or the γ-rays. The THz frequencies range correspond to the wave–length interval of $[30\mu m; 3mm]$; the formula $f = c/\lambda$ connects the frequency, f, to the wavelength, λ, by constant c, the speed of light in vacuum.

(i.e. water, carbon monoxide or ozone). Each molecule, infact, has unique and particular rotational, vibrational and translational modes which lie in the THz range thus allowing the identification of chemical substances [2]. Once the spectral feature of a particular molecule is precisely known, other scientists would use this information to identify the same molecules in remote or non-accessible environments. Formation of stars and stellar clouds has been studied by radio-astronomers who could identify remotely native elements [3, 4]. Atmospheric scientists are able to measure the constituents of the upper atmosphere from airborne platforms and satellites (i.e. ozone cycles or pollution levels) [5, 6].

A new door was also opened in medical field when biomaterials were found to interact with terahertz radiation [6]: DNA analysis and disease detection could be performed much easier than in the past [7]. An important factor, favoring the TeraHertz–rays, is the relative weakness of the photons in this range ($\hbar\omega \approx 4 - 40$ meV) with respect of X–rays ($\hbar\omega \approx 0.120 - 120$ KeV) normally used for imaging [8, 9].

THz imaging finds application also in security sector: when a terahertz system is properly used, concealing barriers (such as packaging, cardboard, walls or clothing) could be penetrated and literally seen through [10, 11]. In such a condition, the hidden materials (smuggled good, drugs, explosives or chemical agents) could be detected basing the search on a known spectral fingerprint [2, 12].

Precious advantages can be brought to short–range radar–sensing by T–rays: they can sustain significantly higher bandwidth (than microwaves) thanks to the relatively short wavelength; on the other hand, the T–rays' wavelength is long

enough to reduce the interferences due to the Rayleigh scattering[2] thus penetrating through smoke or fog (further than optical radiation). Short-range battlefield communication, where smoke blocks the infrared transmission (IR), is a very likely application.

The advantage of THz over IR for indoor applications is that it occupies an extremely quiet band without noise or background clutter. Conventional wireless techniques for communication use microwaves at very low power. THz could increase the rate of information–transfer as well as the volume as now–a–days wireless communication technology requires more bandwidth for communication and data transfer [13]. Although the high atmospheric attenuation at terahertz frequencies makes it difficult to have long range mobile–communication, high–bandwidth, short–range and line–of–sight wireless link is completely realisable.

Terahertz technology development find applications also in plasma diagnostics [14] and adaptive cruise control systems [15].

1.2 The Technology Gap

It is clear that there are many present and future applications of terahertz technology: in fact, if one considers the many applications of the microwave and infrared bands that bound the terahertz spectral region, it seems logical that terahertz

[2]Rayleigh scattering, named after the British physicist Lord Rayleigh, is the elastic scattering of electromagnetic radiation by particles much smaller than the relative wavelength. The particles may be individual atoms or molecules. It can occur when radiation travels through transparent solids and liquids, but is most prominently seen in gases.

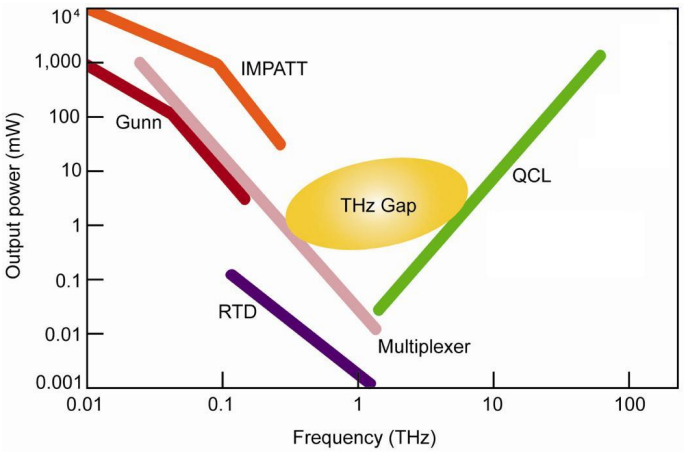


FIGURE 1.2: Terahertz power performance of different sources around the THz gap. RTD: Resonant tunnel diode. IMPATT: Impact ionization avalanche transit-time diode. Gunn: Gunn diode. QCL: Quantum-cascade laser.

technology should eventually play a role in society as large as these more developed bands. The main reason this has not occurred to date has been the difficulty of achieving a mature technology in this spectral region. The terahertz frequency band spans from classical electronics to quantum optics, or photonics. Below 100 GHz, electronic devices such as amplifiers and oscillators are common. Above 10 THz, solid-state lasers, light-emitting diodes, and optical detectors are available. However, in between, neither of these technologies is particularly practical and the result is the so–called Technology Gap (or Terahertz Gap) in terms of both sources and detectors. Figure 1.2 describes the power–performances of some THz sources as a function of frequency. More details about the single devices will be provided in the next two sections.

1.3 Terahertz Sources

There are two approaches to *fill* the THz–Gap: trying to extend the microwave techniques to higher frequencies or extending IR photonic to longer wavelengths.

Traditional photonic devices (such as semiconductor laser diodes) generate radiation through radiative recombination of electron-hole pairs. The emission frequency of these devices is determined by the energy band gap of the semiconductor used to fabricate the laser diode. For THz emission, the band gap has to be smaller than 40 meV (corresponding to 10 THz); state–of–the–art Quantum Cascade Lasers (QCLs) have been reported to achieve 4 THz at a temperature of 160 K and 1.2 THz at 69 K [16]. Temperature is the main issue in THz optoelectronic devices: the whole system must be cooled down to cryogenic temperatures to make it work.

On the other hand, microwave semiconductor devices can work at room temperature but they are limited by the transit time and parasitic RC time constants of the electronics. The power level of these devices decreases as ν^{-4} with frequency[3]: above 1 THz, microwave based sources result in microwatt level power [17, 18].

Newest applications (like new–generation mobile phones or non–invasive imaging systems) requires compactness, portability and a reduced power consumption: for these reasons researchers focused on a broad variety of technological families: Hetero–junction Bipolar Transistors (HBT) [19, 20], High Electron Mobility Transistors (HEMT) [21, 22], Schottky Diode Multiplier Chains [23, 24], IMpact Avalanche Transit–Time (IMPATT) diodes [25, 26], Gunn diodes [27, 28], TUNNEl injection Transit-Time (TUNNETT) [29, 30] and Resonant Tunneling Diodes (RTD) [31, 32, 33] are amongst the most used THz sources.

All these devices can be grouped into two big families: three– and two–terminals

[3]The spatial frequency, ν, is a measure of how often sinusoidal components of the structure repeat per unit of distance. The SI unit of spatial frequency is cycles per meter. It is defined by $\nu = 1/\lambda$, hence $\nu = f/c$.

devices. Although three–terminals have higher DC–to–RF conversion efficiency[4] compared to two–terminals ones [34, 35], they have higher phase noise values compared to the two–terminals [36]. Another great advantages of active two-terminals devices is their simpler fabrication: they typically have vertical layer structures and no fabrication compromises need to be made (for example in HBTs fabrication the base access resistance significantly affects the frequency limits of amplifiers and oscillators).

Heterojunction bipolar transistors (HBT) have been reported to provide 126 mW and 63 mW respectively at 100 GHz and 311 GHz [19, 20]. The highest power efficiency (34%) was recorded with an output power of 160 mW at a frequency of 2 GHz [35].

Along with HBTs, high electron mobility transistors (HEMTs) have shown the best (up–to–date) DC–to–RF efficiency: 36% [34, 35]. The fastest oscillators based on HEMTs provided oscillating frequencies at 254, 314, and 346 GHz. The measured output was 158, 46, 25 μW, respectively [21, 22]. The oscillators utilized HEMTs with a newly developed 35-nm gate process.

Schottky diode multiplier chains have been exploited for many decades at sub-millimetre wavelengths, in particular when all–solid–state solutions were required. Their DC–to–RF conversions efficency is 30% at 190 GHz which drops to 9% to 750 GHz and 4% to 1.5 THz [24]. The highest frequency achieved is 2.55 THz with 0.1 μW of output power [37]; while the highest power levels achieved in the THz region are 150 mW at 100 GHz and 35 mW at 200 GHz [24].

[4]The efficiency is defined as the ratio of the Useful Power Output by the Total Electrical Power Consumed, typically denoted by the Greek letter η. $\eta = P_{RFout}/P_{DC} = P_{RFout}/(V_{DC} \times I_{DC})$

Impact avalanche transit–time (IMPATT) diodes have given a maximum output power of 78 mW at 185 GHz with an efficiency of 2.3%; at 285 GHz, an output power of 7.5 mW with an efficiency of 0.35% was obtained [25, 26]. The highest frequency measured to date is 394 GHz, which provided less of 200μW [25]. Higher RF power and oscillation frequency were achieved by cooling the diode to 77K (liquid nitrogen): devices generate 2 – 7.5 mW power in the 300 – 400 GHz frequency range [38].

Also Transferred Electron Devices (aka Gunn diodes) have a very low DC–to–RF efficiency: 0.07 – 0.24% [39]. On the other hand, Gunn diodes provide an acceptable output level: 80 mW at 152 GHz that drastically drops to 25 mW at 162 GHz [40]. State-of-the-art Gunn devices generate 77 μW power at 400 – 560 GHz frequencies [41].

Presently the maximum operating frequency range of TUNNETT devices is 355 GHz with power output of 140 μW [30]. Higher valuers of output power have been achieved at 103 GHz: 200 mW that corresponds to 2.3% DC–to–RF efficiency [42].

1.4 RTD-based Sources

In 1974 the first resonant tunnelling structure was proposed by Chang, Esaki and Tsu [43]; since then, the RTDs have been deeply studied and improved in order to exploit the negative differential resistance (NDR) of the devices current–voltage characteristic for the well–known purposes: the development of THz sources. The two key–parameters that researchers focused on are the oscillating frequency and

the output power.

Twenty years ago Brown et al. reported a fundamental frequency of 712 GHz [44]. It has been the highest frequency ever achieved until 2012 when Feiginov et al. published their work: 1.11 THz with an output power of 1 μW [32]; the same year Teranishi et al. set another record of 1.08 THz with a good output power level of 5.5 μW [45].

From the point of view of the oscillating frequency, RTDs have shown their potential in reaching the THz–region [46]; on the other hand more problems were found in achieving usable power–levels: average DC–to–RF efficiency is still below 1%: even though very early theoretical studies [47] have demonostrated that RTDs are able to achieve efficiencies up to 20% [5].

The reasons for the low output power (of the RTD-based oscillators) can be connected to the low–frequency parasitic bias oscillations [48] and the inefficient oscillator circuit topologies employed [49]. Different oscillators layout have been implemented in either *waveguide* or *planar* technology in order to suppress or reduce these oscillation and reach the THz–gap. Two brief examples of those topologies now follow.

In waveguide RTD oscillators [31, 50], parasitic oscillations were minimized by a *lossy transmission line* section along the DC bias line: in this way oscillations were suppressed. Figure 1.3 shows the schematic diagram of a quasi-optical oscillator designed for the 0.1 THz region. The RTD is mounted in a rectangular waveguide (WR-6) that opens abruptly to a round diameter coupling hole within

[5] The maximum output power achieveable by a RTD is P_{MAX}, where $P_{MAX} = (3/16) \cdot \Delta V \cdot \Delta I$. More details will be provided in chapter 4

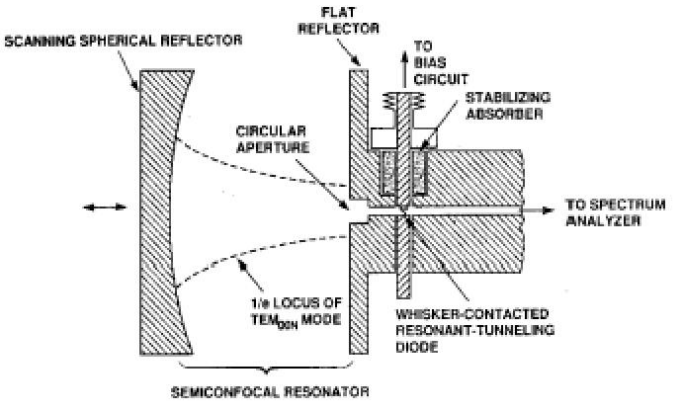


FIGURE 1.3: Cross-section of a quasioptical waveguide resonant tunnelling diode oscillator.

the middle of a flat metallic plate. This plate forms one reflector of a semiconfocal open resonator. The waveguide portion of the oscillator is typical of RTD waveguide oscillators. The diode is dc biased by a coaxial circuit that suppresses spurious oscillations by means of a very lossy section of transmission line placed in close proximity to the diode chip. The lossy material is an iron-loaded epoxy. Waveguide oscillators achieve efficiencies of $\approx 1.6\%$ well below the theoretical predictions. The lossy line diminishes the signal level of parasitic oscillations so that they do not significantly interfere with the main oscillations, but the presence of bias oscillations means that there is less power available from the device for conversion into an RF signal.

Planar RTD oscillators, on the other hand, eliminate parasitic bias oscillations in an oscillator circuit by employing an external resistance in parallell to the NDR device [51, 52]. A non-linear (diode) resistor was first used instead of a linear resistor to reduce the DC power consumption of the stabilizing resistor [52]. Later on Schottky diodes [51, 53] were employed for the same purpose (fig. 1.4).

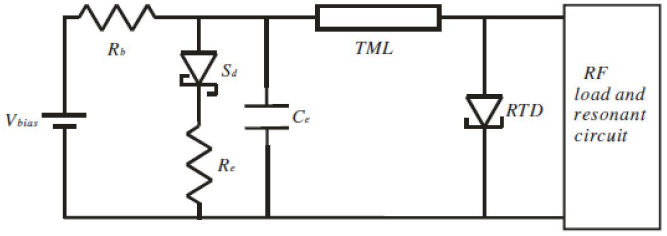


FIGURE 1.4: Bias stabilization scheme for sub–mm–wave RTD oscillators. R_b is the resistance of the bias line. S_d, R_e and C_e are the Schottky diode, external resistor and external capacitor, respectively, which form the stabilising circuit. TML is the quarter-wave length transmission line at the oscillation frequency.

On this line of the research, two works should be mentioned: in [54, 55] good results were shown in terms of frequency achievement (respectively, 1 500 GHz and 650 GHz) but in both cases the output power was very limited (40 μW and 28 μW respectively).

In [55] an inaccuracy in the design of the components[6] actually lead the oscillator to operate at a different (lower) oscillation frequency. As a consequence the impedance of the oscillator is not matching the measurement equipment and the power (which dependes on impedance match) is strongly reduced.

Another example of an inefficient planar RTD oscillator topology is the 50 GHz RTD oscillator circuit described in [54] where the RF power is taken across the DC stabilising resistor. The circuit used is similar to that described fig. 1.4, but with no Schottky diode used for stabilisation. The circuit also has no decoupling capacitor and no explicit RF load or resonant circuit. In the circuit, the stabilising resistor R_e was 5 Ω, which also acted as the load resistance. With the generated RF power being taken across the stabilising resistance, large losses occur due to

[6] It has been shown by [56] that the quarterwave transmission line ($l = \lambda/4$), used to decouple the RTD oscillator circuit from the DC bias circuit, does not act as a RF open circuit but together with the RTD capacitance and any resonator used combine to determine the frequency of the oscillation.

impedance mismatch to a 50 Ω load.

Despite the extensive literature, RTD–based sources have not yet achieved the best figures expected in terms of both design frequency [51] and output power [57]. To increase the output power, series–connected [58, 59] or parallel–connected RTDs [60, 61] have been employed in oscillator circuits. In the first topology, DC instability was the main issue as it was not possible to bias all the devices at the same bias point in the NDR region. In the second topology, the currents flowing into the devices were adding–up thus reducing the negative differential resistance of the whole circuit making it more difficult to suppress the low–frequency parasitic oscillations.

Recent works [33, 49, 56, 62][7] have shown how parasitic bias oscillations can be reduced by a circuitry comprising properly designed resistor and capacitor (fig. 1.5). This method yields accurate results as long as the external resistance magnitude suppresses the parasitic oscillations. The external capacitor is used to short circuit the RF signal from the NDR device to ground, which further improved the DC stability.

The results achieved by this topology showed interesting experimental results: hydrid–technology oscillators[8] provided 6.81 μW at 7.15 GHz and 4.87 μW at 17.5 GHz [56].

The present work is a study aimed to develop the idea presented in [56]: integrating the whole circuit, the same oscillating frequencies are expected to provide

[7] These publications have been generated by a line of research still active at the University of Glasgow.
[8] Hybrid: monolithic fabricated RTD soldered on a PCB along with non–integrated components

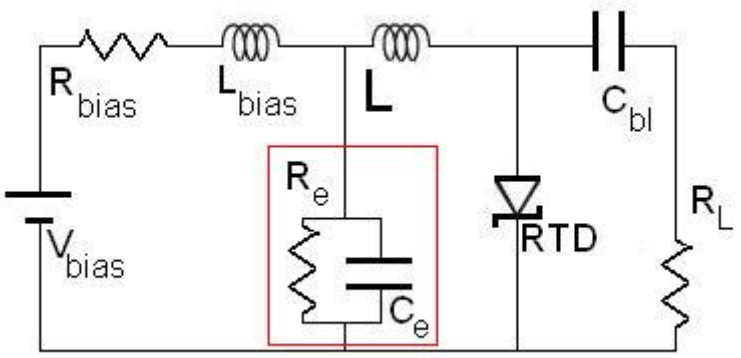


FIGURE 1.5: Single device RTD-based oscillator; the red box highlights the RC decoupling circuit.

output power levels 1000 times higher ($P_{out} \approx 10$mW).

1.5 Thesis Layout

The present thesis is divided in five chapters; following the introduction, the standard nano–fabrication techniques used in the project will be presented.

Chapters 3 and 4 will provide aspects about two terahertz sources: heterojunction bipolar transistors and resonant tunneling diodes. Specifically, chapter 3 will deal with the fabrication, characterisation and modelling of medium–scaled HBTs. Good DC and AC performances have been achieved from an in–house fabricated samples (f_T of 10 GHz for a 6x10 μm^2 emitter area): those (and other) figures will show the practical issues in the attempt of reaching very high frequencies. Nonetheless, an important result was achieved: a new modelling technique [63], that gave good results with other types of transistors [64], has been applied to HBTs with good results.

Chapter 4 will introduce briefly the theory of RTDs: quantum well, tunnelling

transports and relation between the layer–structure and the I–V charateristic. After that, some promising results on RTDs will be presented: I–V and RF characterisation.

Chapter 5 will provide further considerations and development in this line or research.

Chapter 2

Nano–fabrication Processes

This chapter describes the standard steps for the fabrication of semiconductor devices as they have been realized throughout this project. The process involves several basic techniques (or a variation on them) that, as a whole, are called *lithographic processes*[1]. The present chapter will describe each step in its own general aspects: further data will be provided in the following chapters (specifically sect. 3.2 for HBTs and sect. 4.2 for RTDs) and in the appendix A.

All the fabrication activity was carried out in the *James Watt Nanofabrication Centre* – JWNC: this facility houses the clean–room of the School of Engineering.

2.1 Sample Preparation and Cleaning

The current project involved several multiple epilayer structures of InP-based materials: this aspect increased the cost of the materials themselves and led to the choice of cleaving the original wafers (diameter: 3 inches) into smaller samples (12x12 mm^2). The wafer was scribed using a diamond–tipped scribing tool to

[1] Also called photolithography, optical lithography or simply lithography.

keep the size and shape regular. Each sample was then cleaned in an ultrasonic bath of solvent: the small bubbles, created by the phenomenon of cavitation, collapse at high speed physically removing any particulate or contaminant. Acetone is the standard solvent with which organic contaminants are removed from semiconductor substrates. It is an organic compound (propanone, CH_3COCH_3) and hence will not react with any of the materials used. A further ultrasonic bathing in isopropyl alcohol (IPA or propan–2–ol, C_3H_8O) and a reverse osmosis (RO)–water rinse will remove any polar contaminants completing the procedure.

2.2 Photolithography

Photolithography is a process used in microfabrication to selectively remove parts of a thin film (or the bulk of a substrate). It uses light to transfer a geometric pattern from a mask[2] to a light–sensitive film (photoresist, or simply *resist*) on the substrate. A series of chemical treatments then engraves the exposed pattern into the material underneath the photoresist. It is used because it affords accurate control over the shape and size of the objects it creates, and because it can create patterns over an entire surface simultaneously.

After being cleaned, the sample is attached to a spinner in a laminar air flow (LAF) cabinet by means of a vacuum, it is covered in photoresist and, then, made to spin. The rapid acceleration, the rotational speed and the duration determine

[2]In JWNC, a mask for photolithography is four to five inches wide and can fit from ten up to twenty $1x1cm^2$ patterns

TABLE 2.1: Processing informations for photoresist S1818: the Provider [65] suggests the *standard* recipe; it has been improved in–house into two new processes for negative (for metal deposition) and positive (for etching) profile. The positive profile recipe is composed of two subsequent steps.

	Angular Acceleration	**Angular Speed**	**Duration**	**Average Thickness**
	rpm^2	rpm	Seconds	nm
Standard	10000	4000	60	1800
Negative profile	20000	10000	120	1100
Positive profile	500/2000	500/4000	5/30	1800

the thickness of the resist. Following on from the spinning of resist, a pattern must be written into it using a lithography tool. The pattern is then developed using a solvent to selectively remove the exposed (positive tone) areas of the thin film. After rinsing away the solvent, any residual of resist in the developed areas can be removed using a barrel asher. This generates a low power oxygen plasma in a barrel–shaped chamber. The reactive oxygen ions etch the resist at a rate insignificant to the remaining film but which cleans off the residues.

The resist used for this project was Shipley S1818. This is a positive tone resist which is photosensitive to UV light whose wavelength goes from 350 nm to 450 nm. The thickness could be controlled by the spin speed/acceleration: in this project, it was chosen to slightly diverge from the standard processing (see table 2.1). After photoresist spinning, the sample (along with the desired pattern of the photomask) was put onto an hotplate for 90 seconds at 120 ℃[3] and then loaded onto the mask–aligner *Karl Suss MA6*. The sample was then accurately aligned to the pattern and exposed to UV light (365nm for 5 seconds), the exposed resist was then developed using Shipley Microposit Developer Concentrate, hence completing the operation.

In order to get the negative profile, immediatly before UV–ligth exposure, the

[3]This step is called soft–bake, to differentiate from the hard–baking in the oven.

sample was dipped into the developing solution for 75 seconds: it re–hydrates the upper part of the photoresist after the soft–bake, thus reducing the sensibility to the UV–ligth. In this way there will be a different response to development of the top and bottom of the photoresist itself (see fig. liftoffpic for details).

2.3 Metallisation

Metallic thin films are required to form device contacts, passive elements, and interconnections. Electron beam evaporation was the method employed in this project to deposit metals onto the surface of the samples. A Plassys MEB 450 Electron Beam Evaporator (Plassys I) and two Plassys MEB 550S (Plassys II, IV) are the available evaporators in the JWNC facility. Once the sample is loaded into the evaporator, the chamber is pumped down to the process pressure ($2 \cdot 10^{-6}$ Torr) and the desired metal scheme is selected via control software. Following the provided order, the relative metal crucibles inside the evaporator are heated over the melting point by a beam of electrons and metals evaporate onto the wafer where they condensed.

The sample was formerly processed for a *negative sidewalls profile* photoresist: in this way, metals adhere (ideally) only onto the horizontal surfaces and not onto the vertical ones (non–conformal behaviour)[4]. The negative profile of the photoresist's sidewalls creates discontinuities in the metal surfaces such that, when the sample is dipped in resist–stripper solution, all the regions covered in photoresist are

[4] A conformal film defines a morphologically uneven interface with another body and has a thickness that is the same everywhere along the interface. Conformal films may be deposited by thin-film deposition methods, such as plating, chemical vapor deposition or atomic layer deposition.

removed (along with the covering metals) leaving untouched the metals laying directly on the wafer surface. The whole procedure is summarised in figure 2.1

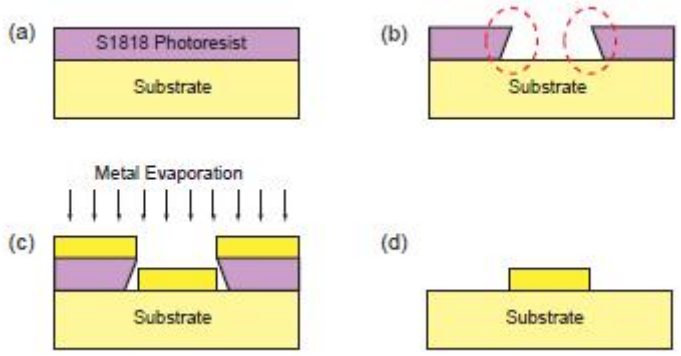


FIGURE 2.1: Summary of the lithographic process of metal deposition. S1818 is deposited onto the wafer (a): it is then soft–baked and *pre–developed*. After UV exposure and development the photoresist gets a negative profile (b). The sample is now ready for blanket evaporation of metals (c): the undercut generates discontinuities in the metallic layer thus allowing the solvent to penetrate and remove the photoresist along with the unwanted metals (d).

Available solutions to remove photoresist are:

- Acetone;
- Shipley **SVC–14**, based on Dimethyl Sulfoxide (or DMSO);
- Shipley **1165**, based on N–methyl–2–pyrrolidine (or NMP);

The surface of semiconductors has a natural tendency to oxidize, hence all the samples were briefly dipped into a mild solution of diluted hydrochloric acid (HCl:RO–H_2O; 1:5) before metallisation. In order to improve the quality of the contacts and their adhesion to the substrate, a Jipelec, Jet First Processor, was used. This process is called Rapid Thermal Annealing (or RTA) [66]: the samples, after being loaded in a nitrogen chamber, are rapidly heated to 280 ℃, the process lasts 60

seconds. During annealing, the individual layers of the contact alloy together and diffuse into the semiconductor, improving the quality of the contact resistance.

2.3.1 Metal–Contact Assessment and Transfer length Method

The quality of metal contacts is extremely important for semiconductor devices and it was constantly monitored in the present project.

When electrons are forced through an interface between two materials (as it happens in the metal–semiconductor contact), the electrons experience a further resistance other than the intrinsic resistance of the materials themselves: the contact resistance (R_{cont}). Transfer length Method[5] (or TLM) was used to assess R_{cont} [67, 68]. Two other important parameters must be introduced before explaining TLM and R_{cont}: sheet resistance, R_{sh} and specific contact resistance, ρ_c. The electrical resistance of a uniform conductor is:

$$R = \rho \cdot \frac{L}{t \cdot W} = \frac{\rho}{t} \cdot \frac{L}{W} = R_{sh} \cdot \frac{L}{W} \tag{2.1a}$$

with

$$R_{sh} = \frac{\rho}{t} \tag{2.1b}$$

where ρ is the electrical resistivity of the contact substrate (MKS: $\Omega \cdot m$), L is the length of the conductor, t and W are, respectively, the thickness and the width of its cross-sectional area. Definition of ρ_c is:

$$\rho_c = \left[\frac{\partial J}{\partial V}\right]^{-1}_{V=0} \tag{2.2}$$

[5] Some Authors prefer to use the name Transmission Line Measurement

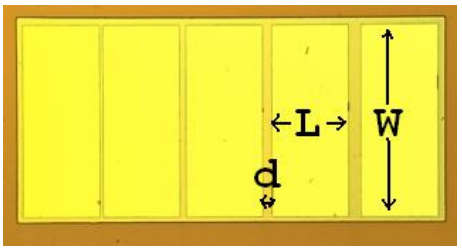


FIGURE 2.2: Optical microscope snapshot of the metal pads for TLM measurement. Each pad measures 100x250 μm^2, the spacings are: 4, 6, 8, 10 μm. It is noticeable the characterist edge of the mesa profile around the interested area ($\approx 5\mu m$).

where J is the current density flowing through the contact (MKS: $\Omega \cdot m^{-2}$).

The TLM consists in placing a series of metallic pads spaced by increasing distances (fig. 2.2): once current is forced to flow between pad_i and pad_{i+1}, voltage drop is recorded hence the corresponding total resistance is calculated. In this way it is possible to measure the total resistance (R_{TOT}) between any two adjacent pads. The R_{TOT} is composed by the contact resistances of the two pads (which will be constant for any pads) and by the resistance of the substrate between the pads themselves (which is proportional to the spacing):

$$R_{TOT} = 2 \cdot R_{cont} + R_{sh} \cdot \frac{d_i}{W_{pad}} \tag{2.3}$$

where d_i is the spacing between pad_i and pad_{i+1}, W_{pad} is the width of the pads and L_{pad} is the length of the pads. The results will be then plotted as in figure 2.3. Interpolating the experimental data, it is possible to compare the regression line with equation 2.3: the expression $2 \cdot \text{R}_{cont}$ is the y–intercept of this equation and its value corresponds to the ideal case of two consecutive pads (no gap in between). From the slope of same graph is possible to calculate the value of R_{sh}. Another extracted parameter is the transfer length L_T (from where the name for the method). It is defined as the distance over which most of the current transfers

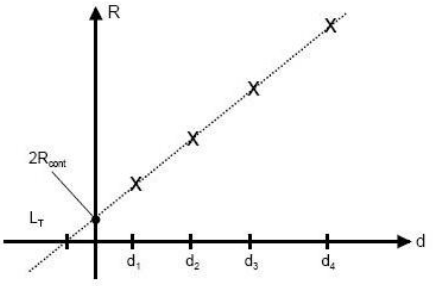


FIGURE 2.3: Ideal behaviour of the total contact resistance R_{cont} versus spacings d_i: real values are compared with the regression line.

from the semiconductor into the metal or from the metal to the semiconductor.

$$L_T = \sqrt{\frac{R_{cont}}{R_{sh}}} \tag{2.4}$$

By using the information above, tests could be performed to improve the contacts. As already stated good ohmic contact is a key–element in devices' fabrication: in the last ten years, a extensive literature has been published by different groups: table 2.2 shows the most common schemes for the layers used in the current project.

TABLE 2.2: Review of the most common metal schemes for InP/InGaAs substrates.

Layer	**Metal Scheme**	**Specific Contact resistance** $\Omega \cdot cm^2$	**Reference**
InP (n–doped)	Ti/Pt/Au	$3.5 \cdot 10^{-7}$	[69, 70]
InGaAs (n–doped)	Ti/W	$7 \cdot 10^{-9}$	[71]
InGaAs (n–doped)	Au/Ge/Ni/Au	$3.8 \cdot 10^{-6}$	[72]
InGaAs (p–doped)	Mn/Au/Ti/Au	$1 \cdot 10^{-5}$	[73]
InGaAs (p–doped)	Ti/Pt/Au	$5.5 \cdot 10^{-7}$	[70, 74]

For practical[6] and technical reasons it was decided to slightly diverge from the Ti/Pt/Au–recipe using Pd instead of Pt [75]: Pt evaporates at a high temperature with a low vapor pressure[7] thus prolonging the process time; these could cause problems such as *out–gassing* of impurities within the evaporation chamber and

[6]Platinum and Palladium are available on two different PLASSYS evaporators in the JWNC clean–room; the machine that provides Pt is an older piece of equipment than the one with Pd: it is frequently under maintenance and the quality of the products is not constant in time. This was one of the practical reasons that induced the use of Pd over Pt.

[7]Vapor pressure, or equilibrium vapor pressure, is the pressure exerted by a vapor in thermodynamic equilibrium with its condensed phases (solid or liquid) at a given temperature in a closed system. The equilibrium vapor pressure is an indication of a liquid's evaporation rate. It relates to the tendency of particles to escape from the liquid (or a solid). A substance with a high vapor pressure at normal temperatures is often referred to as volatile.

the hardening of photoresist. On the other hand, Pd has better physical attributes and similar electrical properties of Pt; table 2.3 shows the relative figures.

TABLE 2.3: Comparison of Pd/Pt physical and electrical properties.

	Platinum	**Palladium**
Melting point	1 769 C	1 552 C
Vapor pressure	0.14 μmHg	0.26 μmHg
Resistivity	$9.85 \cdot 10^{-8}\,\Omega \cdot$ m	$9.93 \cdot 10^{-8}\,\Omega \cdot$ m
Workfunction	5.32 eV	4.99 eV

Some trial–test were run on the Ti/Pd/Au metal scheme for both n– and p– doped substrates: the measured resistance were plotted against the spacings (fig. 2.4) and then the relevant figures were extracted (table 2.4).

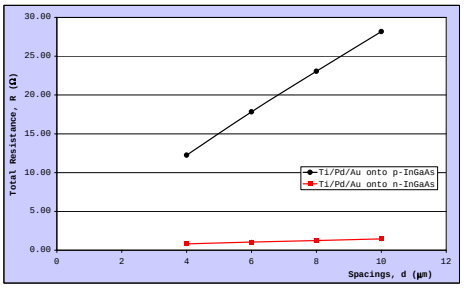


FIGURE 2.4: TLM: measured pad–to–pad total resistances plotted against the spacings: the black line represents the data from the Ti/Pd/Au onto p–doped InGaAs while the red line represents the data from the Ti/Pd/Au onto n–doped InGaAs.

Those numbers will be then use for an initial esteem of the parasitic resistances (par. 3.4.5).

Substrate	Correlation	Gradient	Intercept	Sheet Resistance	Contact Resistance	Transfer Lenght	Specific Contact Resistance
	-	R_{sh}/W_{pad}	$2 \cdot R_{cont}$	R_{sh}	R_{cont}	L_T	ρ_c
	-	$\Omega/\mu m$	Ω	$\Omega/\square$	Ω	μm	$\Omega \cdot \mu m^2$
p–InGaAs	0.999	2.65	1.79	397.24	0.90	0.34	45.41
n–InGaAs	0.999	0.11	0.40	16.00	0.20	1.88	56.69

TABLE 2.4: Extracted parameters from TLM plots, fig. 2.4.

2.4 Thin Film Deposition

In the fabricational process of semiconductor devices, it can be necessary to deposit thin films of dielectric materials, for different purposes:

- hard–masking for etching;
- insulating from external interferences;
- dielectric for integrated capacitors (MIM capacitors[8]);
- planarization.

Table 2.5 summarizes the main dielectrics used in the current project and their relative use. Description follows in the following paragraphs.

TABLE 2.5: List and relative employment of dielectric in use for the project. Two different *recipes* are provided for polyimide.

Material	**Employment**	**Deposition Technique**	**Mask**	**Etching Method**	**Dielectric Constant**
SiO_2	mask and insulator	PECVD	S1818	RIE80+	6.7
Si_3N_4	MIM capacitors	ICP–CVD	S1818	RIE80+	3.9
HSQ	planarization	Spinning	SiO_2	RIE80+	6.7
Polyimide	planar. and insulator	Spinning	S1818	wet etch	3.5
Polyimide	–	–	Aluminum	RIE80+	–

2.4.1 Chemical Vapour Deposition

In the present project, thin film of solid materials (dielectrics) were deposited by the process called Chemical Vapour Deposition (CVD). In a CVD process, the wafer is exposed to one or more volatile precursors, which react on the substrate surface to produce the desired deposit (most commonly silicon dioxide (SiO_2) or

[8]MIM capacitors are integrated passive devices composed of two parallel plates with a dielectric layer between the plates; the fabrication procedure consists in the three following deposition of a metal layer, dielectric insulating film and another metal layer (hence the name Metal–Insulator–Metal): the magnitude of the capacitance is given by: $C = \epsilon A/d$; where ϵ is the permittivity of the dielectric in use while A and d are area and the thickness of the dielectric itself defined by photolithograpy.

silicon nitride (Si_3N_4)). Films deposited in this way are highly conformal over varying device topography. The different types of CVD employed in the present research into semiconductor devices include:

- Low Pressure CVD, *LP–CVD*;
- High Vacuum CVD, *UHV–CVD*;
- Plasma Enhanced CVD, *PE–CVD*;
- Inductively Coupled Plasma, *ICP*.

2.4.2 Spin on glass: HSQ and Polyimide

Another process used to deposit thin film is called spin–on–glass (or SoG): it is used to deposit Hydrogen Silsesquioxane (HSQ) or Polyimide. HSQ and Polyimide are stored in liquid phase: by the means of a pipette, they are transfered from the storage bottle to the sample and then spun.

TABLE 2.6: HSQ and Polyimide: processing informations.

	Angular Acceleration	**Angular Speed**	**Duration**	**Average Thickness**
	rpm^2	rpm	Seconds	nm
HSQ	10000	5000	60	400
Polyimide	10000	1000/4000	30/60	1900

Following hard–baking converts the dielectrics to the solid state; hence they can be patterned using photolithograpy and wet (or dry) etch. Table 2.6 provides details to process HSQ and Polyimide.

2.5 Etching

Etching is the physical and/or chemical process that removes the uppermost layer of the substrate in the areas that are not protected by hard- or soft–masks. According to the application, there may be different demands for etching, like higher etch rates, uniformity, an/isotropy, selectivity, less damage to the surface. Especially for HBT and RTD applications, low etch rate, uniformity, selectivity, anisotropy and less damage to the surface are the main demands.

- *Uniformity* is defined as the etch rate deviation across the wafer.
- *Selectivity* is defined as the ratio of the etch rate for different materials.
- *Isotropy* is the etch rate difference for different crystallographic directions.

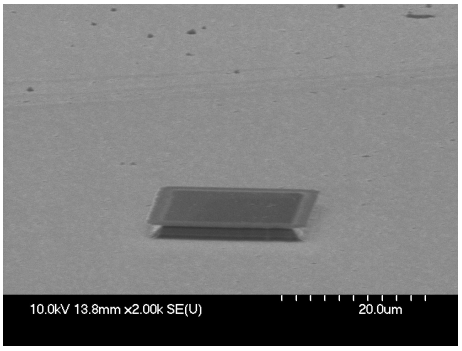
The degree of isotropy A, is depicted as:

$$A = 1 - \frac{R_l}{R_v} \tag{2.5}$$

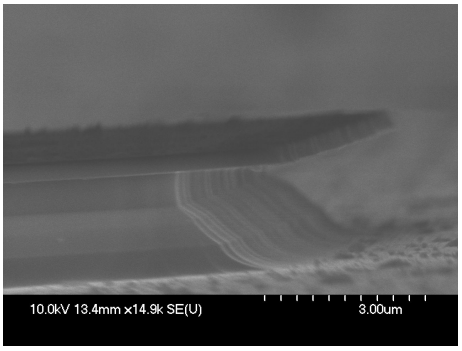
where R_l is the lateral etch rate and R_v is the vertical etch rate. When $A = 0$ then the etching is *perfectly isotropic* and when $A = 1$ it is called as perfectly *anisotropic*.

2.5.1 Wet Etch

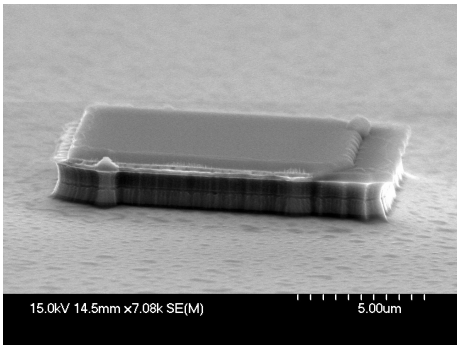
In wet chemical etching, acids are used to erode the material from the surface. Basically, acids react with the exposed material and form other compunds; these compounds are then removed from the surface. Wet etching of III-V materials



(a) Wet etch, top view.



(b) wet etch, side view.



(c) Dry etch.

FIGURE 2.5: SEM snapshots of the resulting structures after wet etch (a, b) and dry etch (c). In the wet–etch, where metal contact acted as hard mask, the process can be strongly isotropic ($R_L \cong R_V$). In the dry–etch, where SiO_2 was used as hard mask, the process can be almost anisotropic ($R_L \ll R_V$).

involves liquid chemical etchants, usually acid solutions in water: an oxidising agent creates an oxide layer at the material surface and then another agent removes the oxide in a reduction reaction.

TABLE 2.7: List of semiconductors alloys and relative etchants in use for the project.

Layer	**HCl:H_2O**	**H_3PO_4:H_2O_2:H_2O**
	3:1	1:1:38
InP	600 nm/min	not etched
InGaAs	not etched	100 nm/min
InAlGaAs	not etched	100 nm/min
InAlAs	not etched	150 nm/min

The selectivity needed to consistently and accurately achieve the correct etch depth in the InP–based materials is available with solutions based upon orthophosphoric

acid (H_3PO_4) or hydrochloric acid (HCl). Table 2.7 provides an overview of the properties of these solutions.

2.5.2 Dry Etch

Dry etch processes are a combination of a physical process, whereby the actual bombardment of the semiconductor by ions etches the semiconductor away, and a reactive process, where the chemical reactivity of the ion removes the material at the surface. Depending on the process conditions, the plasma constituents and the material, one process may dominate. The particles that are etched away are prevented from re-deposition by the low process pressure in the chamber. The advantage of dry etching techniques is the high levels of **anisotropy** (i.e., vertical sidewalls, fig. 3.7), **uniformity** and **etch–rate** control.

A general dry etching process can be divided into the following four steps [76]:

1. Production of active gas species by RF discharge of the reactive gases pumped into the chamber.

2. Diffusion of the active species from the bulk plasma of the discharge onto the surface of wafer.

3. Reaction steps:

   Absorption of the radicals on the surface and ion bombardment.

   Chemical reaction between the reactive species and the substrates.

4. Pump-out of volatile chemical byproducts.

The two main classes of dry etching available in JWNC facilities are:

1. **Reactive Ion Etching** (RIE): is a technique that remove the undesired material through the combined effect of chemical and physical interaction with accelerated ions. RIE can produce both strongly anisotropic profiles and good selectivity between the mask layer and the material to be etched. RF power applied to two parallel plates is used to control both plasma generation and ion acceleration: by this means it is possible to control the etch rate. The consequence is an increase in the ion bombardment energy and hence a deterioration of the etching selectivity and increased sample damage.

2. **Inductively Coupled Plasma** (ICP): The plasma generation is separated from the etching chamber and there are two different RF power generators coupled to the plasma to control ion energy and ion density independently. An inductively coupled RF generator controls the density of the plasma through the chamber walls and, in addition, capacitively coupled RF supply is used to vary the ion acceleration towards the material. The separate process for ions generation and acceleration make it possible to choose a well balanced set of operating parameters, in order to produce the etch profiles required.

Chapter 3

Heterojuntion Bipolar Transistors

The recent advancements of the epitaxial growth techniques allowed a massive exploitation of heterostructures in semiconductor devices. Indium Phosphide-based heterojunction bipolar transistors (HBTs) find applications in very wide-band digital and mixed-signal ICs [77]. Compared to Si-based BJTs and FETs, InP HBTs have higher bandwidth for the same lithographic feature size and higher breakdown voltage for a given bandwidth [78, 79]. These advantages originate from the high electron mobility of the InGaAs base, the high base doping and the high peak electron velocity [80].

This chapter contains a discussion of the physical and electrical properties of heterojunction bipolar transistors. It starts with a description of the theory and the ideal operations of an HBT, sect. 3.1. The chapter goes on to look at the layer structure in use and at the fabrication process, sect. 3.2. DC and RF measurement of fabricated devices are then provided and analysed, sections 3.3, 3.4. From those experimental data a novel extraction technique for the elements of the small–signal–model is finally presented, par. 3.4.4.

3.1 Theory and Operations

As already mentioned, the HBT is an engineering improvement of the Bipolar Junction Transistor (BJT): in these terms the most important of the enhancements is the restriction of the base–emitter back injection current[1] by a dedicated design of the energy–bands of the materials used in the epitaxial structure. From this point–of–view, the *heterojunction* is the pivotal element.

3.1.1 The Heterojunction

The metallurgical junction of a semiconductor that has different dopant types on each side is called a p-n (homo)junction. On the other hand, heterojunctions are formed of two dissimilar materials having different energy gaps[2]. It is customary to use the capital letter N or P to denote the doping type in the layers with larger energy gaps. If high level of doping are reached then one or more plus–signs (or minus–sign) can be added to the doping indications (i.e., p^{-} or N^{++}).

When a p-n homojunction is created, the Fermi–levels on the two sides of the interface align: in this way the conduction band, E_C, and valence band, E_V, on both sides, will arrange themselves accordingly. Since both sides of the junction are made of the same compound (hence same band–gap) the relative differences

[1] Later on this current will be indicated with I_B^*.

[2] From the energy diagram theory, the band–gap is the minimum energy that an electron requires to *jump* from the valence band to the conduction band, this energy is written as E_G

(called *off–set*) between the two E_C and E_V will be the same:

$$\Delta E_C = E_{Cn} - E_{Cp} \tag{3.1a}$$

$$\Delta E_V = E_{Vp} - E_{Vn} \tag{3.1b}$$

$$\Delta E_C = \Delta E_V \tag{3.1c}$$

where ΔE_C is conduction band off–set given by the difference between the values of the conduction bands on both sides of the interface (E_{Cn} and E_{Cp} for the n–doped and p–doped side). Same definition applies to the valence band.

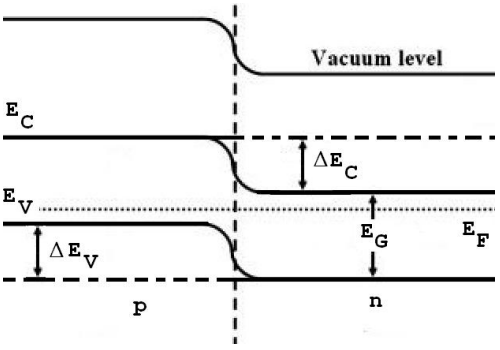


FIGURE 3.1: Homojunction. The left side is p–doped, the right side is n–doped. The Fermi level is aligned: the conduction band and the valence band off-sets (ΔE_C and ΔE_V) are properly formed.

In heterojunctions this *symmetry* is no longer respected: the difference in the band–gaps leads to different arrangements for ΔE_C and ΔE_V ($\Delta E_C \neq \Delta E_V$). For the semiconductors used in this project (InGaAs and InP) Anderson's rule[3] has proven to give accurate predictions for the band offsets [81].

The relevant properties of the involved materials, that must be taken into account for a precise description of the charge dynamics at the heterojunction, are: the

[3]Anderson's rule is used for the construction of energy band diagrams of the heterojunctions; it states that, when constructing an energy band diagram and before the junction itself is formed, the vacuum levels of the two semiconductors on both side of the heterojunction should be aligned at the same energy.

band–gap, E_G and the *electron affinity*, χ. The energy difference between the valence band (E_V) and conduction band (E_C) is the band–gap, while the electron affinity of each material represents the energy difference between the conduction band and the vacuum level.

A new set of equations can be written then:

$$\Delta E_C = \Delta\chi = \chi_{wide} - \chi_{narrow} \tag{3.2a}$$

$$\Delta E_V = \Delta E_G - \Delta\chi \tag{3.2b}$$

where

$$\Delta E_G = E_{G,wide} - E_{G,narrow} \tag{3.2c}$$

Table 3.1 provides the figures for the InGaAs/InP heterostructure in use[4].

	InGaAs eV	InP eV	Interface (InGaAs/InP) eV
E_G	0.75	1.35	-
χ	4.60	4.38	-
ΔE_G	-	-	0.60
ΔE_C	-	-	0.22
ΔE_V	-	-	0.38

TABLE 3.1: Figures for the InGaAs/InP heterojunction.

The direct conseguence of this *asimmetry* is that the carriers flowing from one material to another will experience different potential barriers: in the case in use (as figure) the holes from the p–InGaAs will find a bigger step (0.38 eV) than the electrons (0.22 eV) moving in the opposite direction.

[4]See table 3.3 for HBT epitaxial layers.

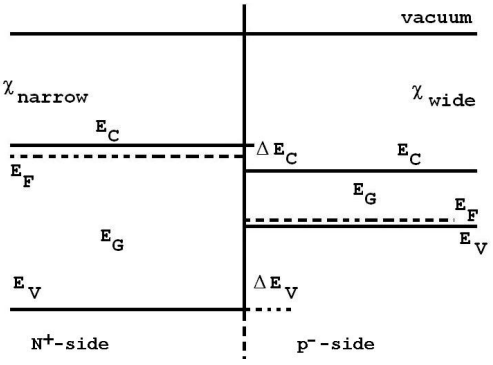


FIGURE 3.2: Example of an ideal p^-N^+–heterojunction. The left side is n–doped, the right side is p–doped. The junction is not formed already and the Fermi levels are not aligned yet: the conduction band and the valence band off-sets (ΔE_C and ΔE_V) are formed as expected.

3.1.2 The Heterojunction Bipolar Transistor

A bipolar transistor mainly consists of three layers: the emitter, the base and the collector. The emitter and the collector are doped to the same type, while the base layer in between has the opposite type of doping. These three layers form two p–n junctions, connected *back–to–back*.

In the present project, the base–emitter (b–e) interface is grown with a InGaAs/InP system thus creating the heterojunction. It will be shown in the next paragraphs how the engineering of the b–e junction improves the performances of the devices. Following on, the base–collector junction is made of InGaAs on both side (see table 3.3 for details). About the doping type, a n–p–n pattern was used for emitter, base and collector respectively: in such a configuration, the electrons are the conducting carriers between the emitter and the collector. As the electrons mobility is higher than the hole mobility[5] for semiconductor materials, a given npn transistor tends to be *faster* than an equivalent pnp type (see paragraph 3.1.5 for details).

[5]Electrical mobility (or simpler mobility) is the ability of charged particles to move through a medium in response to an electric field that is pulling them; for electron and holes in $In_{0.47}Ga_{0.53}As$ it is 7000 and 600 $cm^2/(V*S)$ respectively.

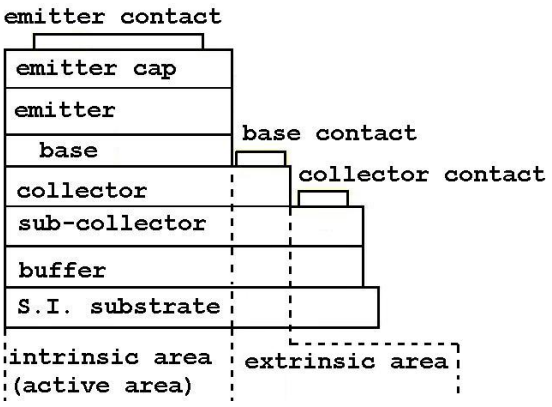


FIGURE 3.3: Ideal cross–section of an HBT. The emitter width and length (W_E and L_E) define the active area or *intrinsic*; extrinsic area is also highlighted.

Four combinations of biasing are available the HBTs: table 3.2 summarizes the operating conditions.

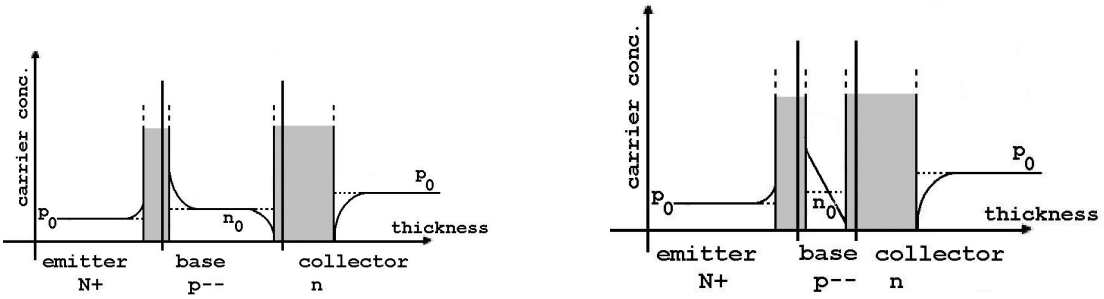
TABLE 3.2: Standard biasing regions for bipolar transistors.

$\mathbf{V}_{BE}$ V	$\mathbf{V}_{BC}$ V	**Region**
> 0	< 0	Forward–Active
> 0	> 0	Saturation
< 0	< 0	Cut–Off
< 0	> 0	Reverse

The concentration of the minority carriers in the base is the key element in the proper operation of an HBT.

In the forward–active biasing regions, the minority carriers (the electrons) concentration in the base is higher at the emitter side then at the collector side (here is nearly zero). According to the thickness of the base layer, the excess–electron concentration will decrease with a different rate from the emitter to the collector: in fig. 3.4(b) the base is made very thin by epitaxy, so this concentration decreases rapidly without saturing[6] as in fig. 3.4(a).

[6] Base–lenght should be one tenth of the diffusion lenght of the carriers to avoid them saturate within the base itself [82].



(a) Thick base minority carrier concentration

(b) Thin base minority carrier concentration

FIGURE 3.4: Minority carrier concentrantrion for an HBT in thick (a) and thin (b) base case. The proper operation of the devices relies on the impossibility for the electron to saturate (at the value n_0) in the base layer. Grey areas in the figures represent the depletion layers across the junctions.

The electron diffusion current in the base is obtained by differentiating the minority electron profile: since the latter is almost linear[7], the electron current will be constant (at the emitter side is the same as the collector side).

The base current is relatively small in a thin–base transistor because only a *thin* base allows the injected electrons from the emitter to diffuse quickly through the base without recombining. This physically means that the electrons flowing through the base–emitter interface will not exit the device by the base contact but they will be forced to the base–collector depletion layer where they will be *collected* by the collector terminal (hence the name).

3.1.3 Base current considerations

As already mentioned, the HBT represents an improvement of the bipolar transistor because, thanks to band–engineering, it maximize the electron flow and minimises the hole current in the same time.

[7]Strictly speaking the trend is exponential but in the approximation of *thin base* it can be considered linear.

The electrical set–up associated to the diagram represented in fig. 3.5 is called

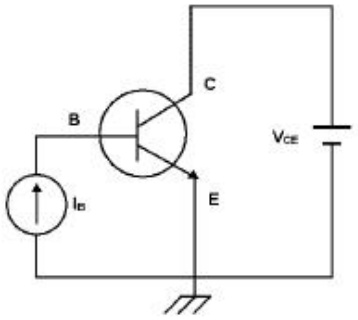


FIGURE 3.5: Common emitter configuration: while I_B is forced into the base, the potential across the collector and the emitter is varied. The I_C is then recorded in function of V_{CE}.

common–emitter configuration [80]; as a parameter to assess the quality of the operations of a bipolar transistor, the common–emitter current gain, β, is defined and used:

$$\beta = \frac{I_C}{I_B} \tag{3.3}$$

where β is the ratio between the collector current, I_C, and the current provided by the external base contact, I_B. For completeness it has to be noticed that the base current has five main contributors:

- the backward–injection current, I_B^*;
- the surface recombination current, $I_{B,surf}$;
- the interface recombination current, $I_{B,cont}$;
- the bulk recombination current in the base region, $I_{B,bulk}$;
- the space-charge recombination current in the base-emitter depletion region, $I_{B,scr}$.

$$I_B = I_B^* + I_{B,surf} + I_{B,cont} + I_{B,bulk} + I_{B,scr} \tag{3.4}$$

where the backward–injection current, I_B^* is the current of the holes that are back–injected from the base to the emitter; I_B^* is the main contributor of I_B. $I_{B,surf}$ is defined as the recombination of the minority carriers injected from emitter with the base majority carriers at the surfaces. This is proportional to the emitter periphery. Especially for small devices having large perimeter area ratio, this component becomes dominant: this effect is also named as *emitter size effect*[8]. $I_{B,cont}$ becomes important when the base–emitter contact spacing is low. In this case, the minority carriers flow also laterally to the base contact and recombine with the majority carriers and increase the $I_{B,cont}$. If the base–emitter contact spacing is large, the minority carrier concentration reaching the base contact is nearly zero. The recombination of the carriers in the depletion region results in increase of $I_{B,scr}$.

As mentioned before, the main element in eq. 3.4 is the backward–injection current, so it is possible to re–write eq. 3.3 [80] in:

$$\beta = \frac{I_C}{I_B^*} = \frac{N_E T_E DnB}{N_B T_B DpE} \cdot e^{\frac{\Delta E_V}{kT}} \tag{3.5}$$

where, N_E and N_B are the emitter and base doping levels; T_E and T_B are the emitter and base thickness; D_{pE} and D_{nB} are the minority hole diffusion coefficient in the emitter and the minority electron diffusion coefficient in the base, respectively. ΔE_V is valence band discontinuity at base–emitter interfaces.

[8]The dc current gain decreases with the decreasing emitter area.

3.1.4 HBT Small–Signal Equivalent Circuit

Last paragraph has shown how DC characteristics for an HBT depend only on the semiconductors used in the epitaxial layer structure. No geometrical considerations[9] are involved. On the other side AC performances are deeply dependant on such variables (see next par.). For this reason it is necessary to introduce the small–signal equivalent circuit (SSEC).

The SSEC is a method by which it is possible to represent a non–linear device by an equivalent linear network; this linearization is formed about a given DC bias point of the device and can be accurate only for small excursions about this point[10].

This *linearisation* process allows the non–linear device to be associated to a network of finite elements (such as capacitors, resistors and inductors) for the given bias point above–mentioned. Even though those finite elements are a mere representation of the reality, each of them can be connected to a physical parameter of the device[11]; once the SSEC model is known, it is possible to associate its elements to the electrical performances of the device: in this way there will be a link between the *physics* of the device and the relative electrical behaviour.

A common form for the small signal equivalent circuit for transistors is to use a two-port network to model the device: in this model, the three terminal transistor, considered as a two port network, must have one of its terminals in *common* to

[9] i.e., emitter contact dimensions, distance between the base contact and the emitter mesa, metal schemes for the contacts.

[10] Under this condition the non–linear behaviour of the device is approximated to a liner one.

[11] For exemple, there is no real resistance between the metal contact and the substrate but there is a real voltage–drop across the interface that can be reproduced by a so–called access resistance.

both ports. In the present case a common emitter configuration will be considered[12]. Figure 3.6 represent the small-signal equivalent circuit of an HBT [63] in

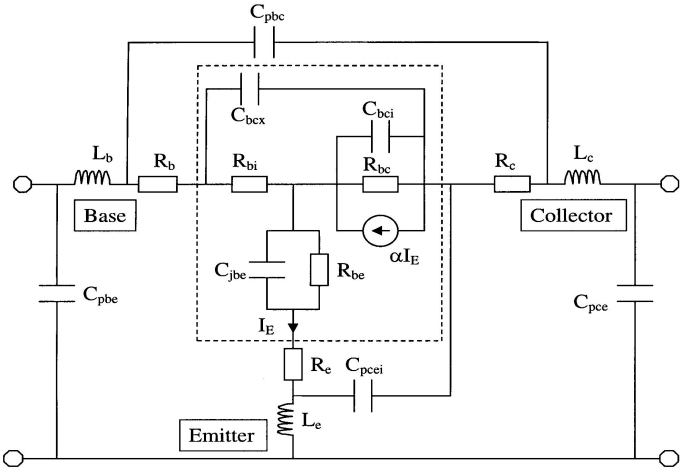


FIGURE 3.6: Small-signal equivalent circuit of an HBT; it is divided into two areas: the internal one, called *intrinsic*, whose origin is in the HBT itself and the external area, whose origin depends from the interconnections between the measurement system and the HBT.

common emitter configuration. A description of the single elements follows:

- Emitter access resistance, R_e; it includes:

  Emitter contact resistance, $R_{e,cont}$;

  Emitter bulk resistance, $R_{e,bulk}$;

- Base access resistance, R_b; it includes:

  Base contact resistance, $R_{b,cont}$;

  Base–emitter gap resistance, $R_{be,gap}$;

- Collector access resistance, R_c; it includes:

  Collector contact resistance, $R_{c,cont}$;

  Collector bulk spread resistance, $R_{c,spread}$;

[12] In this case the input 2–terminals port will be the b–e junction and the output will be taken across the collector and the emitter.

- Resistance of base–emitter junction (depletion layer), R_{be};
- Capacitance of base–emitter junction (depletion layer), C_{jbe};
- Intrinsic base resistance, R_{bi};
- Common–emitter gain, α;
- Emitter current, I_E;
- Resistance of intrinsic base–collector junction (depletion layer)R_{bc};
- Capacitance of intrinsic base–collector junction (depletion layer), C_{bci};
- Capacitance of extrinsic base–collector junction (depletion layer), C_{bcx}.

Figure 3.6 also includes:

- Parasitic inductances given by the connections: L_e, L_b, L_c;
- Parasitic capacitances given by the overlapping connections: C_{pbe}, C_{pce};
- Parasitic capacitances: C_{pcei}[13], C_{pbc}.

Before showing the connections of these parameters to the RF figures of merit, two formulas must be introduced [83]: total base resistance, R_{bb}, and total capacitance

[13]The value of C_{pcei} is negligeable with respect of the other parasitics so it will not be taken in account for future calculations.

of the Base–Collector junction, C_{bc}.

$$R_{bb} = R_{b,cont} + R_{be,gap} + R_{bi} \tag{3.6a}$$

$$R_{b,cont} = \frac{\sqrt{R_{sh} \cdot \rho_C}}{2 \cdot L_E} \tag{3.6b}$$

$$R_{be,gap} = \frac{\rho \cdot W_{BE}}{2 \cdot L_E} \tag{3.6c}$$

$$R_{bi} = \frac{\rho \cdot W_E}{12 \cdot L_E} \tag{3.6d}$$

where R_{sh} is the sheet resistance of the substrate below the contact, ρ_C is the specific contact resistance for the metal scheme onto the base, L_E is the emitter length (which defines the base–length too), W_E is the emitter width and W_{BE} is the distance between the emitter and the base (generally called *gap*).

$$C_{bc} = C_{bci} + C_{bcx} \tag{3.7}$$

where the total base–collector junction capacitance[14], C_{bc}, is given by the sum of *intrinsic* base–collector junction capacitance, C_{bci}, and of *extrinsic* base–collector junction capacitance, C_{bcx}.

3.1.5 RF ideal performances

There are two important figures of merit for the high frequency characterization of HBTs': current gain cut–off frequency f_T, and the maximum oscillation frequency f_{max}. The current–gain cut–off frequency is the frequency at which the AC current–gain becomes unity, while the maximum oscillation frequency is the

[14]Given the particular shape of a step pyramid for an HBT, the base–collector junction is larger than the base–emitter junction; since the active area of the device is the one beneath the emitter contact, only a part of the base–collector junction will be part of the device (hence the name *intrinsic).*

frequency at which the *unilateral* power gain of the transistor rolls off to unity. The current–gain cut–off frequency is determined by the transit time of the electrons from emitter to collector [83] and is given by:

$$\frac{1}{2 \cdot \pi \cdot f_T} = \tau_E + \tau_B + \tau_{SC} + \tau_C \tag{3.8}$$

that can be simplified as

$$f_T = \frac{1}{2 \cdot \pi \cdot (\tau_E + \tau_B + \tau_{SC} + \tau_C)} \tag{3.9}$$

where

- τ_E is the emitter charging time;
- τ_B is the base transit time;
- τ_{SC} is the space–charge transit time;
- τ_C is the collector charging time.

The following equations associate each of these parameters to the relative elements in the SSEC (see fig. 3.6):

$$\tau_E = \frac{kT}{qI_C} \cdot (C_{be} + C_{bci}) \tag{3.10a}$$

$$\tau_B = \frac{T_B^2}{2D_{nB}} \tag{3.10b}$$

$$\tau_{SC} = \frac{T_{dep}}{2\nu_{sat}} \tag{3.10c}$$

$$\tau_C = (R_e + R_c) \cdot C_{bci} \tag{3.10d}$$

where kT/q is the thermal voltage, T_B and T_{dep} are the thickness of base and depletion region on base–collector junction, D_{nB} is the electron diffusion coefficient in the base, ν_{sat} is the saturation velocity of carriers in the base–collector depletion region. Considering equations 3.10, equation 3.9 is written as:

$$f_T = \frac{1}{2 \cdot \pi \cdot \left[\frac{kT}{qI_C} \cdot (C_{be} + C_{bci})\right] + \left[\frac{T_B^2}{2D_{nB}}\right] + \left[\frac{T_{dep}}{2\nu_{sat}}\right] + [(R_e + E_c) \cdot C_{bci}]} \tag{3.11}$$

The equation shows now how the f_T is influenced mainly by the vertical design of HBTs.

Another important figure of merit for high frequency performance is the maximum oscillation frequency f_{max}: it is defined as the frequency where the power gain drops to unity:

$$f_{max} = \sqrt{\frac{f_T}{8\pi R_{bb} C_{bc}}} \tag{3.12}$$

3.2 Layer Structure for fabricated InP HBTs

The epilayer structure of the wafers used in the project was grown from InP substrates using MBE. Table 3.3 shows the layer structure used in the fabrication of Npn InP/InGaAs single heterojunction bipolar transistor.

TABLE 3.3: Layer structure in use for InP–based heterojunction bipolar transistors.

	Material	**Thickness** nm	**Doping Type**	**Doping Concentration** cm^{-3}
Emitter cap	$In_{0.53}Ga_{0.47}As$	40	n^{++}	Si: $3 \cdot 10^{19}$
Emitter cap	InP	80	N^{++}	Si: $3 \cdot 10^{19}$
Emitter	InP	10	N^{+}	Si: $8 \cdot 10^{17}$
Emitter	InP	40	N^{+}	Si: $3 \cdot 10^{17}$
Base	$In_{0.53}Ga_{0.47}As$	40	p^{--}	C: $6 \cdot 10^{19}$
Collector	$In_{0.53}Ga_{0.47}As$	400	n	Si: $2 \cdot 10^{16}$
Etch Stop	InP	10	N^{++}	Si: $1 \cdot 10^{19}$
Sub-collector	$In_{0.53}Ga_{0.47}As$	200	n^{++}	Si: $3 \cdot 10^{19}$
Buffer	InP	200	N^{++}	Si: $2 \cdot 10^{19}$
SI Substrate	InP	-	-	-

The emitter cap is the topmost layer; because of the smaller band-gap of InGaAs it is appropriate to make very good ohmic contact. A second, less doped, cap layer is necessary to reduce the potential barrier step between the first cap and the upper layer of the emitter: the main consequence is to reduce the emitter access resistance. Like the cap, the emitter level is split into two sections: the upper one is made of highly doped InP to provide large numbers of electrons for injection into the base. The lower layer of InP is less doped to reduce charge storage which would lead to higher intrinsic base-emitter junction capacitances. Such a difference in doping concentration narrows band-gaps and reduces (slightly) the potential barrier between the two InP layers.

Between the emitter and base a very thin layer of undoped InGaAs is generally grown: it is called *spacer layer* and should reduce the possibility of diffusion of the base dopant into the emitter region. Because of the high temperature reached during the whole process (from the growth of the epistructure to the annealing of the metal contacts), some of the p-type dopant could migrate from the base to the emitter, filling the latter of holes[15]. This occurance has been sensibly reduced by using carbon as p–dopant for the base; so in the present project, the so–called spacer has not been used[16].

The base layer is the most critical layer in HBTs. This layer defines the current gain of the transistor (eq. 3.5). Its thickness should be chosen as thin as possible to prevent excessive recombination and loss of electrons trying to reach the

[15] It would make void all the efforts for an highly doped emitter

[16] Berillium and Zinc have higher diffusivity than Carbon; they require a space layer of several hundreds of Angström (300 $\AA$ in [84, 85])

collector[17]. On the other hand, an extremely thin base would lead to two major issues: the two depletion layers could overlap (*punch–through*) and the base resistance would increase (generally resistance is inversly proportional to the thickness). Moreover, the thinner the base the shorter the transit time is: this leads to higher cut-off frequency (eq. 3.10b). In order to mantain a very thin base and to avoid the overlapping of the depletion layers, the base should be doped as highly as possible[18].

The upper collector layer is made of lightly doped InGaAs as there is no requirement for high carrier concentration. There is a reason why high concentration of carriers is not desirable: the collector region acts also as the *dielectric* in the capacitor formed between the base metal contacts and the highly doped sub-collector. A lightly doped collector means that the collector is fully depleted under normal operation and so the base collector capacitance is minimised. The subcollector layer is highly doped to facilitate the formation of ohmic contacts (as for the emitter cap). To improve the quality of the active device layers grown on the InP substrate a 200nm buffer layer of InP is grown using the same growth conditions as the uppermost layers.

3.2.1 Fabrication Process

The present section deals with the specific process steps as they were generally explained in section 2.2.

[17] Theoretically the base layer thickness should be less than a tenth of the diffusion length of minority carriers in the base.

[18] The depletion layer thickness depends on the doping levels of both sides, p– and n–: $N_A \cdot W_p = N_D \cdot W_n$ where N_A and N_D are respectively the concentrations of acceptors and donors in the doped semiconductors forming the junction, while W_p and W_n are the extensions of the depletion region within the p– and n– doped semiconductors respectively.

TABLE 3.4: HBT fabrication steps: optical lithography is the process in use.

Step	Step name	Procedure	Notes
01	Emitter metal	Metal deposition / lift off	Ti/Pd/Au
02	Emitter mesa	Wet etch	Layers 1-4
03	Base metal	Metal deposition / lift off	Ti/Pd/Au
04	Base mesa	Wet etch	Layers 5-7
05	Collector metal	Metal deposition / lift off	Ti/Pd/Au
06	Collector mesa	Wet etch	Layers 8-9
07	Insulator	Deposition and w.e.	Polyimide
08	Bond-pads	Metal deposition / lift off	Ti/Au

HBTs are generally named after their emitter size, so the following devices were fabricated: 6x10 μm^2, 10x15 μm^2, 20x30 μm^2; table 3.4 provides an overview of the fabrication process.

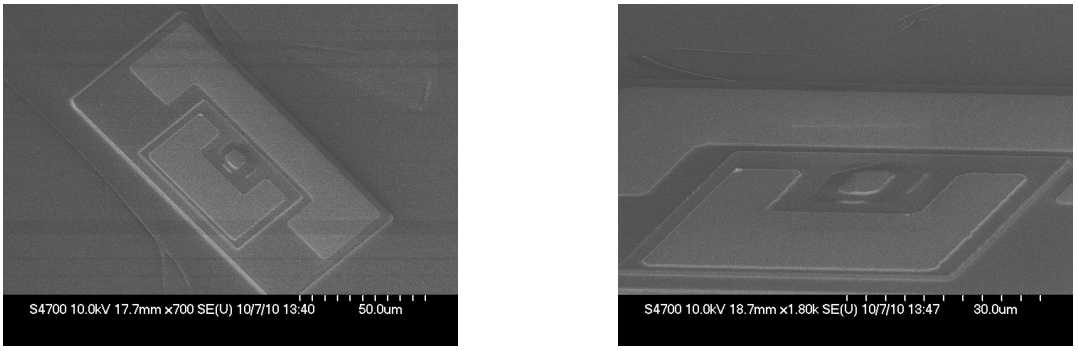
The first stage of fabrication was to define the emitter metal layer using photolithography. As already stated in par. 2.3.1, the Ti/Pd/Au metal scheme[19] was chosen for all the contacts (n– and p–doped); alignment markers were deposited along with the topmost contact of the HBTs. The emitter mesa structure was then defined. As this step required etching, two alternative ways were available:

- Use the metal contact itself as protective mask (recommended for wet–etch, it does not suit dry–etch);
- Spin (and pattern) a layer of photoresist to act as mask (recommended for dry–etch, it actually suits both dry– and wet–etch).

For emitter mesa formation the former method was chosen: even though wet–etch does not provide vertical sidewall (as in the dry–process, see fig. 2.5), the relative undercut is negligeable with respect of the emitter area[20].

[19] The exact scheme is a blanket deposition of 20nm of titanium, 30nm of palladium and 80nm of gold.

[20] The total emitter vertical depth is 170nm; it means that in worst case scenario (isotropic process) the 6*10 μm^2 emitter will become 5.7*9.7 μm^2 which corresponds to the 92% of the designed area.



(a) Top view

(b) Later view

FIGURE 3.7: SEM snapshots of 6x10 μm^2 emitter area HBT device

The next step was to define the base metal pattern. After that the metal scheme (Ti/Pd/Au) could be evaporated and then the base mesa was defined. For the base mesa formation the relative areas were covered in photoresist and then wet–etched. The collector metal deposition and mesa formation were carried out as for the base. Figures 3.7 show samples of the HBTs after collector mesa: after that stage an insulating layer of Polyimide and the bond–pads[21] were deposited onto the devices. Figure 3.8 shows how finished devices are. The full fabrication process is given in the Appendix A.

3.3 DC chacterization

After completing the fabrication process, all devices were tested for DC characteristics; measurements were carried out with Cascade Microtech M150 probe station by the means of the B1500A semiconductor device analyzer. Devices under test (DUTs) were fabricated in common-emitter configuration (fig. 3.8 and 3.9): in such a configuration (and for a given bias voltage, V_{CE}) the transistor behaves as

[21] Metal scheme: Ti/Au, 50nm of titanium and 150nm of gold.

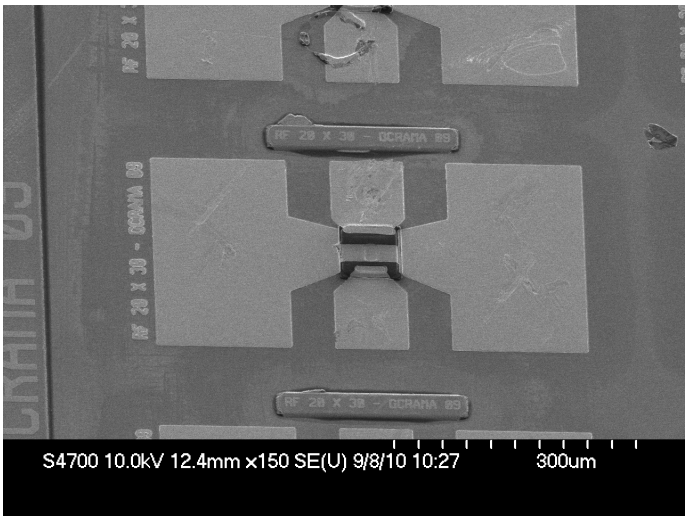
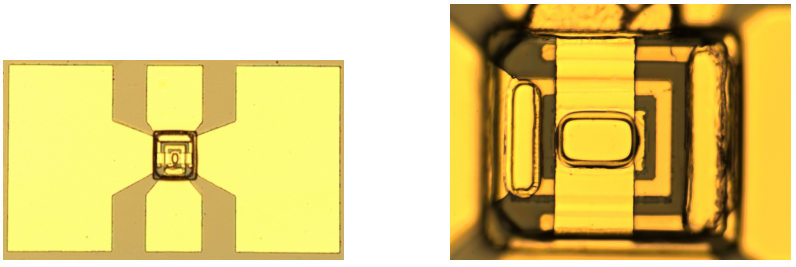


FIGURE 3.8: Completely fabricated sample: 20x30μm^2 HBT. It has to be noticed: the mesa structure in the centre (it is black because the dielectric that is used to insulate ICs is opaque in Scanning Electron Microscpes), the GSG–CPW for connections (base at the bottom, collector at the upper).

a current–controlled current–generator (I_C=f(I_B)). The current gain, β, has been already defined in equation 3.3.

For a proper DC analysis of an HBT, the common emitter I–V characteristics and the Gummel Plots are required; following paragraphs will show these methods and the data acquired on the fabricated samples.



(a) Fully fabricated HBT in common emitter configuration.

(b) Magnification of the HBT as in fig (a).

FIGURE 3.9: Completed fabrication of a 20x30 μm^2 device.

3.3.1 Common Emitter Current–Voltage Characteristic

The current–voltage characteristic, taken from the devices in common emitter configuration, is shown in fig. 3.10. Forcing a given current I_B into the base, the collector current, I_C, is measured in function of the collector voltage (V_{CE}). This operation is iterated for several different values of I_B (from 0 to 500 μA).

The I_C vs V_{CE} plot shows the expected behaviour: while $V_{CE} \leq 0.45V$ (linear region), the current gain is directly proportional to V_{CE} and to I_B, in the forward active region ($V_{CE} \geq 0.45V$) the current gain is constant with respect of V_{CE} and proportional only to I_B.

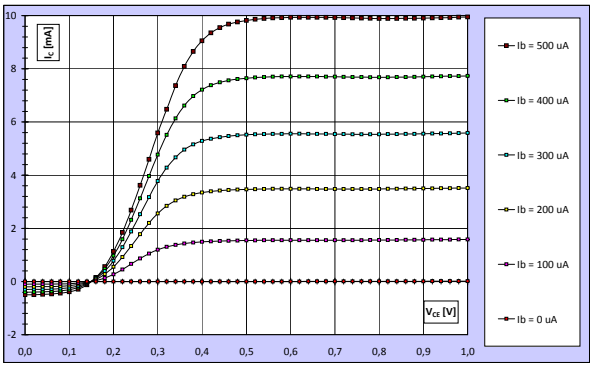


FIGURE 3.10: I-V Characteristic for an HBT in common emitter configuration.

It is possible to extract the current gain β from eq. 3.3: for $V_{CE} \geq 0.45V$, a maximum value of $\beta = 20$ was obtained.

3.3.2 Gummel Plots

To evaluate the behaviour of the devices, Gummel Plots have been taken. Gummel plot is a plot of the I_B and I_C in function of V_{BE} when V_{BC} is kept at 0 V (fig. 3.11).

In this way, each junction is tested while the whole transistor is operative; if a

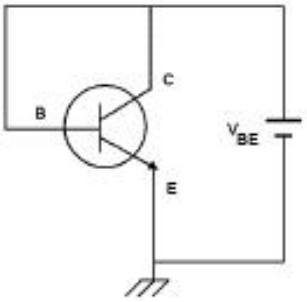


FIGURE 3.11: Diagram of a npn transistor in standard configuration for Gummel Plot. While $V_{BE} = V_{BE}$, The bias is varied and the two currents I_B and I_C are recorded.

simple two-probes I–V measure was performed separately on the two junctions it wouldn't have been possible to highlight all the current contributions that turn on in a transistor[22]. For example, the space-charge recombination is a negligible effect in the diode, while it is relevant in the transistor.

Along with Gummel Plot, it is necessary to introduce the concept of the ideality factor. The ideality factor is a fundamental parameter of the matematical expressions of I_B and I_C (equations 3.13). It can be used as a tool to assess the quality

[22]When a junction is tested for I–V characteristic, while the other terminal is left floating, it will be tested as a simple p–n diode.

of the HBT.

$$I_C = I_S \cdot e^{\frac{qV_{BE}}{\eta \cdot kT}} \tag{3.13a}$$

$$I_B = \frac{I_S}{\beta} \cdot \left(e^{\frac{qV_{BE}}{\eta \cdot kT}} - 1 \right) \tag{3.13b}$$

where I_S is the transport saturation current. The value of η can be extracted by exponential interpolation of the measured data. The expected value for the I_C and I_B ideality factor is 1. According to the influence of its specific component, measured base current (eq. 3.4) might have an higher ideality factor. The following table lists the expected values of the ideality factor for each components of I_B with relative connections to devices' geometry.

TABLE 3.5: Expected values of the ideality factor for each components of I_B.

Current Component	**Expected Ideality Factor**	**Proportinality**	**Relevant for**
I_B^*	1	Area	Abrupt BE junc.
$I_{B,surf}$	1	Perimeter	Small geometry
$I_{B,cont}$	1	Perimeter	Self-aligned BE cont.
$I_{B,bulk}$	1	Area	Heavy base doping
$I_{B,scr}$	2	Area	Large numb. traps

Figure 3.12 is an example of a Gummel Plot extracted from the fabricated devices: the red line, I_C, shows an ideality factor of 1.18, very close to the ideal case ($\eta = 1$); the yellow line, I_B, shows an η close to the expected value ($\eta > 1$).

It is possible to identify the three main regions:

- Cut-off [0 300 mV];
- Linear [300mv 600mV];

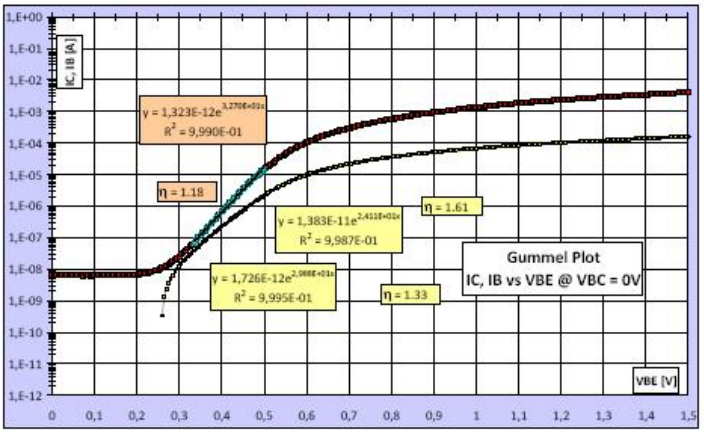


FIGURE 3.12: Example of Gummel plot measured from a fabricated HBT: red line is the collector current, I_C, yellow line is base current, I_B.

- Forward active [over 600mV].

In the plot, I_B follows the expected behaviour only in the second half ($V_{BE} \geq$ 0.45V). For $V_{BE} \leq 0.45$V a very high ideality factor is shown. This suggest that the $I_{B,scr}$ dominates the other components of I_B (see eq. 3.4). The ratio between the I_C and I_B in forward active regime can be used to evaluate the current gain, β obtained previuosly: in this is case a magnitude of 20 has been confirmed.

3.4 RF Characterization

RF characterization is performed by measuring S–parameters[23]: it was carried out on the probe station Cascade Microtech M150 by the means of the Two Port E8361 PNA network analyzer and the B1500A semiconductor device analyzer. S-parameters can be used to describe the electrical behavior of linear electrical networks when undergoing various steady state-stimuli by electrical signals.

[23] The S stands for scattering.

Before specifying the properties and the use of S-parameters, it is important to define a related subject: the two–port network. A two–port network is an electrical circuit[24] with two pairs of terminals to connect to external circuits: signals will be applied to this *black box* and relative outputs will be recorded; a matrix will be then created to match the inputs and the outputs. This allows the response of the network to signals applied to the ports to be calculated easily, without solving all the internal voltages and currents in the network. According to possible combinations of input/output signals, several models of matrices are available: *scattering* matrix, *impedance* m., *admittance* m., *hybrid* m. are the most used; a brief description of those models follows in the next paragraphs.

3.4.1 S-parameters

For the case in use[25], 2–port S–parameters will be considered. In a 2–port network, often port 1 is considered the input port and port 2 is considered the output port. Figure 3.13 represent a standard 2–port network for S-parameters extraction: $\mathbf{a}_n$ and $\mathbf{b}_n$ represent the incident and reflected *power waves* respectively[26].

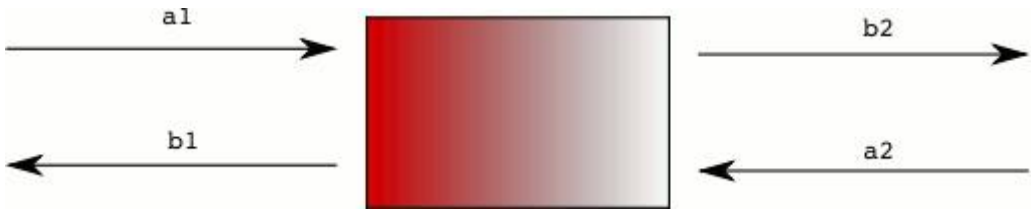


FIGURE 3.13: Ideal rapresentation of a 2-Port network for S-parameters extraction. Incident and reflected power waves, a_n and b_n, are highlighted.

[24] It can be a single device or a complex network of devices.
[25] As mentioned, DUTs were fabricated in common emitter configuration.
[26] The indeces $_1$ and $_2$ denote the port where the respective waves are applied

At the test frequency each element of the S–matrix is represented by a unitless complex number (magnitude and angle or amplitude and phase) as a function of the frequency: expression (3.14) is the matrix that describes the system.

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{3.14}$$

that can be expanded into:

$$b_1 = S_{11}a_1 + S_{12}a_2 \tag{3.15a}$$

$$b_2 = S_{21}a_1 + S_{22}a_2 \tag{3.15b}$$

If one considers an incident power wave at port 1 (a_1) it may result in an exiting wave from either port 1 itself (b_1) or port 2 (b_2). If port 2 is terminated in a load identical to the system impedance (Z_0, in the case 50 Ω) then it will be totally absorbed making a_2 equal to zero. Therefore:

$$S_{11} = \frac{b_1}{a_1} \tag{3.16a}$$

$$S_{21} = \frac{b_2}{a_1} \tag{3.16b}$$

If port 1 is terminated in the same fashion, the a_1 becomes zero, giving:

$$S_{12} = \frac{b_1}{a_2} \tag{3.17a}$$

$$S_{22} = \frac{b_2}{a_2} \tag{3.17b}$$

where: S_{11} is the *input port voltage reflection coefficient*, S_{12} is the *reverse voltage*

gain, S_{21} is the *forward voltage gain*, S_{22} is the *output port voltage reflection coefficient*.

3.4.2 Z–, Y–, h– parameters

The S-matrix can be used as a starting point to extract other 2–ports models that can highlight different electrical properties of the network in the study:

- *Impedance* matrix, or *Z–matrix*;
- *Admittance* matrix, or *Y–matrix*;
- *Hybrid* matrix, or *h–matrix*.

In the first case, Z–matrix transforms the input–output currents vector, $\mathbf{I} = (I_1; I_2)$, into the input–output voltages vector, $\mathbf{V} = (V_1; V_2)$. The elements of the matrix will have the dimension of an impedance[27].

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \tag{3.18}$$

[27] Impedance, Z, is a complex quantity used to describe Ohm's Law for AC domain. It is possible to write $Z = R + jX$ where R is the resistance (DC domain) and X is the reactance. Z, R and X are expressed in Ohms (Ω)

that provides:

$$Z_{11} = \frac{V_1}{I_1} \qquad \text{for } I_2 = 0 \tag{3.19a}$$

$$Z_{21} = \frac{V_2}{I_1} \qquad \text{for } I_1 = 0 \tag{3.19b}$$

$$Z_{12} = \frac{V_1}{I_2} \qquad \text{for } I_2 = 0 \tag{3.19c}$$

$$Z_{22} = \frac{V_2}{I_2} \qquad \text{for } I_1 = 0 \tag{3.19d}$$

In the same fashion it is possible to get the admittance matrix: it connects the input–output voltages vector $\mathbf{V} = (V_1; V_2)$ to the input–output currents vector $\mathbf{I} = (I_1; I_2)$. The elements of the matrix will have the dimension of an admittance[28].

$$\begin{bmatrix} I_1 \\ \\ I_2 \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{12} \\ \\ Y_{21} & Y_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ \\ V_2 \end{bmatrix} \tag{3.20}$$

that provides:

$$Y_{11} = \frac{I_1}{V_1} \qquad \text{for } V_2 = 0 \tag{3.21a}$$

$$Y_{21} = \frac{I_2}{V_1} \qquad \text{for } V_2 = 0 \tag{3.21b}$$

$$Y_{12} = \frac{I_1}{V_2} \qquad \text{for } V_1 = 0 \tag{3.21c}$$

$$Y_{22} = \frac{I_2}{V_2} \qquad \text{for } V_1 = 0 \tag{3.21d}$$

The hybrid matrix is generally used for current amplification measurement; it is named after the hybrid nature of the vectorial quantities of the two vectors that

[28] Admittance, Y, is a complex quantity. It is generally defined as $Y = Z^{-1}$, it is also possible to write $Y = G + jB$ where G is the conductance (DC domain, $G = R^{-1}$) and B is the susceptance. Y, G and B are expressed in Siemens or Ohms^{-1} (S or Ω^{-1})

this matrix relates[29]: $(I_1; V_2)$ and $(I_2; V_1)$.

$$\begin{bmatrix} V_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ V_2 \end{bmatrix} \tag{3.22}$$

that provides:

$$h_{11} = \frac{V_1}{I_1} \qquad \text{for } V_2 = 0 \tag{3.23a}$$

$$h_{21} = \frac{I_2}{I_1} \qquad \text{for } V_2 = 0 \tag{3.23b}$$

$$h_{12} = \frac{V_1}{V_2} \qquad \text{for } I_1 = 0 \tag{3.23c}$$

$$h_{22} = \frac{I_2}{V_2} \qquad \text{for } I_1 = 0 \tag{3.23d}$$

All the aforementioned matrices are interchangeable: starting from the measured set of S–parameters (eq. 3.14) it is possible to obtain the others[30].

3.4.3 Calibration

Before each measurement session, it is necessary to calibrate the apparatus: this is to remove the effect of connecting cables, define the measurement reference plane to the probe tips and remove measurement errors from, for instance, coupling between the ports. The calibration of the network analyzer improves the accuracy and repeatability of measurements. This process involves the measurement of known standard parameters and using the data to compensate for systematic errors. After making these measurements, the network analyzer computes some

[29] It is called hybrid because that matrix gathers figures with different physical dimension(i.e., Ohms, Siemens, unitless).

[30] All the CAD applications (Agilent's ADS in the project) provides specific tools for automated conversion from–to all the possible combinations of S–, Z–, Y–, h–parameters.

correction values to produce the expected answer[31]. Calibrations can be simple (such as compensating for transmission line length) or can involve methods that compensate for losses, mismatches, and feedthroughs.

Some of the factors that contribute to measurement errors are repeatable and predictable over time and temperature, and can be removed, while others are random and cannot be removed:

- **Systematic Error**
  - Directivity and crosstalk (relating to signal leakage)
  - Source and load impedance mismatch (relating to reflections)
  - Frequency response errors (caused by reflection)
  - Transmission tracking within the test receivers
- **Random Error**
  - Instrument noise
  - Switch repeatability
  - Connector repeatability
- **Drift Error**
  - Enviromental variation (temperature)

Figure 3.14 shows the assumed general model for the effects of the systematic errors. The matrix $[S^m]$ will define the S–parameters actually measured by the VNA

[31] For answers that are supposed to be zero, the analyzer can subtract the residual. For non-zero values, the analyzer could calculate complex factors that will compensate for both phase and amplitude errors.

during each single run of measurement. These include all the above mentioned errors and the DUT properties.

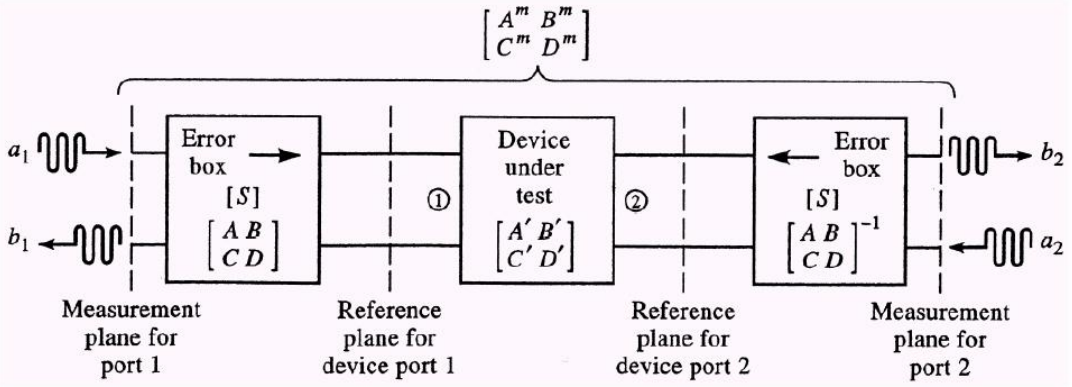


FIGURE 3.14: General Model which embeds the DUT + the removeable systematic errors.

The two error boxes (one to the left and one to the right) and the relative S-parameters gather all the possible systematic errors while the S-parameters of the DUT lie within the matrix $[S']$ (along with the non-removeable random errors). The purpose ot the network analyzer calibration is to determine the numerical values of $[S]$ in the error model at each frequency of interest.

There are two basic types of calibration used to correct the systematic error: SOLT and TRL[32]. The differences in the calibrations are related to the types of calibration standards they use and how the standards are defined. They each have their advantages, depending on frequency range and application. In this case the SOLT calibration method was chosen. SOLT calibration is easy to perform, and is used in a broad variety of environments and it is the most widely used choice for coaxial

[32]The names stand respectively for Short Open Load Through and Through, Reflect Load.

measurements [86]. It can also be used with fixtures and probes. SOLT inherently provides a broadband calibration, essentially from DC to the upper frequency limit of the connector type being used[33].

As mentioned before, the calibration consists in the measurement of the response of given standard substrate. For the equipment in use, the impedance standard substrate (ISS) provided by Cascade Microtech was used. The ISS–SOLT calibration is based on the following calibration standards:

- *short* circuit, where a vertical metallized line shorts the three tips of the RF probe together;
- *open* circuit, where the probes are elevated above the substrate (h $\approx$ 200 μm);
- *load* structure which is matched to the 50Ω charcteristic impedance of the system;
- *thru* structure which is essentially a 50Ω short line connecting the two probes by a line of a given length.

3.4.4 Extraction of the small–signal equivalent circuit elements

Accurate small-signal modelling is a crucial part in the process and development of HBTs to allow for improved device evaluation and modelling [63, 87, 88, 89, 90, 91, 92, 93]. Early trends in this field tended towards the numerical optimization of the S-parameters extracted from the model to let them match the S-parameters

[33]The highest measured frequency was 67 or 110 GHz, according to the available instruments

measured from the real device. The consequence of this approach however, was to extract non–realistic circuit values. Since the 90s, direct extraction techniques have prevailed over optimization [63, 87, 89, 90, 91, 92, 93].

Direct extraction is a simple and fast–to–realise technique where few measurements on the device are sufficient to extract the circuit element values for the small–signal model. The small–signal model shown in figure 3.6 is the standard T–model commonly employed for HBTs [63].

The T–topology is directly related to the physics of the device and allows for checking of the consistency of the extracted parameters: hence it is not only useful for circuit design, but also for device optimization and technology development.

It is possible to identify two major parts of the model: the *external* and the *intrinsic*. The external part is composed of parasitic pad capacitances C_{pce}, C_{pbe}, C_{pbc}, inductances L_c, L_b, L_e, and transistor access resistances, R_c, R_b, R_e. The intrinsic part is composed of capacitances C_{be}, C_{bcx}, C_{bci}, resistances, R_{be}, R_{bi}, R_{bc} and the current gain α.

The parasitic resistances are commonly determined by open-collector measurements [63, 91, 92, 93]. In this way, the base-emitter junction is heavily forward-biased while the collector current is kept at zero. At high base current densities the influence of the pad capacitance is negligible. This method, besides requiring an extra measurement, is as accurate as the assumption placed on the intrinsic device under the high base current density.

Pad inductances and capacitances are commonly determined from the RF pad test-fixture [63, 92]. Besides requiring an additional dedicated structure and extra

measurements, this method is inaccurate for the following reasons:

- uncertainty in probe placement on the test structure;
- uncertainty about reproducibility of the device;
- differences with the layout of device.

Any combination of these differences will introduce errors in the extraction process.

3.4.5 The *new* extraction method for the Small–Signal Equivalent Circuit.

An accurate small-signal modelling technique is crucial to the steadfast development and improvement of the semiconductor devices: it will indeed provide valuable feedback for process optimisation. It is also essential for reliable circuit design. This paragraph presents the extraction of a physically realistic small–signal equivalent circuit for in–house fabricated HBTs; the small–signal model has been already presented in fig. 3.6: this is a physically based small–signal equivalent circuit model[34]. It is possible to separate two main areas in the small–signal model: intrinsic and external. There are nine *lumped* elements in the external area: three capacitors, three inductors and three resistors as well. All those elements are bias–independent and this property will be taken in the due account. The other intrinsic parameters include: three capacitors, three resistors and a current generator; those elements are bias dependent.

The new extraction approach requires an intimate knowledge of the device layout: pad capacitances and inductances are estimated using 3D electromagnetic

[34]Extensive details has been given in paragraph 3.1.4

numerical simulations. Test structures (see figure 3.15(a) and 3.16(a)) are actually simulated in Agilent's Momentum[35] application that includes information about the epilayer geometry and relative dielectric constants.

It possible then to estimate the magnitude of the metal–semiconductor contacts with straight–forward consideration on the geometry: photolithography allows, infact, a precise control over the dimensions of the metal deposited.

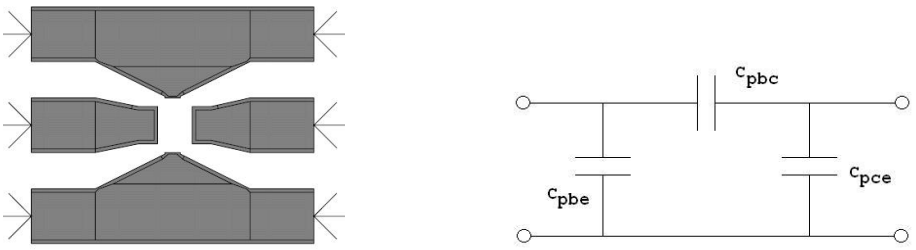
Knowledge of the estimated values of the external elements allows all the intrinsic elements to be estimated analytically. With good estimates of all the elements of the small–signal model, optimisation of these will follow to determine the actual element values. This approach is based on the fact that in multivariable optimisation, a starting vector close to the actual solution leads to quick convergence and determination of the correct optimised solution [95].

3.4.5.1 The parasitic capacitances

The parasitic capacitances C_{pbe}, C_{pce} and C_{pbc} are initially estimated from the test structure (fig. 3.15(a)) used to connect the co-planar waveguide (CPW) probes to the device: the test structure is an open pad structure that is modelled by a π–model comprising the parasitic capacitances C_{pbe}, C_{pbc} and C_{pce} (fig. 3.15(b)).

The simulations produce S–parameters which can be transformed into Y–parameters

[35] Momentum is a 3D electromagnetic simulator within Agilent's Advanced Design System (ADS) suit used for passive circuit modelling and analysis. It uses a technique called method of moments to solve Maxwell's electromagnetic equations for planar structures embedded in a multilayered dielectric substrate [94].



(a) Momentum Screenshot.

(b) Equivalent circuit.

FIGURE 3.15: Simulated test structure for parasitic capacitance estimations: (a) is a screenshot for Agilent Moment, (b) is the equivalent circuit for the simulated structure.

for analysis. This operation is necessary because the latter allows a straight analysis of the schematic in fig. 3.15(b):

$$Y_{network} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} = \begin{bmatrix} Y_{pbe} + Y_{pbc} & -Y_{pbc} \\ -Y_{pbc} & Y_{pce} + Y_{pbc} \end{bmatrix} \tag{3.24}$$

The magnitude of the single elements is hence determined by:

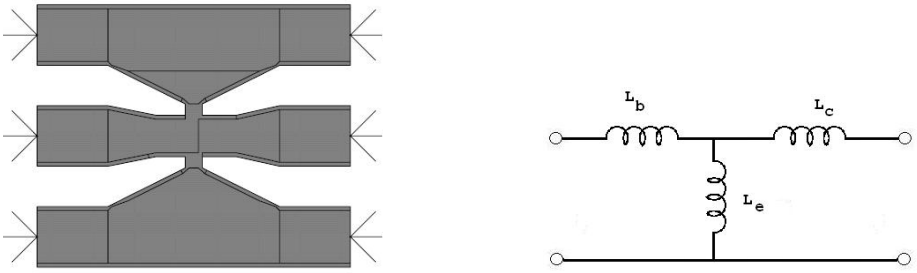
$$Y_{pbc} = -Y_{12} \tag{3.25a}$$

$$Y_{pbe} = Y_{11} + Y_{12} \tag{3.25b}$$

$$Y_{pce} = Y_{21} + Y_{22} \tag{3.25c}$$

As mentioned, these parameters are expressed in magnitude and angle (or amplitude and phase) in function of the test frequency; given the definition of admittance for a capacitor[36], the actual magnitude of the capacitors can extracted in this way:

[36]The admittance Y of a capacitor C is $Y = j \cdot \omega C$



(a) Momentum Screenshot.

(b) Equivalent circuit.

FIGURE 3.16: Simulated test structure for parasitic inductors estimations: (a) is a screenshot for Agilent Moment, (b) is the equivalent circuit for the simulated structure.

$$C_{pbc} = \frac{1}{Im(Y_{pbc}) \cdot \omega} \tag{3.26a}$$

$$C_{pbe} = \frac{1}{Im(Y_{pbe}) \cdot \omega} \tag{3.26b}$$

$$C_{pce} = \frac{1}{Im(Y_{pce}) \cdot \omega} \tag{3.26c}$$

The extracted figures are listed in table 3.6: they will be used as starting data for optimisation.

3.4.5.2 The parasitic inductances

After parasitic capacitors, lead inductances L_b, L_c, L_e are then estimated by a similar method, except that a short circuit is placed where the device would sit (fig. 3.16(a)).

The values of the inductances of fig. 3.16(b) are then estimated by using the extracted Z-parameters.

$$Z_{network} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} = \begin{bmatrix} Z_b + Z_e & Z_e \\ Z_e & Z_c + Z_e \end{bmatrix} \tag{3.27}$$

The parasitic inductances L_b, L_e and L_c are:

$$Z_e = Z_{12} \tag{3.28a}$$

$$Z_b = Z_{11} - Z_{12} \tag{3.28b}$$

$$Z_c = Z_{22} - Z_{21} \tag{3.28c}$$

Given the definition of impedance for an inductor[37], the actual magnitude of the inductors can extracted in this way:

$$L_e = \frac{Im(Z_e)}{\omega} \tag{3.29a}$$

$$L_b = \frac{Im(Z_b)}{\omega} \tag{3.29b}$$

$$L_c = \frac{Im(Z_c)}{\omega} \tag{3.29c}$$

The extracted figures are listed in table 3.6: they will be used as starting data for optimisation.

3.4.5.3 The parasitic resistances

The access resistances to the base, emitter and collector are estimated gathering the knowledges of HBTs layout, specific contact resistance (ρ_c) and sheet resistance (R_{sh}). These data are extracted by TLM measurements (table **??**).

[37] The impedance Z of an inductor L is $L = j \cdot \omega L$

The emitter contact resistances (R_E) can be initially estimated from:

$$R_E = \frac{\rho_c}{W_E \cdot L_E} \tag{3.30}$$

where W_E and L_E are the width and the length of the emitter, respectively, while the base and collector resistances can be estimated from:

$$R_{TOT} = R_{CNT} + R_{GAP} \tag{3.31}$$

$$R_{CNT} = \sqrt{\frac{\rho_c \cdot R_{sh}}{W_E}} \tag{3.32}$$

$$R_{GAP} = \frac{R_{sh} \cdot d}{L_E} \tag{3.33}$$

where R_{TOT} is the total value for the base or collector resistance (R_b and R_c in fig. 3.6), R_{CNT} is the metal/substrate contact resistance, d is the distance between the pad edges and the active (intrinsic) region of the device, and R_{gap} is the resistance of this portion of substrate. Table 3.6 gives the starting values for the estimated parasitic elements for a 6x10 μm^2 device.

3.4.5.4 The intrinsic model's parameters

Once all the external components of the small-signal model are known, they can be de-embedded from the measured S–parameters [63]. The *de-embedding* technique removes the aforementioned parasitic elements by an algebric subtracion of the parasitics themselves from the S–parameters [96].

This new set of data ideally represent the S–parameters of the intrinsic part of

the model (fig. 3.6) as if they where measured directly to the device without any parasitics. Once again it is possible to transform the S–matrix into a Y–matrx and Z–marix in order to exploit the relative properties.

The intrinsic circuit can be expressed in Z–paramaters:

$$Z_{intrinsic} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} = \begin{bmatrix} R_{bi} + 1/Y_{be} & 1/Y_{be} \\ 1/Y_{be} - \alpha/Y_{bc} & 1/Y_{be} + (1-\alpha)/1/Y_{bc} \end{bmatrix} \tag{3.34}$$

where

$$Y_{be} = \frac{1}{R_{be}} + j \cdot \omega C_{be} \tag{3.35a}$$

$$Y_{bc} = \frac{1}{R_{bc}} + j \cdot \omega C_{bci} \tag{3.35b}$$

$$\alpha = \frac{\alpha_0 e^{-jw\tau_1}}{1 + jw\tau_2} \tag{3.35c}$$

where α is the common–base high–frequency current gain, α_0 is the DC common–base gain[38], τ_1 and τ_2 model the collector and base transit times respectively. From the last equations the next ones are derived:

$$R_{bi} = Z_{11} - Z_{12} \tag{3.36a}$$

$$R_{be} = \frac{1}{Re(Z_{12})} \tag{3.36b}$$

$$C_{jbe} = \frac{Im(Z_{12}^{-1})}{\omega} \tag{3.36c}$$

$$R_{bc} = \frac{1}{Re(Z_{22} - Z_{21})} \tag{3.36d}$$

$$C_{bci} = \frac{Im(Z_{22} - Z_{21})^{-1}}{\omega} \tag{3.36e}$$

$$\alpha = \frac{Z_{12} - Z_{21}}{Z_{22} - Z_{21}} \tag{3.36f}$$

[38]The common–base gain is defined as $\alpha = I_C/I_E$. Hence it is possible to write $\beta = \alpha/(1-\alpha)$.

The values extracted are reported as estimated values in table 3.7.

One further element, C_{bcx}, the base–collector extrinsic capacitance is still to be determined. Following the same fashion of extracting C_{bcx} by S–parameters it would lead to a complicate process. As it has been done for the parasitics, the base–collector extrinsic capacitor will be determined by geometrical considerations The base collector junction is reversed biased (when the device is active) and the $400nm$ thick collector layer is fully depleted (because it is lightly doped), hence the total base–collector capacitance (C_{bc}) which consists of a component underneath the emitter C_{bci} and another under the base contacts C_{bcx} can be estimated from:

$$C_{bc} = \frac{\varepsilon_r \cdot \varepsilon_0 \cdot W_B \cdot L_B}{d} \tag{3.37a}$$

$$C_{bci} = \frac{\varepsilon_r \cdot \varepsilon_0 \cdot W_E \cdot L_E}{d} \tag{3.37b}$$

$$C_{bcx} = C_{bc} - C_{bci} \tag{3.37c}$$

where ε_0 is the vacuum permittivity, ε_r is the relative dielectric constant, W_B and L_B are the dimensions of the base mesa and d is the thickness of the collector layer.

3.4.5.5 The Optimisation

The estimation process is completed when the external elements are de-embedded and the intrinsic elements are analytically computed [63]. The estimated equivalent circuit elements are then optimised using Agilent ADS to give the best fit to the measured S-parameters. Optimization is necessary since uncertainties like

fabrication tolerances and placement of the probes-tips make the whole set of computed elements mere estimates.

Gradient optimizer was chosen for the case in object; it uses the gradient search method to arrive at new parameter values using the gradient information of the network's error function [86]. The gradient of the error function indicates the direction to move a set of parameter values in order to reduce the error function. For each iteration, the error function and its gradient is evaluated at the initial point. Then the set of parameter values is moved in that direction until the error function is minimised.

The gradient optimizer uses the Least-Squares error function to minimise the average weighted difference for the desired responses[86]. So the value for the error function represents the average weighted[39] difference for the desired responses and the value of zero indicates that all of the intended performance goals have been reached.

The optimisation goals were defined in order to minimise magnitude and phase error targets:

$$Mag(S_{ij,modelled}) - Mag(S_{ij,measured}) \rightarrow 0 \tag{3.38a}$$

$$Phase(S_{ij,modelled}) - phaseMag(S_{ij,measured}) \rightarrow 0 \tag{3.38b}$$

where i and j are the indices of the S–matrix ($i, j = 1, 2$). The following equation shows how it was calculated the percentage of error between the measured and

[39] The application allows also to increase (or reduce) the *weights* for the single S–par. in the weighted average.

modelled S-parameters:

$$\text{Error } (\%) = \frac{S_{ij,measured} - S_{ij,modelled}}{S_{ij,measured}} \tag{3.39}$$

This new approach relies on the fact that the starting assumptions on the magnitudes for the SSEC elements are reasonably close to the real values: to ensure that, the user needs to set a tight variational range for most of the elements. Only the elements that exhibited a large variation in the calculated values were allowed wider excursions in the optimisation procedure[40].

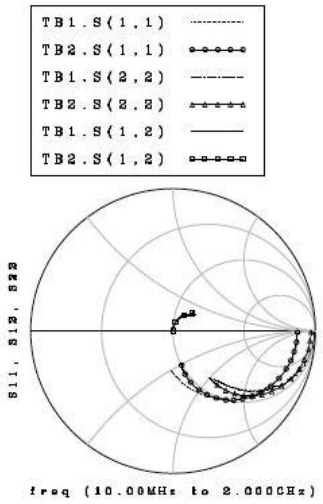
The optimisation program was then executed and it results in a better fit between the simulated and measured S-parameters: the optimisation goals were satisfied when the difference between the measured and modelled S-parameters equalled or approached zero (according to eq. 3.38).

Figures 3.4.5.5 show a very good fit between measured and modelled S-parameters from 0.01 to 2 GHz. S_{11}, S_{12}, S_{22} are shown in a Smith Chart while S_{21} on a rectangular plot.

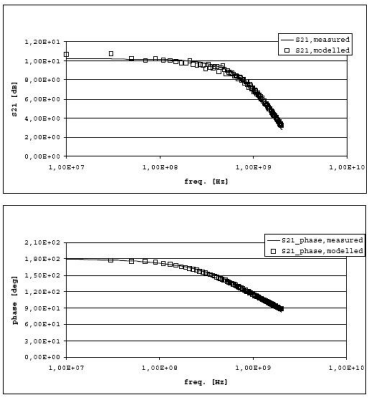
Tables 3.6 and 3.7 give details on the values on the components pre– and post–optimization.

Some considerations can be done on table 3.6: the external inductances show a good *pre–post ratio*, this means that the analytical approach was correct. Some

[40]The parameters that needed higher ranges are the ones that are connected to the emitter mesa dimensions, L_E and W_E. The wet etch process indeed can change sensibly these dimensions thus affecting all the related parameters. It does not happen for the base–collector related parameters because the b–c junction is much wider than the possible undercut so the relative variation is negligeable.



(a) S-parameters (S11, S12, S22): comparison between measures and simulations.



(b) S-parameters (S21, magnitude and phase): comparison between measures and simulations.

FIGURE 3.17: Scattering parameters, comparison between measures and simulations.

TABLE 3.6: Estimated and optimized parasitic components with relative $\%error$.

	R_e Ω	R_c Ω	R_b Ω	L_e pH	L_c pH	L_b pH	C_{pbc} fF	C_{pce} fF	C_{pbe} fF
Estimated	0.40	1.91	28.14	8.41	110	100	2.06	25.48	26.31
Optimized	0.39	2.18	20.01	8.14	108	96.9	2.25	28.37	27.82
Relative Error	2.50%	14.14%	28.89%	3.21%	1.82%	3.10%	9.22%	11.34%	5.74%

improvements should be done for the parasitic capacitances: they are mainly generated by the insulating layer of Polyimide (in this case) and their geometry is much more difficult to be determined than other elements; the triple–mesa structure makes difficult to establish average thicknesses and areas. The three access resistors (R_e, R_b and R_c) have a wide variation in the relative errors: a precise value for R_e was *easy* to calculate because the stream of electrons at the interface metal–semiconductor is *perfectly* vertical and the interface itself is relatively

undercut–safe[41]. The R_c revealed to be slighty higher than expected: further improvement can be made by taking into account for current crowding effects [97, 98] at the edges of the b–c interface. Base access resistance present an error of $\approx 30\%$: R_b strongly depends from emitter dimensions and even a slight diversion from the design values for L_E and W_E will affect the estimation process. Current crowding should be taken in account for the base too [98].

Table 3.7 presents the data for the estimated and optimized intrinsic components of the SSEC model.

TABLE 3.7: Estimated and optimized intrinsic components.

	α	R_{be} Ω	R_{bi} Ω	R_{bc} Ω	C_{be} pF	C_{bcx} fF	C_{bci} fF
Estimated	0.964	30	147	22	4.5	440	40
Optimized	0.966	23.2	100	20.9	3.82	450	31
Relative Error	0.21%	22.67%	31.97%	5.00%	15.11%	2.05%	22.50%

Comments on table 3.7 can be similar to the ones on table 3.6: the emitter–area–independent parameters: α, R_{bc} and C_{bcx} have been precisely determined in the early stage of estimation. The others, R_{be}, R_{bi}, C_{be} and C_{bci}, need larger optimisation ranges in order to get a good approximation between $S_{modelled}$ and $S_{measured}$. For this reason it is recommended to consider undercut side–effect in future modeling.

[41] The emitter metal lies on an InGaAs/InP substrate, the selective wet–etch for this structure is very slow for the InGaAs and very quick for InP.

3.4.6 Current–Gain and Power–Gain cut–off frequencies

For a complete analysis of the devices under test, current–gain and power–gain cut–off frequencies (f_T and f_{max}) were taken.

As mentioned the frequency f_T is the frequency at which the a.c. current–gain becomes unity: it is possible to measure precisely this quantity by operating some calculations on the measured S–parameters.

Starting from the definition of current gain:

$$\text{Current Gain} = \frac{i_{output}}{i_{input}} \tag{3.40}$$

the expression $\frac{i_{output}}{i_{input}}$ has been already found in paragraph 3.4.2, eq. 3.23b; so the former equation can be written:

$$\text{Current Gain} = \frac{i_{output}}{i_{input}} = h_{21} \tag{3.41}$$

transformation between from S–matrix to h–matrix is possible by:

$$h_{21} = \frac{-2S_{21}\sqrt{R_{01}R_{02}}}{(1-S_{11})(Z_{02}^{*}+S_{22}Z_{02})+S_{12}S_{21}Z_{02}} \tag{3.42}$$

Also f_{max}, the cut–off frequency for the Power–Gain, can be calculated by S–parameters.

$$\begin{aligned}\text{Power Gain} &= \frac{\text{Power Delivered to Load}}{Power\ Delivered\ to\ Network} \\ &= \frac{|S_{21}|^2\left(1-|\Gamma_L|^2\right)}{\left(1-|S_{11}|^2\right)+\left[|\Gamma_L|^2\left(|S_{22}|^2-|D|^2\right)\right]-2Real(\Gamma_L N)}\end{aligned} \tag{3.43}$$

if Γ_L=0, 3.43 can be simplified in:

$$\text{Power Gain} = \frac{|S_{21}|^2}{1 - |S_{11}|^2} \tag{3.44}$$

From equations 3.41, 3.43 and 3.44 it is possible to draw the relative plots as a function of the frequency.

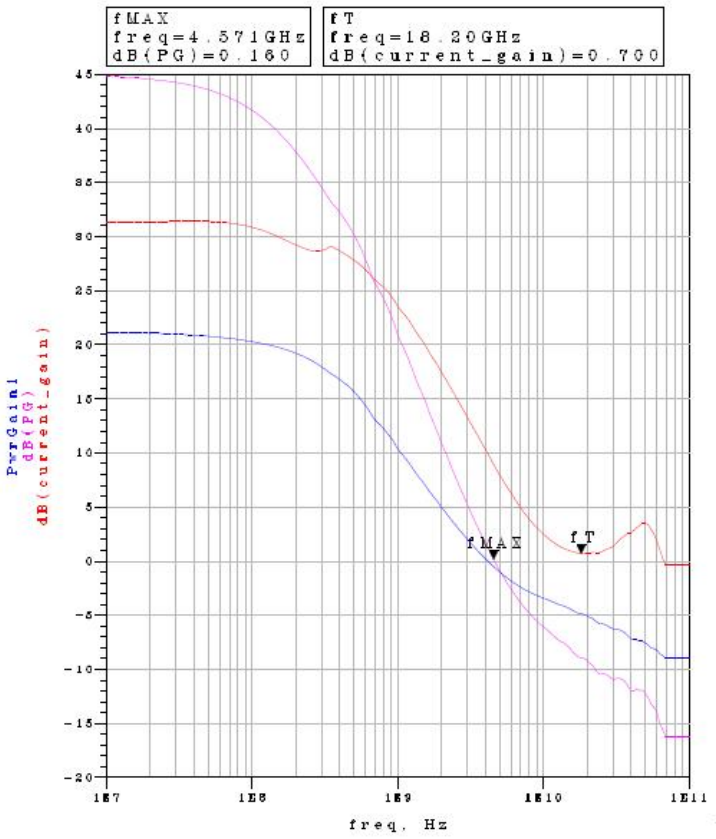


FIGURE 3.18: RF performance for a 6x10 μm^2 HBT. The red line represents the current gain (eq. 3.41) and sets an f_t of 20GHz. The purple and the blu lines plot the power gain in accordance of the different matemathical approaches(eq. 3.43, 3.44): they both agree on the value of f_{max}, 5 GHz.

The plot in figure 3.18 is current gain and the power gain (both expression, plain and simplified). The relative cut–offs happens at 20 GHz and 5 GHz respectively. It should be noticed that f_{max} is not affected by the simplifying assumption of

eq. 3.44. These figures are drastically low compared with the state–of–art HBTs operating frequencies: large improvements would be achieved in scaling down the devices' dimensions.

Chapter 4

Resonant Tunnelling Diodes

A *resonant tunnelling diode* (RTD) is a semiconductor device whose peculiar epitaxial structure creates, in the conduction band, a finite *potential well* buried between two thin finite *potential barriers*: this system goes under the name of *Double Barrier Quantum Well*, or DBQW. The electrons will be able to pass through the DBQW thanks to the tunnelling effect which occurs at specific bias voltages; on the other hand, the quantum well will allow only a precise number of discrete energies (quantisation) for the moving particles [99, 100].

RTDs can be very compact and are also capable of ultra high-speed operations because the quantum tunnelling effect through the very thin layers is a very fast process [27, 53, 56, 101, 102, 103]. The current-voltage characteristic exhibits one, or more, negative differential resistance (NDR) region which enables many applications: MMIC oscillators based on RTDs have already achieved the TeraHertz region [32] but with low output power. future improvement on device scaling and layout techniques will lead to high–power TeraHertz sources.

Section 4.1 describe the theory and operation of a basic resonant tunnelling diode;

section 4.2 will show the design and fabrication of a single device. DC and RF measurement on single devices were carried out too: section 4.3 will describe the results obtained.

4.1 Theory and operation of Resonant Tunnelling Diodes

As mentioned RTDs have a typical current–voltage characteristic: figure 4.1 shows an example; it is possible to identify two points in the diagram: (V_{peak}, I_{peak}) and (V_{valley}, I_{valley}). They determine three regions:

i) $V \leq V_{peak}$ the trend is almost linear;

ii) $V_{peak} \leq V \leq V_{valley}$ current reduces with increasing voltage;

iii) $V \geq V_{valley}$ the current increases and the trend is again almost linear.

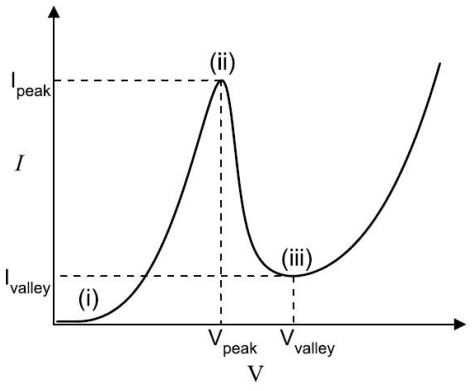


FIGURE 4.1: Typical RTD I–V characteristic. The tunnelling current is plotted against the voltage. In regions (i) and (iii), there is a diode–like behaviour, region (ii) is the NDR region.

The I–V curve, with its distinctive behaviour is determined [43] by the layer structure in use (for details see table 4.1); RTDs' epitaxy is specifically engineered to

exploit two important physical phenomena: the *potential well* and the *tunnel effect*.

4.1.1 Introduction to the Quantum Well

Quantum Physics (QP) is essential in order to understand and design resonant tunnelling semiconductor structure[1]. In the present paragraph, potential quantum well and square quantum well of finite depth will be briefly introduced.

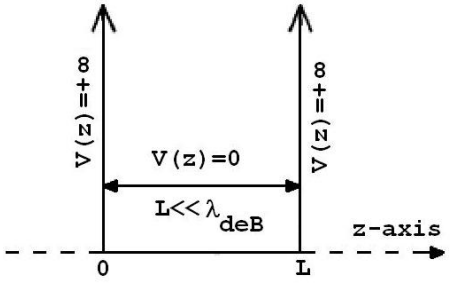


FIGURE 4.2: Standard potential well of thickness L. The barriers are infinitely high while the potential within the well is null.

A potential well is a region surrounding a local minimum of potential energy: in the present study the potential well is a local minimum of the conduction band[2], E_C.

The simplest case provided by QP is called infinite potential well[3]: an example is provided in fig. 4.2. It considers the case that the particle may only move backwards and forwards along a straight line[4] with impenetrable barriers at both

[1] It is necessary to deal with Quantum Physics because the semiconductor structures in use are much smaller than these characteristic length scales: *de Broglie wavelength* of the carriers, *mean free path* of the carriers, *phase–relaxation length*. When a structure satisfies these requirements it is called *mesoscopic*, in the opposite case it shows classical electrical behaviour and it is simply called *ohmic*.

[2] Only the behaviour of electrons in the conduction band will be considered, but similar considerations apply to holes in the valence band.

[3] Also: particle in a box or infinite square well.

[4] For the case of study, the straight line will be superimposed with the z–axis.

ends[5]: the quantum well extends from $z = 0$ to $z = L$, for $L \ll \lambda_{deB}$, where L is the thickness of the well and λ_{deB} is the *deBroglie wavelength*. The walls of a one–dimensional box are conceived as regions of space with an infinitely large potential energy ($V(z) = +\infty$, for $z = 0$ and $z = L$). Conversely, the interior of the box has a constant potential energy ($V = 0$, for $0 \leq z \leq L$): as a consequence, no forces are acting upon the particle inside the box and it can move freely in that region; however, the above mentioned infinitely large forces repel the particle at the walls of the box, preventing it from escaping.

The motion of the electron within the well can be described by the time–independent Schroedinger equation[6] [99, 100]:

$$-\frac{\hbar^2}{2m} \cdot \frac{d^2\psi}{dz^2} + V(z) \cdot \psi(z) = E \cdot \psi(z) \tag{4.1a}$$

if applied to the quantum well (no potential energy, V=0), it is rewritten:

$$-\frac{\hbar^2}{2m} \cdot \frac{d^2\psi}{dz^2} = E \cdot \psi(z) \tag{4.1b}$$

A general solution for the differential equation 4.1b can be written [99, 100]:

$$\psi(z) = A\sin(kz) + B\cos(kz) \tag{4.2}$$

Applying eq. 4.2 to eq. 4.1b, gives:

$$\frac{\hbar^2 \cdot k^2}{2m} \cdot [A\sin(kz) + B\cos(kz)] = E \cdot [A\sin(kz) + B\cos(kz)] \tag{4.3a}$$

[5] This is the simplest example, also called one–dimensional problem.
[6] This treatment is valid only if the mass of the electron is constant.

which simplifies in to:

$$\frac{\hbar^2 \cdot k^2}{2m} = E \tag{4.3b}$$

Consideration of the boundary conditions will yield the, yet unknown, constant k. As mentioned, barriers are supposed to be infinitely high outside the well: the solutions to the Schroedinger equation must remain well–behaved[7] and the term $V(x) \cdot \psi(x)$ can be finite only if $\psi(x) = 0$ at the barriers. This leads to the conditions $\psi_{x=0} = 0$ and $\psi_{x=L} = 0$. These two can be simultaneously satisfied by $\sin(kx)$.

The latter condition can be written as $\sin(kL) = 0$ so kL must be an integer multiple to π:

$$k \cdot L = n \cdot \pi \Rightarrow k = \frac{n \cdot \pi}{L} \tag{4.4}$$

The integer number n is the so–called *quantum number*; the direct consequence of eq. 4.4 is the restriction of the allowed energy levels within the well; this process goes under the name of *quantisation* and the energy levels are called *bound–states*. Then the quantised wave functions and energies are written[8]:

$$\psi_n(z) = A_n \sin\left(\frac{n\pi z}{L}\right) = \sqrt{\frac{2}{L}} \cdot \sin\left(\frac{n\pi z}{L}\right) \tag{4.5a}$$

$$A_n = \sqrt{\frac{2}{L}} \tag{4.5b}$$

$$E = E(k_n) = \frac{\hbar^2 k_n^2}{2m} = \frac{\hbar^2 \pi^2 n^2}{2mL^2} \tag{4.5c}$$

[7] In this case a well–behaved function is a continuous differentiable function.

[8] Equation 4.5b is obtained by the normalisation of the wave–function: $\int_0^L \psi^*(z) \cdot \psi(z) dz = 1$

Figure 4.3 is a representation of the first three bound–states within an idel potential well.

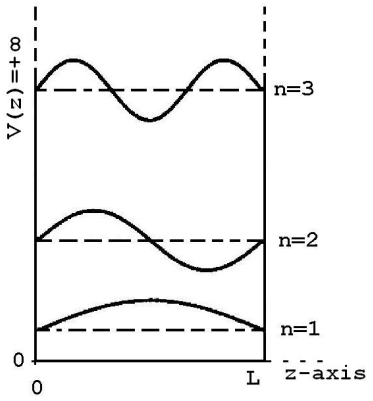
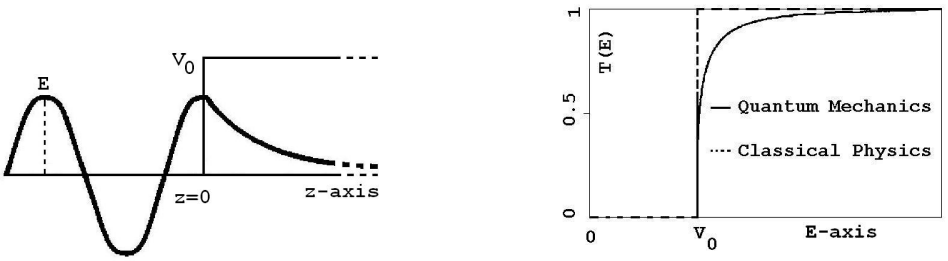


FIGURE 4.3: Ideal depiction of a Quantum Well of thickness L and surrounded by infinitely high potential barriers. The first three bound–states ($\psi_n(z)$ for $n = 1, 2, 3$) have been represented.

4.1.2 Potential Barrier and Tunnel Effect

The *Step Potential* problem is a standard one-dimensional problem typical of Quantum Mechanics (QM): in Classical Physics when an electron e of (kinetic) energy E, travelling in the $+z$ direction, hits the positive face of a potential step ($z_0 = 0$), where the potential $V(z)$ goes from 0 (for $z \leq 0$) to V_0 (for $z \geq 0$), the electron itself will be either reflected by the step ($E \leq V(z_0)$) or will pass over the step ($E \geq V(z_0)$). On the contrary, Quantum Mechanics states that the electron has a finite probability to be transmitted (i.e. to *tunnel*) through the potential barrier (fig. 4.4(a)).



(a) Wave impinging on a Step–Potential

(b) Transmission coeff. for a potential barrier.

FIGURE 4.4: Ideal depiction (a) for a wave–particle hitting a potential–step, V_0. In classical physics a particle with energy E, will be able to pass the step only if $E \geq V_0$: in this case the transmission coefficient (b) will be either $T = 0$ or $T = 1$. In quantum mechanics the particle can penetrate the step with a probability of being transmitted $0 \leq T(E) \leq 1$.

One way of quantifying the proportion of electrons that penetrate the step is in terms of the transmission coefficient, $T = T(E)$, where T is the transmission coefficient as a function of the energy E of the particle. Transmission coefficient is defined as the probability that any single electron impinging on a potential structure will penetrate it. In classical treatment T can be either $T = 0$ or $T = 1$, respectively if the particle is repelled by the step V_0 or if the particle is able to pass above the step itself. On the other hand quantum mechanics allows any value in the range for $0 \leq T(E) \leq 1$: figure 4.4(b) compares the classical and quantistic approach to the potential–step problem.

Potential Barrier

The *potential barrier* is a specific case of the step–potential problem: not only the height of the step is a finite number (V_0) but also its width, d. In this configuration, the particle hitting the barrier ($z = 0$) would emerge from the other side of

the barrier itself ($z = d$): this process is known as *tunnel effect*. An understanding of the potential barrier and the relative tunnel effect is necessary for the study of the DBQW operation and the current flowing through it.

Figure 4.5 shows a schematic representation of the potential barrier problem: in this case, the time–independent Schroedinger equation 4.1a can be *indexed* in order to describe the solution (wave–function, ψ_i) for each region (where the index i indicates the relative region in the figure 4.5):

$$-\frac{\hbar^2}{2m}\cdot\frac{d^2\psi_i}{dz^2} + V_i(z)\cdot\psi_i(z) = E\cdot\psi_i(z) \tag{4.6}$$

The electron in region 1 ($i = 1, z \leq 0$) is represented by the wave–function ($\psi_1(z)$) that gathers the incident wave–function (amplitude A_1) and the reflected wave–function (amplitude B_1), in region 3 ($i = 3, z \geq d$) the electron's wave–function ($\psi_3(z)$) is only transmitted (only amplitude A_3 as $B_3 = 0$). Function $\psi_2(z)$ represents the behaviour of the wave–particle within the barrier ($i = 2, 0 \leq z \leq d$) where components decaying in both directions, with a phase difference between

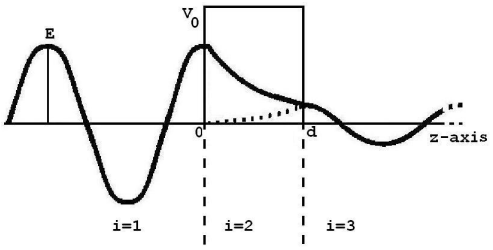


FIGURE 4.5: Potential Barrier problem: an electron, e, of energy E hits a potential barrier of energy V_0 and thickness, d. Three regions are defined: $z \leq 0$ for the incoming/reflected particle, $0 \leq z \leq d$ inside the barrier and $z \geq d$ for the outcoming particle.

them (A_2 and B_2 are the relative amplitudes). The *indexed* wave–function is then:

$$\psi_i(z) = A_i \sin(k_i z) + B_i \cos(k_i z) \tag{4.7}$$

Some considerations on well-behaving of the wave–function $\psi_i(z)$ provides the transmission coefficient across the barrier [99, 100]:

$$T(E) \approx \frac{16E}{V_0} \cdot exp\left(-2d\sqrt{2m\frac{V_0 - E}{\hbar^2}}\right) \tag{4.8}$$

4.1.3 Double Barrier Quantum Well

After the considerations made in par. 4.1.1 and 4.1.2, the double barrier quantum well is defined: growing alternatively narrow and wide band–gap material it is possible to create a region of potential well surrounded by thin barriers[9].

This pattern has been represented in figure 4.6: the tunnel effect allows the

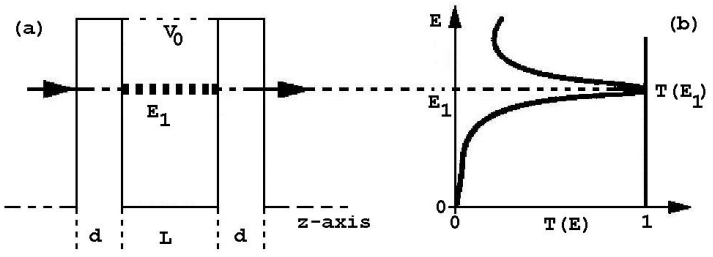


FIGURE 4.6: Ideal representation of a *Double Barrier Quantum Well* (a): the well, of thickness L, is surrounded by two barriers of thickness d. The depth of the well is also the height of the barriers: V_0. One resonant state (E_1) is depicted: it corresponds (b) to the maximum ($T(E_1) = 1$) of the transmission coefficient, T=T(E). Other *maxima* could be found if further resonant–states are present in the quantum well($T(E_n) = 1$ for $n = 1...N$, where N is the last quantum number available).

electrons to penetrate the barriers either to enter or to leave the quantum well. In this configuration the electrons in the well are still forced in the *quantised* states but only for a certain time, called *dwell time* or τ: after this period the

[9] Generally the barriers are 10 to 20 $\AA$ wide and the well is 40 to 60 $\AA$.

electrons have a high probability to leave the quantum well. This time–limited quantised states are called *quasi–bound* or *resonant* state: hence the name *resonant tunnelling* structures.

It is possible to calculate the occupation time of each resonant state, τ_n, by the *Heisenberg's Uncertainty Principle*[10]:

$$\tau_n \approx \hbar / \Delta E_n \tag{4.9}$$

where ΔE_n is the uncertainty for the resonant state itself[11]: in the WBK[12] approximation:

$$\Delta E_n = E_n \cdot exp\left[-2L\sqrt{\frac{2m(V_0 - E_n)}{\hbar^2}}\right] \tag{4.10}$$

The current in the double barrier structure, I_{DBQW}, can be estimated starting from the transmission coefficient [99, 100]:

$$I_{DBQW} \approx \frac{2e}{h} \cdot \int_{-\infty}^{+\infty} T(E) dE \tag{4.11a}$$

$$= \frac{2e}{h} T_{pk} \cdot \int_{-\infty}^{+\infty} \left[1 + \left(\frac{E - E_{pk}}{\Gamma/2}\right)^2\right]^{-1} dE \tag{4.11b}$$

$$= \frac{2e}{h} \frac{\pi}{2} \Gamma T_{pk} \tag{4.11c}$$

where T_{pk} is the value of the transmission coefficient a the resonant state [13], E_{pk} is the energy at which $T(E_{pk})$ is maxed [14], Γ is the width at half maximum of the resonance peak.

[10] $\Delta E \cdot \Delta t \geq \hbar$
[11] This physically means that the level of the resonating electron will have a certain degree of uncertainty.
[12] WKB stands for Wentzel Kramers Brillouin: WKB is a method for finding approximate solutions to linear partial differential equations with spatially varying coefficients.
[13] Ideally $T_{pk} = 1$
[14] Ideally $E_{pk} = E_n$

4.1.4 Effects of applied bias to an DBQW structure

As mentioned in the beginning of the present chapter, the current–voltage characteristic of an RTD is determined by its epitaxial layer structure. It is possible now to associate the I–V behaviour to the DBQW biasing conditions.

It has been highlighted how electrons hitting a double barrier will be able to pass through the structure only if their kinetic energy (E) equals the energies (eq. 4.5c) of the quasi–bound states (E_n) within the well (in an uncertainty range ΔE_n).

The bias voltage across the DBQW bends the conduction band (figures 4.7): as

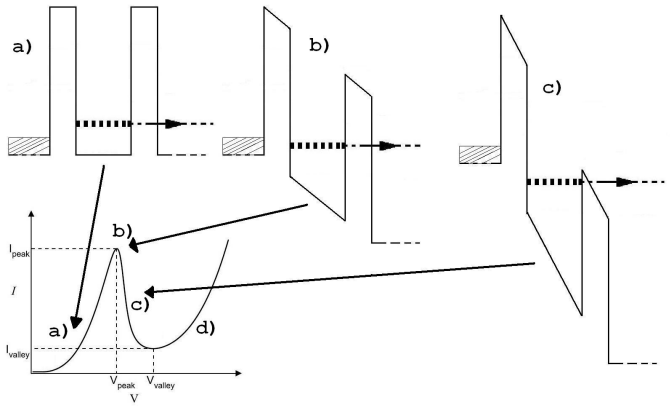


FIGURE 4.7: Biasing conditions of Double Barrier Quantum Well: for $V_{bias} \approx 0\,\mathrm{V}$ there is no current flowing (a). When V_{bias} is increased, the flowing current increases as well, until $V_{bias} \approx V_{peak}$ (b): this is caused by the first quasi–bound state ($E = E_1$) that is lying in correspondance of the Fermi Sea ($E_C \leq E_{electrons} \leq E_F$). Once the quasi–bound state does not match the Fermi Sea ($V_{peak} \leq V_{bias} \leq V_{valley}$), the current decreases (c). For $V_{bias} \geq V_{valley}$, the current increaes again (d).

a consequence, the resonant levels within the well are shifted downwards [99, 100]. When the first level approaches the *Fermi Sea*[15], the electrons start tunnelling (hence the increase of the current against the bias). When the tunnel level will be aligned with the Fermi Sea, the current ($I = I(V)$) will reach the peak, $I_{peak} = I(V_{peak})$, and then will reduce once the resonant level will move away

[15]The Fermi Sea is the portion of conduction band,filled with electrons, between the Fermi Level, E_F, and the bottom of the band itself, E_C. The name *sea* comes after the fictional representation of the electrons filling the lower part of the conduction band as water is filling the sea.

in its downwards movement.

Some *considerations* must be done noticing that, for $V_{bias} \geq V_{valley}$, the current starts increasing again. If the barriers are high enough, there might be further resonant levels (from eq. 4.4, n=2, ..., N): the pattern of peaks and valleys will be replicated as much as there are quasi–bound states.

This is not the only explanation for the current increasing: if the tunnel effect had been the *only* phenomenon acting, this would have led to $I_{valley} = 0\,\mathrm{A}$. Since there is current even though there are not energy levels available to allow the tunnel effect, it means that there must be further transport process running along with the tunnel effect.

Transport phenomena in the DBQW

Resonant tunnelling is the major responsible for current transport in DBQW–based devices but other effects must be taken in account (fig. 4.8).

A description of these mechanisms follows [102]:

- *thermo–ionic emission*: a fraction of electrons in the tail of the Fermi–Dirac distribution will have enough energy to pass over both the barriers;
- *field assisted tunnelling*: electrons with sufficiently high energies can flow non–resonantly or through higher lying quantum levels;

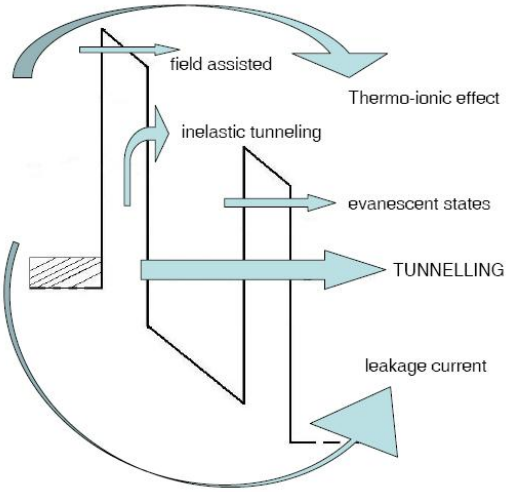


FIGURE 4.8: Transport phenomena in the DBQW: the major effect is the tunnelling. Other effects to take in account are: thermo–ionic emission, field assisted tunnelling, evanescent states, inelastic tunnelling and leakage current.

- *evanescent states*: incident electrons have a small probability to tunnel through the non resonant energy ranges between the resonances;
- *inelastic tunnelling*;
- *leakage current*: current flowing through the periphery of the device mainly generated by surface effects.

4.1.5 RTD small–signal equivalen circuit

The standard small–signal equivalent circuit of an RTD is represented in fig. 4.9 [44, 56, 101, 102].

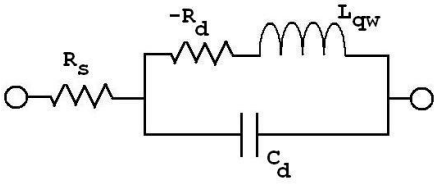


FIGURE 4.9: Small–signal equivalen circuit of an RTD. R_s is the series resistance, R_d is the internal resistance, C_d is the internal capacitance, L_{qw} is the internal inductance.

It includes the following elements:

- series resistance, R_s: arising from the ohmic contacts, the emitter and collector regions and from spreding current effects;
- parallel capacitance, C_d: resulting from charging and discharging of the DBQW and the depletion regions;
- negative differential resistance[16], $-R_d$ representing the corresponding negative differential conductance, $-G_d = -R_d^{-1}$;
- inductive element, L_{qw}: associated with the delay of the current with respect to the voltage, which arises from the necessary time to bild–up the charge in the quantum well[17].

Maximum oscillation frequency for an RTD

When a current $i = i(f)$ flows into the RTD (fig. 4.10), the electrical power spent in the diode is given by $Re(Z_{in}) \cdot i(f)^2$, where the first term is the real part of the impedance on the diode as a function of the frequency.

If $Re(Z_{in})$ is negative, the RTD supplies electrical power to the external circuit, thus showing *gain* [56, 102].

The maximum oscillation frequency, f_{max}, is defined as the frequency at which the real part of the impedance is null [56, 102].

[16] The resistance showed by the device is negative only when the RTD is biased in $V_{peak} \leq V_{bias} \leq V_{valley}$.
[17] The inductance, L_{qw}, is chosen in order to satify: $\tau = L_{qw} \cdot G_d$ where τ is the dwell time.

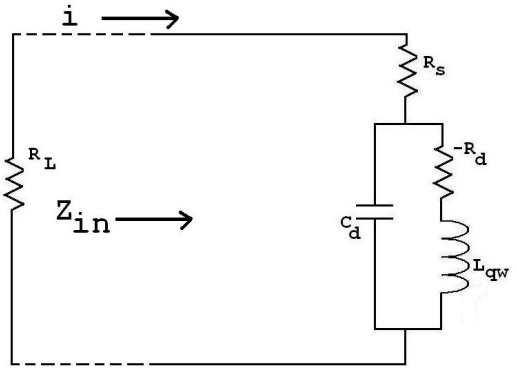


FIGURE 4.10: Schematic representantion of a current i flowing into an RTD with a load, R_L, in parallel.

$$f_{max} = \frac{1}{2\pi}\sqrt{\left[\frac{1}{L_{qw}C_d}\left(1 - \frac{C_d}{2L_{qw}G_d^2}\right)\right]\left[1 - \sqrt{1 - \frac{(1+R_sG_d)/R_sG_d}{(C_d/2L_{qw}G_d^2 - 1)^2}}\right]} \tag{4.12}$$

Negleting the resonant tunnelling time–delay[18], eq. 4.12 is written:

$$f_{max} = \frac{1}{2\pi R_d C_d}\sqrt{\frac{R_d}{R_s + R_L} - 1} \tag{4.13}$$

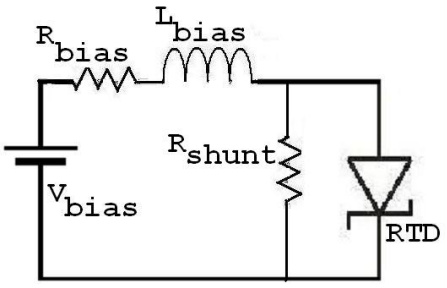
The highest values for the RTD's f_{max} (and zero output power, $R_L = 0$) can be obtained if $R_d = 2 \cdot R_L$:

$$f_{max} = \frac{1}{2\pi R_d C_d} \tag{4.14}$$

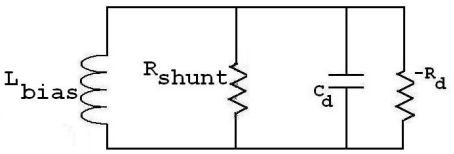
The RTD capacitance, C_d, can be designed by choosing the structural parameters of the DBQW (L, d, V_0) and the doping profile of the layers. The high frequency performaces of RTDs are also improved by increasing G_d or decreasing R_s. Large values of G_d are obtained by high peak current and low valley current[19]. High PVCR can be expected by devices based on InGaAs/AlAs system [56, 101, 102].

[18] Assuming L_{qw} negligeably small.

[19] Actually high PVCR: peak to valley current ratio, $\frac{I_{peak}}{I_{valley}}$.



(a) Schematic diagram of a biased RTD.



(b) Schematic diagram of the equivalent circuit of a biased RTD.

FIGURE 4.11: Self–oscillation and stability of an RTD: a complete schematic including R_{bias} (the internal resistance of the battery) and L_{bias} (the inductance of the bias line) is presented in fig. (a). Equivalent circuit is depicted in fig. (b): the battery is *opencircuited*, R_{bias}, R_s and the L_{qw} are all neglected.

Self–oscillations and stability of RTDs

It has just been showed that the NDR of RTDs puts these devices in a prominent role for oscillator design. This characteristic causes some issues as well: the internal (negative) resistance of the device will tend to *ignite oscillations* if specific conditions are achieved.

Some authors propose the use of an external resistors in order to suppress parasitic bias oscillations: [49, 56, 62]. Starting from the schematic in fig. 4.11(a), it is possible to identify the internal resistance $-R_d$, the internal capacitance C_d, the lumped inductance of the bias line L_{bias} and the shunt resistance R_{ext}; all the other elements are negligible. The Laplace transformed voltage to a current $i(s)$ is:

$$v(s) = \left(\frac{1}{R_{shunt}} - \frac{1}{R_d} + \frac{1}{sL_{bias}} + sC \right)^{-1} \cdot i(s) \tag{4.15}$$

which can be written as:

$$v(s) = \frac{s}{2i\omega_0 C}\left(\frac{1}{s-\gamma-i\omega_0} - \frac{1}{s-\gamma+i\omega_0}\right)\cdot i(s) \tag{4.16}$$

where $\gamma = (1/R_d - 1/R_{shunt})/2C_d$ and $\omega_0 = 1/L_{bias}C_d - \gamma^2$. In time domain (independent variable, t, time) the response to a current pulse $i(t) = I_0\delta(t)$ is:

$$v(s) = I_0\frac{e^{\gamma t}}{C_d}\left(\frac{\gamma}{\omega_0}sin\,(\omega_0 t) + cos\,(\omega_0 t)\right) \tag{4.17}$$

For $\gamma \geq 0$ (that corresponds to $R_d \leq R_{shunt}$) any finite current fluctuation will ignite an oscillation with exponentially increasing amplitude. So the low–frequency stability *rule* can be written:

$$R_{shunt} \leq R_d = \frac{1}{G_d} \tag{4.18}$$

To avoid any possible self–oscillation, one or more shunt resistors will placed in parallel to the device both for DC and RF analysis.

4.2 Design and fabrication of a Resonant Tunnelling Diode

The present section introduces the layer structure used for RTDs in the project; an overview of the photolithographic techniques will be also provided.

4.2.1 Layer structure

The layer structure used for the present work (table 4.1) was defined in a previous project: the wafer was grown by Molecular Beam Epitaxy in a Varian Gen II

system on an semi–insulating InP substrate by IQE Ltd. This material was used as it was available and had reliable NDR characteristics.

TABLE 4.1: Layer structure in use for the resonant tunnelling diodes.

		Material	**Thickness** [nm]	**Doping type**	**Doping concentration** [cm^{-3}]
01	Emitter	$In_{0.53}Ga_{0.47}As$	40	n^{++}	Si:$2 \cdot 10^{19}$
02	Emitter	$In_{0.53}Ga_{0.47}As$	80	n^{++}	Si:$2 \cdot 10^{18}$
03	Spacer	$In_{0.53}Ga_{0.47}As$	50	n^{+}	Si:$5 \cdot 10^{16}$
04	Barrier	AlAs	1.4	undoped	-
05	Well	$In_{0.53}Ga_{0.47}As$	5.5	undoped	-
06	Barrier	AlAs	1.4	undoped	-
07	Spacer	$In_{0.53}Ga_{0.47}As$	50	n^{+}	Si:$2 \cdot 10^{16}$
08	Collector	$In_{0.53}Ga_{0.47}As$	80	n^{++}	Si:$2 \cdot 10^{18}$
09	Etch Stop	$In_{0.52}Ga_{0.48}As$	10	n^{++}	Si:$1 \cdot 10^{19}$
10	Collector	$In_{0.53}Ga_{0.47}As$	200	n^{++}	Si:$3 \cdot 10^{19}$
11	Buffer	$In_{0.53}Ga_{0.47}As$	200	n^{++}	Si:$2 \cdot 10^{19}$
12	SI substrate	InP	$625 \cdot 10^{3}$		

The Emitter is the topmost layer (lay.1): because of the smaller band-gap of InGaAs and the relative high doping level, it is easy to make very good ohmic contact; a less doped emitter layer (lay. 2) is necessary to reduce the potential barrier step between the metal contact and the upper layer of the DBQW (lay. 3). Then double barrier follows (layers from 4 to 6). In the same fashion, a symmetric series of layers (from 7 to 11) matches the bandgap structure between the lower layer of the DBQW and the Collector.

Once the layer structure is defined it is possible to estimate the (maximum) number of resonant states (N) that exist within the quantum well. The following system of equations [99, 100] allows to determine N:

$$\left\{ \begin{array}{c} tan \\ \\ -cot \end{array} \right\} \theta = \sqrt{\frac{\theta_0^2}{\theta^2} - 1} \tag{4.19a}$$

where

$$\theta_0^2 = \frac{mV_0L^2}{2\hbar^2} \tag{4.19b}$$

In eq. 4.19, θ is the *adimensional* independent variable[20], m is the mass of the electron (assumed constant), V_0 is the height of the $AlAs$ barrier with respect of bottom of the conduction band of the undoped $InGaAs$, L is the width of the quantum well and $\hbar$ is the reduced Planck constant[21].

θ_0	m_{eff}	V_0	L	$\hbar$
30.83	$6.74 \cdot 10^{-32}$ kg	$1.48 \cdot 10^{-19}$ V	$8.3 \cdot 10^{-9}$ m	$1.05 \cdot 10^{-34}$ J · s

TABLE 4.2: Physical parameters to solve eq. 4.19.

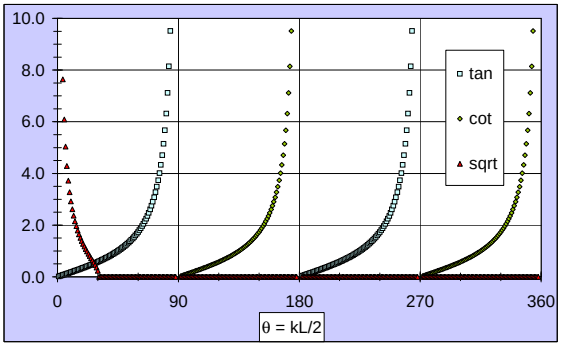


FIGURE 4.12: Number of states within the quantum well: plot of eq. 4.19; the red line intersects only once the system tan/cotan (θ).

Plot 4.12 shows that there is approximately one solution: the red line called *sqrt* intersects only once the system tan/cotan (θ), it implies that there will one pattern of Peak/Valley in the I–V characteristic of the the RTDs fabricated with this epitaxial structure.

[20] It corresponds to: $\theta = k \cdot L/2$
[21] It is equal to the Planck constant divided by 2π: $\hbar = 1.054 \cdot 10^{-34}$ J · s.

4.2.2 Fabrication Process

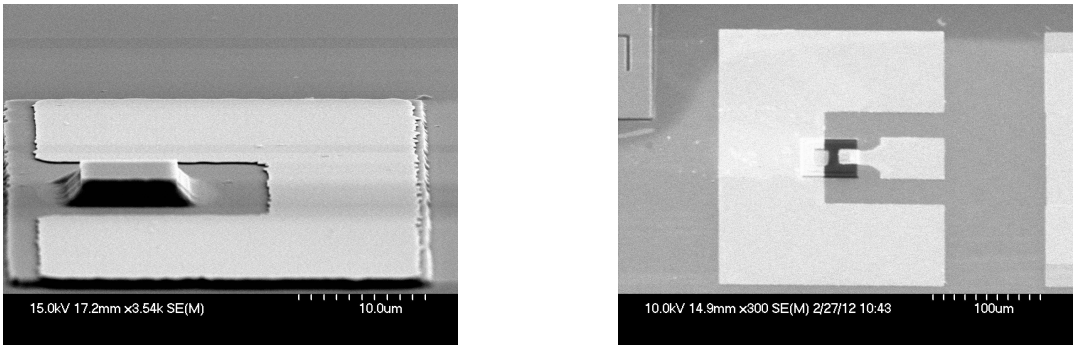
As for HBTs, RTD are named after their emitter size, so the following devices were fabricated: 3x3, 4x4 and 5x5 μm^2; table 4.3 provides an overview of the fabrication process.

The first stage of fabrication was to define the emitter metal layer using pho-

TABLE 4.3: RTD fabrication steps: optical lithography is the process in use.

Step	Step name	Procedure	Notes
01	Emitter metal	Metal deposition / lift off	Ti/Pd/Au
02	Emitter mesa	Wet etch	Layers 1-9
03	Collector metal	Metal deposition / lift off	Ti/Pd/Au
06	Collector mesa	Wet etch	Layers 9-11
07	Insulator	Deposition and d.e.	Polyimide
08	Bond-pads	Metal deposition / lift off	Ti/Au

tolithography. The Ti/Pd/Au metal scheme[22] was chosen both for emitter and collector; alignment markers were deposited along with the topmost contact. The emitter mesa structure was then defined: given the shallow etch (step 2, depth: 300 nm), it was decided to use the metal contacts as protective layers in this step. The next step was to define the collector metal pattern. After that the metal



(a) Completely formed device.

(b) Complete fab. for an RTD.

FIGURE 4.13: SEM snapshots from the fabrication of an RTD.

scheme (Ti/Pd/Au) could be evaporated and then the collector mesa was defined.

[22]The exact scheme is a blanket deposition of 20nm of titanium, 30nm of palladium and 80nm of gold.

Figures 4.13(a) shows a sample of RTD after collector mesa: after that stage an insulating layer of Polyimide and the bond–pads[23] were deposited onto the devices. Figure 4.13(b) shows an SEM picture of a finished device. The full fabrication process is given in the Appendix A.

4.2.3 Stabilisation of an RTD

As mentioned, RTDs are prone to parasitic low–frequency oscillations: eq. 4.18 sets the maximum value for an external resistor to suppress parasitic bias oscillations. As for the eq. 4.18 is mandatory to know the magnitude of the internal negative resistance (actullay the gradient of the NDR region of the I–V plot) in order to *stabilize* the RTD, a literature review of RTDs fabricated starting from the same layer structure was performed [53, 56, 101, 102, 103]. For the present work two magnitudes of resistances were chosen: $5\,\Omega$ and $10\,\Omega$.

Nanofabrication of resistors consists of the deposition of an alloy of metal (NiCr in the present project): given a known sheet resistance ($R_{sh} = 50\,\Omega/\square$ for the NiCr) the magnitude of the resistance is:

$$R_{res} = R_{sh} \cdot \frac{L_{res}}{W_{res}} \tag{4.20}$$

where L_{res} and W_{res} corresponds to the length and width of the resistor.

For practical reasons, the resistors were placed in the gaps between the coplanar waveguide (fig. 4.14): so the relative lengths, L_{res}, were defined by the CPW deposition ($L_{res} \approx 40\,\text{nm}$). The widths, W_{res}, were then defined by eq. 4.20. In

[23] Metal scheme: Ti/Au, 50nm of titanium and 150nm of gold.

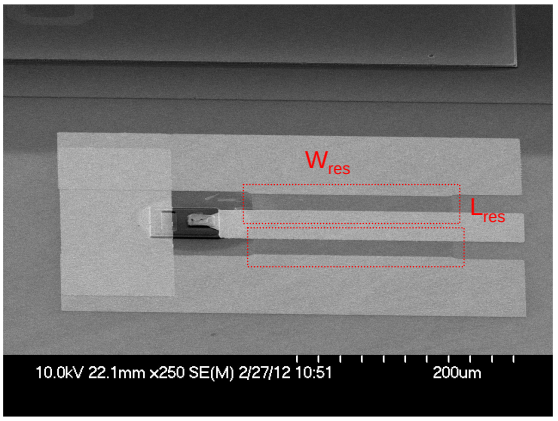


FIGURE 4.14: SEM snapshot of an RTD with a stabilizing resistor: as they are very thin (and opaque) the two resistors have been highlighted in red–dotted squares.

order to dissipate much of the heat produced, it was chosen to use two resistors in parallel instead of one. So two different magnitudes of resistances were actually fabricated: 10 Ω and 20 Ω.

Resistance	R_{sh}	Width	Length
Ω	Ω/□	nm	nm
10	50	40	200
20	50	40	100

TABLE 4.4: Physical parameter for the stabilizing resistances

4.3 Device characterisation

After completing the fabrication process, all devices were tested for DC characteristics; measurements were carried out with Cascade Microtech M150 probe station by the means of the B1500A semiconductor device analyzer (data and procedures are shown in par. 4.3.1).

RF characterization was also performed by measuring S–parameters: it was carried out on the probe station Cascade Microtech M150 by the means of the Two–Port E8361 PNA network analyzer and the B1500A semiconductor device analyzer

(data and procedures are shown in par. 4.3.2).

4.3.1 DC characteristics

As RTDs are two–terminal devices, the configuration for the measurement of the current–voltage characteristic is straightforward: a variable voltage source is applied to the DUT[24] while the current flowing into the device is recorded. Figure 4.15 ideally represents this configuration.

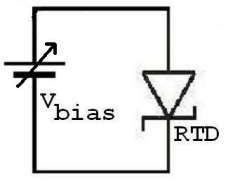


FIGURE 4.15: Schematic representation of an RTD under varying bias voltage.

The device depicted in figure 4.13 is one of the structures fabricated for I–V measurement: the connections between the devices' and the instrumentation were designed to fit RF probes.

Figure 4.16 shows the I–V characteristics for the fabricated RTDs: three different sizes were designed (3x3, 4x4 and 5x5 μm^2). In order to get more data, two devices per dimension were fabricated: the data has been then averaged before plotting.

Some important information can be extracted from the I–V plot: table 4.5 shows these parameters extracted from fig. 4.16.

[24]DUT: Device Under Test

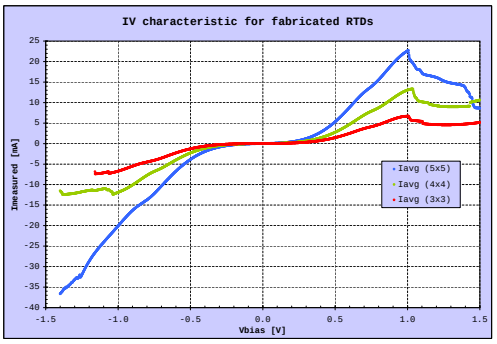


FIGURE 4.16: Current–Voltage characteristic for fabricated devices: the plotted currents (red for the 3x3 μm^2, green for the 4x4, blue for the 5x5) have been averaged amongst same–area devices. The typical NDR is clearly present for $1.0\,\mathrm{V} \leq V_{bias} \leq 1.5\,\mathrm{V}$.

TABLE 4.5: RTD DC parameters extracted from I–V plots.

Area	**Peak**	**Valley**	ΔV	ΔI	**PVCR**
3x3 μm^2	(1 V; 7 mA)	(1.5 V; 5 mA)	0.5 V	2 mA	1.4
4x4 μm^2	(1 V; 14 mA)	(1.5 V; 10 mA)	0.5 V	4 mA	1.4
5x5 μm^2	(1 V; 22 mA)	(1.5 V; 12 mA)	0.5 V	10 mA	1.83

The current flowing in a biased RTD is related to the transport phenomena as described in par. 4.1.4 but it is also linearly dependent of the active area of the device: in order to appreciate the original behaviour of the DBQW, it is necessary to extract the current density against the bias voltage (J–V plot)[25].

Figure 4.17 shows the current densities for the fabricated RTDs: the result seems to converge to the expected peak value of $100\,\mathrm{mA}/\mu m^2$ ($= 100\,\mathrm{KA/cm^2}$). This result was obtained considering as *active area* the designed values (9, 16 and 25 μm^2): it is not completely satisfactory because, as mentioned, the very same values of J_{peak} should have been obtained.

On the other hand, if the relative undercut[26] is taken in the due consideration, the plot of the current densities can be redrawn: in fig. 4.18 the peak values of

[25]The current density, J, is obtained dividing the measured currents by the relative areas, $J = I_{measured}/Area_{dev}$

[26]The mesa structure is $h_{etch} \approx 300nm$, per side: this leads to $L_{shrunk} \approx L_{design} - 2 \cdot h_{etch}$. The new areas will be $L^2_{shrunk} \approx 5.76, 11.56, 19.36\ \mu m^2$

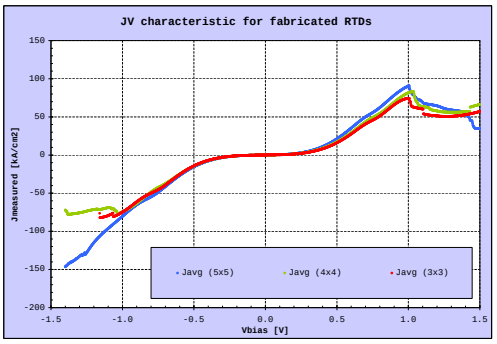


FIGURE 4.17: Current Density–Voltage characteristic for fabricated devices: the values have been obtained from the ratio between the plotted currents (fig. 4.16) and the relative areas (red for the 3x3 μm^2, green for the 4x4, blue for the 5x5).

the current density clearly converges to $J \approx 110\,\mathrm{KA/cm^2}$.

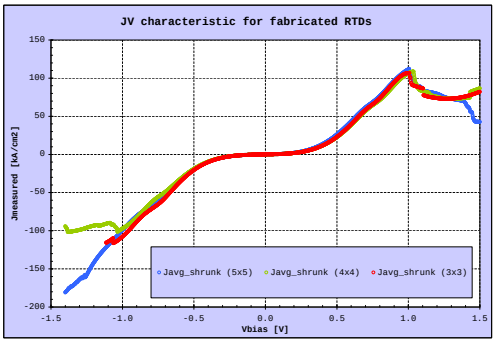


FIGURE 4.18: Current Density–Voltage characteristic for fabricated devices: the values have been obtained from the ratio between the plotted currents (fig. 4.16 and the real active areas (red for the 5.76 μm^2, green for the 11.56, blue for the 19.36).

The plot represented in fig. 4.19 has been taken from stabilized RTDs: as mentioned in par. 4.14, two different magnitudes of resistances were fabricated, 5 Ω and 10 Ω[27].

[27] Further investigations have showed how the actual values for the resistances were 4.75 Ω and 9.5 Ω.

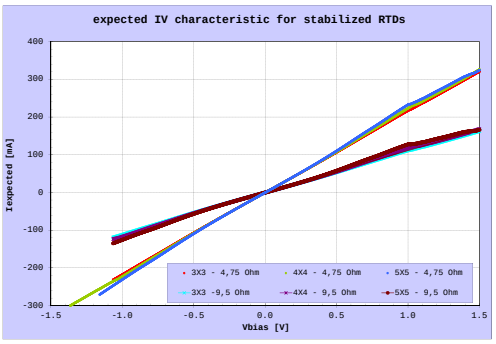


FIGURE 4.19: I–V characteristics for stabilized RTDs.

4.3.2 RF characteristics

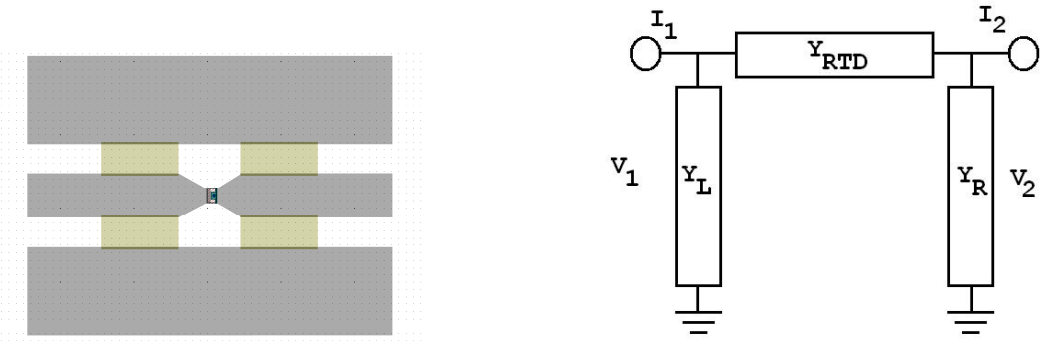
RF characterisation is fundamental for the devices that are supposed to work within high frequency analogue circuits: RTDs have already showed their potential [32, 33, 56] as core of THz oscillators. In the present section. a description of the modelling and small–signal equivalent circuit of resonant tunnelling diodes will be given.

4.3.2.1 RF text fixture

For RF characterisation of an RTD, a stabilising pi-network (in which the device is monolithically embedded) should be employed [56]. Figure 4.20 shows a schematic diagram for the pi-network.

Such a configuration allows the currents and voltages to be associated to the following Y–matrix:

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} Y_L + Y_{RTD} & -Y_{RTD} \\ -Y_{RTD} & Y_R + Y_{RTD} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \quad (4.21)$$



(a) Designed layout for RF–structures.

(b) Schematic diagram of the pi–network.

FIGURE 4.20: Representations of the designed structures for RF measurement.

From eq. 4.21 it is possible to associate currents and voltages as in fig. 4.20(b):

$$I_1 = Y_L V_1 + Y_{RTD}\,(V_1 - V_2) \tag{4.22a}$$

$$I_2 = Y_R V_2 + Y_{RTD}\,(V_2 - V_1) \tag{4.22b}$$

Starting for a measured set of S–parameters, this configuration (fig. 4.20(a)) allows to isolate each single branch from the others in order to study the relative properties. In this way the three impedances will have the following meanings:

- Y_L: the lumped element will gather the properties and behaviour of the parasitic elements (resistance, capacitace and inductance) of the *real* network from the left side of the device (including the stabilizing resistor).
- Y_R: this element is the same as Y_L but from the right side of the device.
- Y_{RTD}: this element equals the whole behaviour of the small–signal model of the RTD.

In more details, the RF test-fixture consists of the RTD device and stabilising resistors R_L and R_R in a pi-network configuration: the resistors are chosen to ensure circuit stability (eq. 4.18). At DC or low frequencies, the right hand terminal of the NDR device is grounded through the VNA internal connections. Here, the interconnects between the NDR device and the two stabilising resistors was reduced to a minimum ($\approx 50\,\mu$m) and the inductance associated with this was included in the diode lead inductance. The stabilizing left resistor (R_L) is effectively in parallel with the NDR device and suppresses bias oscillations from the left hand terminal. The inductance of the internal connections from the right side to the ground would also cause oscillations. Therefore, another resistor R_R was used to suppress the oscillations at the right hand terminal. Using this setup, S–parameter characterisation of the NDR device can be done.

4.3.2.2 Modelling and evalution of the biasing network

The biasing networks from left and right sides are described by the ideal network depicted in fig. 4.21

The magnitude of the lumped elements described in fig. 4.21 can be extracted[28] from the measured S–parameters; the procedure is straightforward and it is very similar the procedure described in the former chapter (see par. 3.4.5.1):

- Measurement of the S–parameter of the whole network;

[28] As mentioned in the former chapter, all the values will be extracted against the frequency, f.

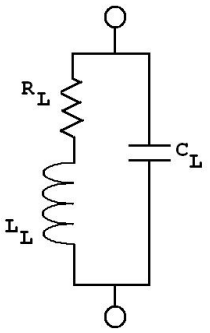


FIGURE 4.21: Equivalent circuit for the left (right) biasing network.

- Conversion of the S–paramaters into Y–parameter
- Extraction of Y_L and Y_R from eq. 4.21;
- Extraction of R_L, C_L and L_L from Y_L (R_R, C_R and L_R from Y_R).

The impedance of the network in fig. 4.21 can be written as[29]:

$$Y_L = \frac{1}{R_L + j\omega L_L} + j\omega C_L \tag{4.23a}$$

$$Re(Y_L) = \frac{R_L}{R_L^2 + \omega^2 L_L^2} \tag{4.23b}$$

$$Im(Y_L) = \frac{\omega R_L^2 C_L - \omega L_L + \omega^3 C_L L_L^2}{R_L^2 + \omega^2 L_L^2} \tag{4.23c}$$

From eq. 4.23b, the magnitude of the stabilizing resistance can be estimated[30]: considering a value of $\omega = 0\,\text{rad/sec}$, the relative real part of the admittance 4.23b becomes: $Re\left(Y_L(0)\right) = \frac{R_L}{R_L^2+0} = 1/R_L$.

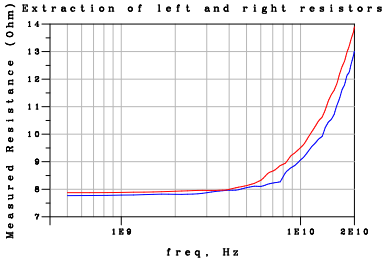
Once R_L is defined, eq. 4.23b can provide the parasitic inductance, L_L.

[29] The equation is given for the left side but is identical for the right side.

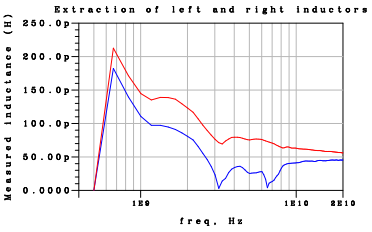
[30] To be noticed that the estimated value for R_L includes any possible (negligible) parasitic resistance on the line.

The same operation, applies for the capacitance C_L: eq. 4.23c has three starting *unknown* parameters: R_L, L_L and C_L. As the first two are already determined, it will be elementar the extraction of the third.

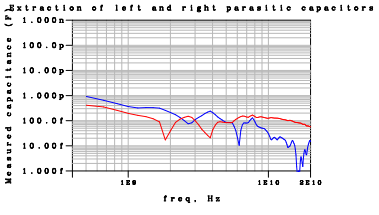
From the more practical point of view, all tha data (S–parameters) were acquired on the probe station by the E8361 PNA network analyzer: after acquisition, all the data were loaded into Agilent ADS; this CAD program allows easy transformation from S–parameters to Z– or Y–parameters and plots all the elements of of the relative matrices against frequency.



(a) Extraction of the resistance of the biasing network.



(b) Extraction of the inductance of the biasing network.



(c) Extraction of the capacitance of the biasing network.

FIGURE 4.22: Extraction of the three lumped elements of the biasing network for $V_{bias} = 1.25\,\mathrm{V}$.

The extracted figures for R_L, L_L and C_L are shown in table 4.6.

Bias–Network	Resistance	Inductance	Capacitance
	Ω	pF	fF
Left side	7.78	138	98.1
Right side	7.88	210	117.9

TABLE 4.6: Extracted lumped elements of the biasing network.

4.3.2.3 Modelling and evalution of the biasing network for RTDs

The small–signal model for the RTD depicted in figure 4.9 is the standard model for such a device [56].

In the same fashion of the biasing network, the extraction of the internal elements for the RTD model follows the procedure:

- Measurement of the S–parameter of the whole network;
- Conversion of the S–paramaters into Y–parameter
- Extraction of Y_{RTD} from eq. 4.21;
- Conversion of Y_{RTD} into Z_{RTD};
- Extraction of R_s, C_d, $-R_d = G_d^{-1}$ and L_{qw} from Z_{RTD}.

The impedance of the network in fig. 4.9 can be written as:

$$Z_{RTD} = R_s + \left(\frac{1}{R_d + j\omega L_{qw}} + j\omega C_d \right)^{-1} \tag{4.24a}$$

$$Re(Z_{RTD}) = R_s + \frac{R_d}{\left(1 - \omega^2 C_d L_{qw}\right)^2 + \left(\omega C_d R_d\right)^2} \tag{4.24b}$$

$$Im(Z_{RTD}) = \frac{\omega L_L - \omega^3 C_d L_{qw}^2 - \omega R_d^2 C_d}{\left(1 - \omega^2 C_d L_{qw}\right)^2 + \left(\omega C_d R_d\right)} \tag{4.24c}$$

If L_{qw} is cosidered negligeagle ($L_{qw} \approx 0$), the magnitude of the stabilizing resistance can be estimated from eq. 4.24b[31]: considering a value of $\omega = 0\,\text{rad/sec}$, the relative real part of the impedance becomes $R_s + R_d \approx R_d$.

Once R_L is defined, eq. 4.23b can provide the parasitic capacitance $C_L = C_L(\omega)$.

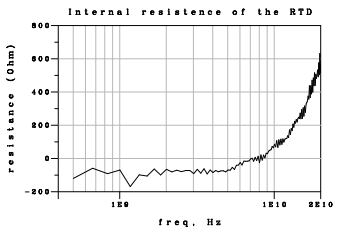
From the more practical point of view, all tha data (S–parameters) were acquired on the probe station by the E8361 PNA network analyzer: the behaviour of the network is bias–dependant mainly because the internal resistance of the RTD is bias–dependant: all the presented measurement were taken for V_{bias} within the NDR region ($1\,\text{V} \leq V_{bias} \leq 1.5\,\text{V}$)

After acquisition, all the data were loaded into Agilent ADS: this CAD program allows easy transformation from S–parameters to Z– or Y–parameters and plots all the elements of of the relative matrices against frequency.

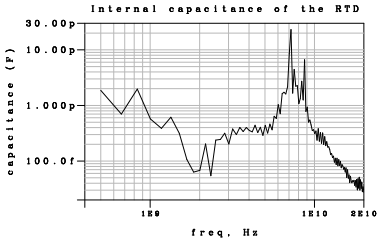
For convenience the value of the series resistance, R_s, was estimated by geometrical considerations[32].

[31] In this case it has been considered that the parasitic series resistance is negligible with respect of the internal device's resistance.

[32] The R_s is mainly given by the access resistance of the metal–semiconductor interfaces of the terminals, $\approx$ one or two Ω.



(a) Extraction of the internal resistance of the RTD.



(b) Extraction of the internal capacitance of the RTD.

FIGURE 4.23: Extraction of the internal lumped elements of the RTD for $V_{bias} = 1.25\,\mathrm{V}$.

The extracted figures for R_d and C_d are shown in table 4.7.

Bias–Network	R_d	C_d
	Ω	fF
Left side	-120	300

TABLE 4.7: Extracted internal elements of the RTD.

The data presented in table 4.6 and in table 4.7 were then put togheter in order to create a complete model of the network (fig. 4.20(a)). Figure 4.24 is a comparison between the measured set of S–parameters of the RF network and the extracted magnitudes of the lumped elements: a good matching has been reached.

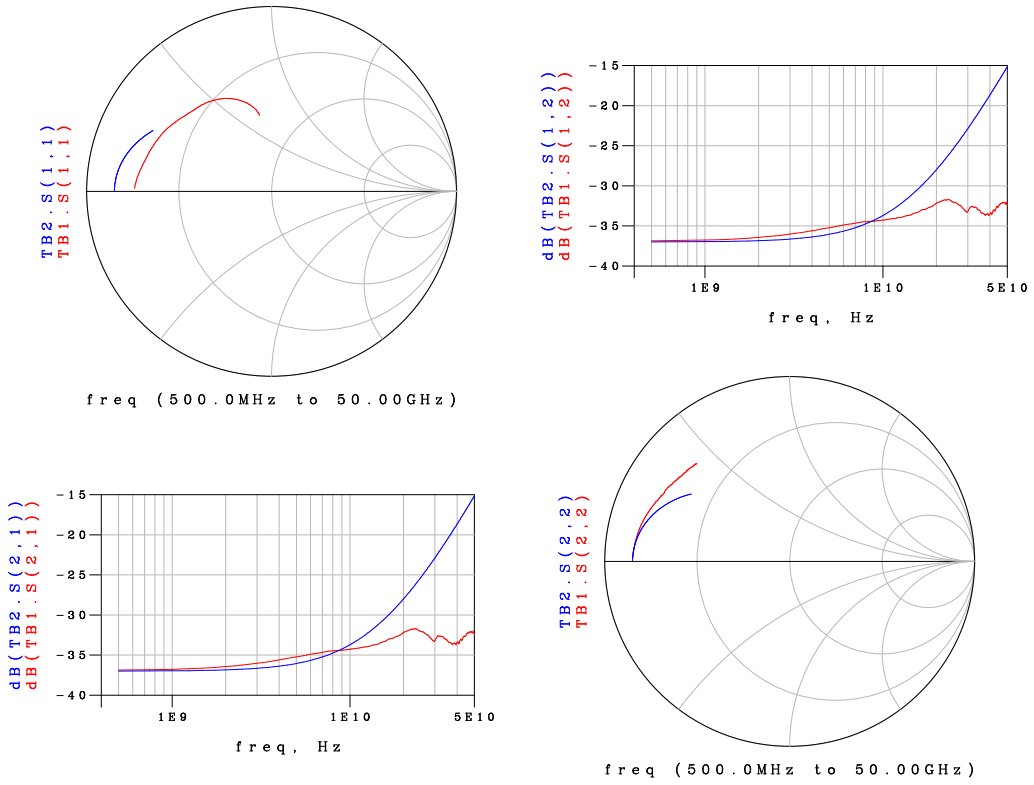


FIGURE 4.24: Comparison between measured (red) and modelled (blue) S–parameters of the RTD RF network for $V_{bias} = 1.25\,\mathrm{V}$.

Chapter 5

Conclusions and future work

The unceasing need for faster, smaller, more robust and powerful *Electronic Sources* was the motivation of this work: in the case of study, devices operating in the TeraHertz frequency region are very likely to invade the market of electronics in the next future. In particular the present thesis investigated two different devices based on the same material system (InGaAs/InP): the heterojunction bipolar transistor (HBT) and the resonant tunnelling diode (RTD).

In the first chapter a literature review was presented: THz waves can be exploited for a broad variety of purposes like sensing, imaging and communications. Many fields would benefit from these developement: spectroscopy and radio–astronomy, bio–medicine and diagnostics, security and defence. The main issue with THz waves is that they lie in a region of the spectrum where there are no sources: below 100 GHz, electronic devices are common, above 10 THz, opto–electronic devices are available. The worldwide research in the area involves several different type

of devices with two or three terminals: for the present work HBTs (3–terminals) and RTDs (2–terminals) were investigated.

The second chapter presented the photolithographic tecniques used to fabricate devices: the standard process was used.

In the third chapter HBTs were studied: in the first part, theory and principles were presented. Some devices were then fabricated and the relative measurements were showed in the second part: from the DC point of view a gain of 20 was obtained. From the RF point of view, an f_{max} of 5 GHz was obtained.

Resonant tunnelling diodes were investigated in chapter 4; after a brief introduction of quantum mechanics principles, some results from fabricated devices were presented: a current density peak of 110 KA/cm^2 with a PVCR of 1.5.

For both HBTs and RTDs a de–embedding technique was used in order to extract small–signal equivalent circuit: it resulted in a good matching between the measured and the modelled data.

5.1 Future work

HBTs and RTDs are possible candidates to realise THz–oscillators: during the project, pros and cons of both devices have been investigated.
The main feature that could favour HBTs is their higher DC–to–RF: 36% (practical result in [34, 35]) for HBTs versus 20% (theoretical result in [46, 47]) for RTDs. On the other hand, fabrication for HBTs is longer (al least one metallisation and one etching more) than RTDs; but the worst aspect is that HBTs performances are

deeply linked to the relative scaling: emitter access resistance, R_e, base–emitter gap resistance, $R_{be,gap}$, base–collector junction capacitance, C_{bc}, need to be extremely small if the device is to be employed in the THz applications. The use of e–beam lithography instead of the photolith. seems mandatory in future development of HBTs.

RTDs' operation depends on the physics of device more than the dimensions so the scaling effort (and the relative cost) is not a primary need in devices' development: fabricated device with standard photolithographic tecniques will provide excellent performances if used in the correct topology.

5.1.1 Introduction to Oscillator design

Published studies on RTD-based oscillator have shown a DC-to-RF conversion efficency less then 1% [33, 56, 104]: this is because of parasitic bias oscillations and inefficient circuit topology. A dedicated design that could stop the parasitic oscillation, would unlock the potential of RTD-based oscillators for mm–wave sources. The same oscillation–stopping topology can be used in a power combining technique [33, 56] in order to get high output power and stable oscillation frequencies at the same time.

Single Oscillator Design

Former studies [56, 62] have shown that parasitic oscillations in RTDs can be stopped by a proper *decoupling circuitry*[1]. Figure 5.1 shows a single–RTD–based oscillator equipped of the decoupling circuit. The components, that will be subject

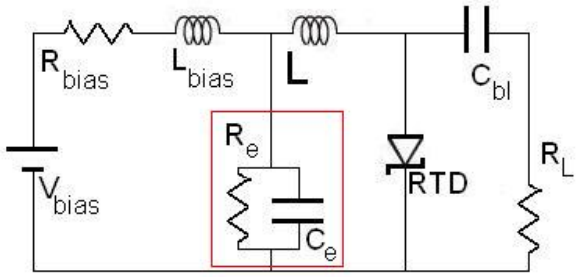


FIGURE 5.1: Single device RTD–based oscillator.

of a dedicated design, are:

- R_e, external (decoupling) resistor;
- C_e, external (decoupling) capacitor;
- L, inductance of the line between the decoupling circuit and the RTD.

When the device is biased in the NDR region using the set up of fig. 5.1, the circuit oscillates if the net resistance in the circuit is negative. In order to eliminate bias oscillations but realize high frequency oscillations, a shunt resistor R_e is used in the oscillator circuit set up. In section 4.1.5 this condition was given:

$$R_e < R_d \tag{5.1}$$

where R_d is the internal resistance of the RTD when it is biased in the NDR region.

[1] The core of the stabilisation method consists in the application of the Van derPol Equation to the design of the oscillator [33, 56].

In [56] and [49] a lower limit for R_e design is presented:

$$\frac{L}{R_d C_n} \leq R_e \tag{5.2}$$

With these two constraints the external resistance would stop the parasitic oscillations without interfering with the signal of the oscillator.

The capacitor C_e has to be large enough to act as an RF short circuit; it is designed in function of the desired oscillating frequency, F_0 and the reactance X by:

$$C_e > \frac{1}{2 \cdot \pi \cdot f_0 \cdot X} \tag{5.3}$$

Power Combining Design

The power combining technique consists in the use of two (or more) RTDs whose output power will be 'added-up' to increase the total output power of the electronic source.

The core idea exploits the property that the topologies in figures 5.1 and 5.2

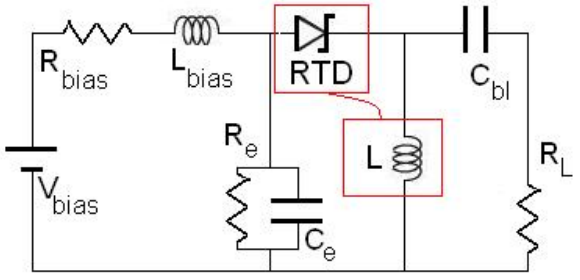


FIGURE 5.2: New topology for single NDR device oscillator with DC stabilizer [33]; red boxes highlight the novel idea: RTD and inductor L are swapped in positions with respect of fig. 5.1; the bias parasitic elements of both side have been omitted.

share the same RF equivalent circuit (fig. 5.4(a)) and they are both free from low-frequency parasitic oscillation [33, 49, 56, 62].

The key advantages of the new topology is that the sub-circuit *RTD and decoupling circuit* can be replicated into the new oscillator design as much times as possible, still keeping the following important aspects:

- RF grounding of the side of the RTD where DC supply is fed;
- RF output from the opposite side;
- dedicated decoupling circuitry per RTD.

This is illustrated in figure 5.3 where it is possible to identify different parts: the the sub–circuit *RTD, RC–docoupling and Bias*, the DC–block (C_{bl}) and the load, the external inductor (L) that will set the oscillating frequency alond with the internal capacitance of the RTD, C_n

$$f_0 = 1/2\pi\sqrt{L \cdot C_n} \tag{5.4}$$

The equivalent circuits for both (single– and double– RTD) oscillators are repre-

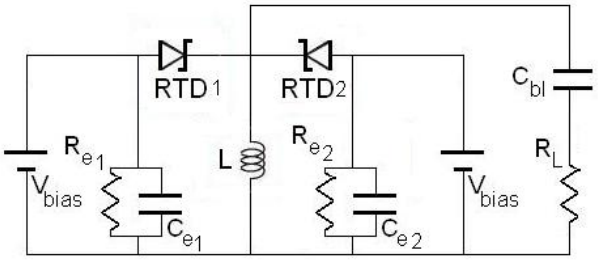
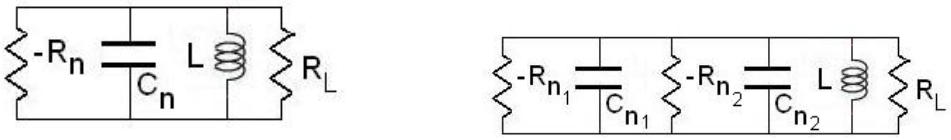


FIGURE 5.3: Two NDR device oscillator topology [49, 56].

sented in fig. 5.4: as for fig. 4.13, the V_{bias} is opencircuited, R_{bias} and L_{bias} are omitted because smaller than the other elements, external capacitors (C_{e1} and C_{e2}) and C_{bl} are opencircuited. In the equivalent circuits are taken in consideration:

the internal elements of the RTD ($-R_n$ and C_n), the external inductor (L) and the applied load (R_L). This is the only way to increase the internal capacitance of the RTD is increased without spoiling the NDR performances[2] in order to improve the oscillating frequency (eq. 5.4).



(a) RF equivalent circuit for 1 RTD-based oscillator.

(b) RF equivalent circuit for 2 RTD-based oscillator.

FIGURE 5.4: RF equivalent circuit for 1 (a) and 2 (b) RTD-based oscillators.

[2]When two or more devices are put together in a parallell configuration, the devices adds up thus reducing the negative differential resistance of the whole circuit making it more difficult to suppress the low frequency parasitic oscillations.

Appendix A

Technical Specifications for Photolithographic Process

A.1 HBT fabrication process

Sample cleaning process

Ultrasonic bath in acetone for 5 min.

Ultrasonic bath in acetone for 5 min.

Ultrasonic bath in acetone for 5 min.

Rinse in RO water.

Blow dry with N2.

Alignment marks and emitter contacts

Clean sample.

Spin S1818 at 10 000 rpm for 120 sec.

Bake on hotplate for 90 sec at 120 ℃.

Pre–develop with 1:1 Microposit Developer Concentrate: H_2O for 75 sec.

Expose using MA6 for 5.0 sec.

Develop with 1:1 Microposit Developer Concentrate: H_2O for 75 sec.

Rinse in IPA.

Rinse in H_2O.

Blow dry with N2.

Ash at 40W for 1 min.

De–oxidise in 5:1 H_2O:HCl for 30 sec.

Rinse in H_2O for 30 sec.

Blow dry with N2.

Deposit Ti/Pd/Au (20/30/80) ohmic contact using electron beam metal evaporator.

Soak in acetone for 2 hours at 50 ℃.

Transfer to IPA.

Blow dry with N2.

Emitter mesa

Clean sample.

Spin S1818 at 500 rpm for 5 sec.

Spin S1818 at 4 000 rpm for 30 sec.

Bake on hotplate for 90 sec at 120 ℃.

Expose using MA6 for 5.0 sec.

Develop with 1:1 Microposit Developer Concentrate: H_2O for 75 sec.

Rinse in IPA.

Rinse in H_2O.

Blow dry with N2.

Ash at 40W for 1 min.

Post–bake on hotplate for 90 sec at 120 ℃.

De–oxidise in 5:1 H_2O:HCl for 30 sec.

Agitation in etching solutions.

30 sec in 1:1:38 – H_2O:H_2O_2:H_3PO_4

13 sec in 1:3 – H_2O:HCl

Rinse in RO–H_2O.

Blow dry in N2.

Base contacts

Clean sample.

Spin S1818 at 10 000 rpm for 120 sec.

Bake on hotplate for 90 sec at 120 ℃.

Pre–develop with 1:1 Microposit Developer Concentrate: H_2O for 75 sec.

Expose using MA6 for 5.0 sec.

Develop with 1:1 Microposit Developer Concentrate: H_2O for 75 sec.

Rinse in IPA.

Rinse in H_2O.

Blow dry with N2.

Ash at 40W for 1 min.

De–oxidise in 5:1 H_2O:HCl for 30 sec.

Rinse in H_2O for 30 sec.

Blow dry with N2.

Deposit Ti/Pd/Au (20/30/80) ohmic contact using electron beam metal evaporator.

Soak in acetone for 2 hours at 50 ℃.

Transfer to IPA.

Blow dry with N2.

Base and collector mesa

Clean sample.

Spin S1818 at 500 rpm for 5 sec.

Spin S1818 at 4 000 rpm for 30 sec.

Bake on hotplate for 90 sec at 120 ℃.

Expose using MA6 for 5.0 sec.

Develop with 1:1 Microposit Developer Concentrate: H_2O for 75 sec.

Rinse in IPA.

Rinse in H_2O.

Blow dry with N2.

Ash at 40W for 1 min.

Post–bake on hotplate for 90 sec at 120 ℃.

De–oxidise in 5:1 H_2O:HCl for 30 sec.

Agitation in etching solutions.

4.5 min in 1:1:38 – H_2O:H_2O_2:H_3PO_4

5 sec in 1:3 – H_2O:HCl

Rinse in RO–H_2O.

Collector contacts

Clean sample.

Spin S1818 at 10 000 rpm for 120 sec.

Bake on hotplate for 90 sec at 120 ℃.

Pre–develop with 1:1 Microposit Developer Concentrate: H_2O for 75 sec.

Expose using MA6 for 5.0 sec.

Develop with 1:1 Microposit Developer Concentrate: H_2O for 75 sec.

Rinse in IPA.

Rinse in H_2O.

Blow dry with N2.

Ash at 40W for 1 min.

De–oxidise in 5:1 H_2O:HCl for 30 sec.

Rinse in H_2O for 30 sec.

Blow dry with N2.

Deposit Ti/Pd/Au (20/30/80) ohmic contact using electron beam metal evaporator.

Soak in acetone for 2 hours at 50 ℃.

Transfer to IPA.

Blow dry with N2.

Sub–collector mesa

Clean sample.

Spin S1818 at 500 rpm for 5 sec.

Spin S1818 at 4 000 rpm for 30 sec.

Bake on hotplate for 90 sec at 120 ℃.

Expose using MA6 for 5.0 sec.

Develop with 1:1 Microposit Developer Concentrate: H_2O for 75 sec.

Rinse in IPA.

Rinse in H_2O.

Blow dry with N2.

Ash at 40W for 1 min.

Post–bake on hotplate for 90 sec at 120 ℃.

De–oxidise in 5:1 H_2O:HCl for 30 sec.

Agitation in etching solutions.

2 min in 1:1:38 – H_2O:H_2O_2:H_3PO_4

20 sec in 1:3 – H_2O:HCl

Rinse in RO–H_2O.

Insulation

Clean sample.

Spin Polyimide at 1 000 rpm for 30 sec.

Spin Polyimide at 4 000 rpm for 60 sec.

Bake in oven for 40 min at 120 ℃.

Spin S1818 at 500 rpm for 5 sec.

Spin S1818 at 4 000 rpm for 30 sec.

Bake on hotplate for 90 sec at 120 ℃.

Expose using MA6 for 10.0 sec.

Develop with TMAH for 15 sec.

Soak in acetone for 2 hours at 50 ℃.

Transfer to IPA.

Blow dry with N2.

Bond pads

Clean sample.

Spin S1818 at 10 000 rpm for 120 sec.

Bake on hotplate for 90 sec at 120 ℃.

Pre–develop with 1:1 Microposit Developer Concentrate: H_2O for 75 sec.

Expose using MA6 for 5.0 sec.

Develop with 1:1 Microposit Developer Concentrate: H_2O for 75 sec.

Rinse in IPA.

Rinse in H_2O.

Blow dry with N2.

Ash at 40W for 1 min.

De–oxidise in 5:1 H_2O:HCl for 30 sec.

Rinse in H_2O for 30 sec.

Blow dry with N2.

Deposit Ti/Au (50/200) ohmic contact using electron beam metal evaporator.

Soak in acetone for 2 hours at 50 ℃.

Transfer to IPA.

Blow dry with N2.

A.2 RTD fabrication process

Sample cleaning process

Ultrasonic bath in acetone for 5 min.

Ultrasonic bath in acetone for 5 min.

Ultrasonic bath in acetone for 5 min.

Rinse in RO water.

Blow dry with N2.

Alignment marks and emitter contacts

Clean sample.

Spin S1818 at 10 000 rpm for 120 sec.

Bake on hotplate for 90 sec at 120 ℃.

Pre–develop with 1:1 Microposit Developer Concentrate: H_2O for 75 sec.

Expose using MA6 for 5.0 sec.

Develop with 1:1 Microposit Developer Concentrate: H_2O for 75 sec.

Rinse in IPA.

Rinse in H_2O.

Blow dry with N2.

Ash at 40W for 1 min.

De–oxidise in 5:1 H_2O:HCl for 30 sec.

Rinse in H_2O for 30 sec.

Blow dry with N2.

Deposit Ti/Pd/Au (20/30/80) ohmic contact using electron beam metal evaporator.

Soak in acetone for 2 hours at 50 ℃.

Transfer to IPA.

Blow dry with N2.

Emitter mesa

Clean sample.

Spin S1818 at 500 rpm for 5 sec.

Spin S1818 at 4 000 rpm for 30 sec.

Bake on hotplate for 90 sec at 120 ℃.

Expose using MA6 for 5.0 sec.

Develop with 1:1 Microposit Developer Concentrate: H_2O for 75 sec.

Rinse in IPA.

Rinse in H_2O.

Blow dry with N2.

Ash at 40W for 1 min.

Post–bake on hotplate for 90 sec at 120 ℃.

De–oxidise in 5:1 H_2O:HCl for 30 sec.

Agitation in etching solutions.

3.5 min in 1:1:38 – H_2O:H_2O_2:H_3PO_4

5 sec in 1:3 – H_2O:HCl

Rinse in RO–H_2O.

Blow dry in N2.

Collector contacts

Clean sample.

Spin S1818 at 10 000 rpm for 120 sec.

Bake on hotplate for 90 sec at 120 ℃.

Pre–develop with 1:1 Microposit Developer Concentrate: H_2O for 75 sec.

Expose using MA6 for 5.0 sec.

Develop with 1:1 Microposit Developer Concentrate: H_2O for 75 sec.

Rinse in IPA.

Rinse in H_2O.

Blow dry with N2.

Ash at 40W for 1 min.

De–oxidise in 5:1 H_2O:HCl for 30 sec.

Rinse in H_2O for 30 sec.

Blow dry with N2.

Deposit Ti/Pd/Au (20/30/80) ohmic contact using electron beam metal evaporator.

Soak in acetone for 2 hours at 50 ℃.

Transfer to IPA.

Blow dry with N2.

Collector mesa

Clean sample.

Spin S1818 at 500 rpm for 5 sec.

Spin S1818 at 4 000 rpm for 30 sec.

Bake on hotplate for 90 sec at 120 ℃.

Expose using MA6 for 5.0 sec.

Develop with 1:1 Microposit Developer Concentrate: H_2O for 75 sec.

Rinse in IPA.

Rinse in H_2O.

Blow dry with N2.

Ash at 40W for 1 min.

Post–bake on hotplate for 90 sec at 120 ℃.

De–oxidise in 5:1 H_2O:HCl for 30 sec.

Agitation in etching solutions.

4.5 min in 1:1:38 – H_2O:H_2O_2:H_3PO_4

Rinse in RO–H_2O.

Resistor deposition

Clean sample.

Spin S1818 at 10 000 rpm for 120 sec.

Bake on hotplate for 90 sec at 120 ℃.

Pre–develop with 1:1 Microposit Developer Concentrate: H_2O for 75 sec.

Expose using MA6 for 5.0 sec.

Develop with 1:1 Microposit Developer Concentrate: H_2O for 75 sec.

Rinse in IPA.

Rinse in H_2O.

Blow dry with N2.

Ash at 40W for 1 min.

De–oxidise in 5:1 H_2O:HCl for 30 sec.

Rinse in H_2O for 30 sec.

Blow dry with N2.

Deposit nICr (33) ohmic contact using electron beam metal evaporator.

Soak in acetone for 2 hours at 50 ℃.

Transfer to IPA.

Blow dry with N2.

Insulation

Clean sample.

Spin Polyimide at 1 000 rpm for 30 sec.

Spin Polyimide at 4 000 rpm for 60 sec.

Bake in oven for 40 min at 120 ℃.

Spin S1818 at 500 rpm for 5 sec.

Spin S1818 at 4 000 rpm for 30 sec.

Bake on hotplate for 90 sec at 120 ℃.

Expose using MA6 for 10.0 sec.

Develop with TMAH for 15 sec.

Soak in acetone for 2 hours at 50 ℃.

Transfer to IPA.

Blow dry with N2.

Bond pads

Clean sample.

Spin S1818 at 10 000 rpm for 120 sec.

Bake on hotplate for 90 sec at 120 ℃.

Pre–develop with 1:1 Microposit Developer Concentrate: H_2O for 75 sec.

Expose using MA6 for 5.0 sec.

Develop with 1:1 Microposit Developer Concentrate: H_2O for 75 sec.

Rinse in IPA.

Rinse in H_2O.

Blow dry with N2.

Ash at 40W for 1 min.

De–oxidise in 5:1 H_2O:HCl for 30 sec.

Rinse in H_2O for 30 sec.

Blow dry with N2.

Deposit Ti/Au (50/200) ohmic contact using electron beam metal evaporator.

Soak in acetone for 2 hours at 50 ℃.

Transfer to IPA.

Blow dry with N2.

Bibliography

[1] J. Wiltse, "History of Millimeter and Submillimeter Waves," *IEEE Transactions on Microwave Theory and Techniques*, vol. 32, pp. 1118 – 1127, sept. 1984.

[2] Y. Watanabe, K. Kawase, T. Ikari, H. Ito, Y. Ishikawa, and H. Minamide, "Component analysis of chemical mixtures using terahertz spectroscopic imaging," *Optics Communications*, vol. 234, no. 16, pp. 125 – 129, 2004.

[3] R. Nitsche, R. Titz, and E. Biebl, "A low-noise 2.5 thz heterodyne receiver with tunable reflector antenna for atmospheric OH-spectroscopy," in *IEEE MTT–S International Microwave Symposium Digest*, vol. 3, pp. 1727 – 1730, jun. 1998.

[4] J. Waters, "Submillimeter-wavelength heterodyne spectroscopy and remote sensing of the upper atmosphere," *Proceedings of the IEEE*, vol. 80, pp. 1679 – 1701, nov. 1992.

[5] T. Phillips and J. Keene, "Submillimeter astronomy heterodyne spectroscopy," *Proceedings of the IEEE*, vol. 80, pp. 1662 – 1678, nov. 1992.

[6] P. Siegel, "Terahertz technology," *IEEE Transactions on Microwave Theory and Techniques,*, vol. 50, pp. 910 – 928, mar. 2002.

[7] P. Siegel, "Terahertz technology in biology and medicine," *IEEE Transactions on Microwave Theory and Techniques*, vol. 52, pp. 2438 – 2447, oct. 2004.

[8] J. Nishizawa, T. Sasaki, K. Suto, T. Yamada, T. Tanabe, T. Tanno, T. Sawai, and Y. Miura, "THz imaging of nucleobases and cancerous tissue using a GaP THz–wave generator," *Optics Communications*, vol. 244, no. 16, pp. 469 – 474, 2005.

[9] S. Wang, B. Ferguson, D. Abbott, and X.-C. Zhang, "T-ray Imaging and Tomography," *Journal of Biological Physics*, vol. 29, pp. 247–256, 2003.

[10] N. Karpowicz, H. Zhong, C. Zhang, K.-I. Lin, J.-S. Hwang, J. Xu, and X.-C. Zhang, "Compact continuous-wave subterahertz system for inspection applications," *Applied Physics Letters*, vol. 86, no. 5, 2005.

[11] R. Appleby and R. Anderton, "Millimeter-wave and submillimeter-wave imaging for security and surveillance," *Proceedings of the IEEE*, vol. 95, pp. 1683–1690, aug. 2007.

[12] H.-B. Liu, H. Zhong, N. Karpowicz, Y. Chen, and X.-C. Zhang, "Terahertz Spectroscopy and Imaging for Defense and Security Applications," *Proceedings of the IEEE*, vol. 95, pp. 1514 – 1527, aug. 2007.

[13] T. Nagatsuma, A. Hirata, M. Harada, H. Ishii, K. Machida, T. Minotani, H. Ito, T. Kosugi, and T. Shibata, "Millimeter–wave photonic integrated

circuit technologies for high-speed wireless communications applications," in *IEEE International Solid-State Circuits Conference*, pp. 448 – 449, feb. 2004.

[14] B. Deng, C. Domier, A. Donne, K. Lee, J. Luhmann, N.C., E. Mazzucato, T. Munsat, H. Park, and M. van de Pol, "THz techniques in plasma diagnostics," in *2002 IEEE MTT-S International Microwave Symposium Digest*, vol. 3, pp. 1587–1590, 2002.

[15] A. Farster, J. Stock, S. Montanari, M. I. Lepsa, and H. Lath, "Fabrication and characterisation of GaAs Gunn Diode chips for applications at 77 GHz in automotive industry," *Sensors*, vol. 6, no. 4, pp. 350–360, 2006.

[16] S. Kumar, "Recent progress in terahertz Quantum Cascade Lasers," *IEEE Journal of Selected Topics in Quantum Electronics,*, vol. 17, pp. 38–47, jan.-feb. 2011.

[17] S. Fathololoumi, *Terahertz Quantum Cascade Lasers: towards high performance operation*. PhD thesis, Electrical and Computer Engineering, University of Waterloo, 2011.

[18] M. Asada and S. Suzuki, "Terahertz oscillators using electron devices: an approach with resonant tunneling diodes," *IEICE Electronics Express*, vol. 8, no. 14, pp. 1110–1126, 2011.

[19] M. Rodwell, M. Urteaga, T. Mathew, D. Scott, D. Mensa, Q. Lee, J. Guthrie, Y. Betser, S. Martin, R. Smith, S. Jaganathan, S. Krishnan, S. Long, R. Pullela, B. Agarwal, U. Bhattacharya, L. Samoska, and M. Dahlstrom, "Submicron scaling of hbts," *Electron Devices, IEEE Transactions on*, vol. 48,

pp. 2606–2624, nov 2001.

[20] V. Radisic, D. Sawdai, D. Scott, W. Deal, L. Dang, D. Li, J. Chen, A. Fung, L. Samoska, T. Gaier, and R. Lai, "Demonstration of a 311-ghz fundamental oscillator using InP hbt technology," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 55, pp. 2329–2335, nov. 2007.

[21] V. Radisic, X. Mei, W. Deal, W. Yoshida, P. Liu, J. Uyeda, M. Barsky, L. Samoska, A. Fung, T. Gaier, and R. Lai, "Demonstration of sub-millimeter wave fundamental oscillators using 35-nm InP hemt technology," *Microwave and Wireless Components Letters, IEEE*, vol. 17, pp. 223–225, march 2007.

[22] S. Kudszus, T. Berceli, A. Tessmann, M. Neumann, and W. Haydl, "W-band hemt-oscillator mmics using subharmonic injection locking," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 48, pp. 2526–2532, dec 2000.

[23] J. Siles and J. Grajal, "Physics-based design and optimization of schottky diode frequency multipliers for terahertz applications," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 58, pp. 1933–1942, july 2010.

[24] G. Chattopadhyay, E. Schlecht, J. Ward, J. Gill, H. Javadi, F. Maiwald, and I. Mehdi, "An all–solid–state broad-band frequency multiplier chain at 1500 GHz," *IEEE Transactions on Microwave Theory and Techniques*, vol. 52, pp. 1538–1547, may 2004.

[25] M. Ino, T. Ishibashi, and M. Ohmori, "C.w. oscillation with p+-p-n+ silicon impatt diodes in 200 ghz and 300 ghz bands," *Electronics Letters*, vol. 12, pp. 148–149, 18 1976.

[26] H. Heisele, *Terahertz sources and systems*. Chateau de Bonas, France: NATO Science Series, proceedings of the nato advenced research workshop on teraherts sources and systems ed., 2000.

[27] H. Eisele, "State of the art and future of electronic sources at terahertz frequencies," *Electronics Letters*, vol. 46, pp. s8–s11, 23 2010.

[28] J. Gunn, "Microwave oscillations of current in III–V semiconductors," *Solid State Communications*, vol. 88, pp. 883–886, 1993.

[29] M. Pobl, C. Dalle, J. Freuer, and W. Harth, "Cw mm-wave GaAs tunnett diode," *Electronics Letters*, vol. 26, pp. 1540–1542, aug. 1990.

[30] H. Eisele, "355 ghz oscillator with GaAs tunnett diode," *Electronics Letters*, vol. 41, pp. 329–331, march 2005.

[31] K. Stephan, E. Brown, C. Parker, W. Goodhue, C. Chen, and T. Sollner, "Resonant-tunnelling diode oscillator using a slot-coupled quasioptical open resonator," *Electronics Letters*, vol. 27, pp. 647–649, april 1991.

[32] M. Feiginov, C. Sydlo, O. Cojocari, and P. Meissner, "Resonant-tunnelling-diode oscillators operating at frequencies above 1.1 thz," *Applied Physics Letters*, vol. 99, no. 23, p. 233506, 2011.

[33] L. Wang and E. Wasige, "A design procedure for tunnel diode microwave oscillators," in *Microwave and Millimeter Wave Technology, 2008. ICMMT 2008. International Conference on*, vol. 2, pp. 832–834, april 2008.

[34] Y. Kwon, G. Ng, D. Pavlidis, R. Lai, T. Brock, J. Castagne, and N. Linh, "High efficiency monolithic Ka-band oscillators using InAlAs/InGaAs hemts," in *Gallium Arsenide Integrated Circuit (GaAs IC) Symposium, 1991*, pp. 263–266, oct. 1991.

[35] C.-H. Lee, A. Sutono, and J. Laskar, "Development of a high-power and high-efficiency HBT MMIC VCO," in *Radio and Wireless Conference, RAWCON 2001.*, pp. 157–160, 2001.

[36] A. van der Ziel, "Noise in solid-state devices and lasers," *Proceedings of the IEEE*, vol. 58, pp. 1178–1206, aug. 1970.

[37] F. Maiwald, S. Martin, J. Bruston, A. Maestrini, T. Crawford, and P. Siegel, "2.7 thz waveguide tripler using monolithic membrane diodes," in *Microwave Symposium Digest, 2001 IEEE MTT-S International*, vol. 3, pp. 1637–1640 vol.3, 2001.

[38] T. Ishibashi, M. Ino, T. Makimura, and M. Ohmori, "Liquid-nitrogen-cooled submillimetre-wave silicon impatt diodes," *Electronics Letters*, vol. 13, pp. 299–300, 12 1977.

[39] H. Eisele, "InP gunn devices for 400-425 ghz," *Electronics Letters*, vol. 42, pp. 358–359, march 2006.

[40] H. Eisele, "Second-harmonic power extraction from InP gunn devices with more than 1 mw in 260-320 ghz frequency range," *Electronics Letters*, vol. 34, pp. 2412–2413, dec 1998.

[41] H. Eisele, "480 ghz oscillator with an InP gunn device," *Electronics Letters*, vol. 46, pp. 422–423, 18 2010.

[42] H. Eisele, A. Rydberg, and G. Haddad, "Recent advances in the performance of InP gunn devices and GaAs tunnett diodes for the 100-300-ghz frequency range and above," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 48, pp. 626–631, apr 2000.

[43] L. L. Chang, L. Esaki, and R. Tsu, "Resonant tunneling in semiconductor double barriers," *Applied Physics Letters*, vol. 24, no. 12, pp. 593–595, 1974.

[44] E. R. Brown, J. R. Söderström, C. D. Parker, L. J. Mahoney, K. M. Molvar, and T. C. McGill, "Oscillations up to 712 ghz in inas/alsb resonant-tunneling diodes," *Applied Physics Letters*, vol. 58, no. 20, pp. 2291–2293, 1991.

[45] A. Teranishi, K. Shizuno, S. Suzuki, M. Asada, H. Sugiyama, and H. Yokoyama, "Fundamental oscillation up to 1.08thz in resonant tunneling diodes with high-indium-composition transit layers for reduction of transit delay," *IEICE Electronics Express*, vol. 9, no. 5, pp. 385–390, 2012.

[46] L. Varani, C. Palermo, J. Millithaler, J. Vaissire, E. Starikov, P. Shiktorov, V. Gruinskis, J. Mateos, S. Prez, D. Pardo, and T. Gonzlez, "Numerical modeling of terahertz electronic devices," *Journal of Computational Electronics*, vol. 5, pp. 71–77, 2006. 10.1007/s10825-006-8822-5.

[47] C. Kim and A. Brandli, "High–frequency high–power operation of tunnel diodes," *IRE Transactions on Circuit Theory*, vol. 8, pp. 416–425, dec. 1961.

[48] "Bias circuit instabilities and their effect on the d.c. current-voltage characteristics of double-barrier resonant tunneling diodes," *Solid-State Electronics*, vol. 34, no. 2, pp. 149–156, 1991.

[49] L. Wang and E. Wasige, "Tunnel diode microwave oscillators employing a novel power combining circuit topology," in *2010 European Microwave Conference*, pp. 1154–1157, sept. 2010.

[50] E. Brown, C. Parker, K. Molvar, and K. Stephan, "A quasioptically stabilized resonant tunneling diode oscillator for the millimeter– and submillimeter–wave regions," *IEEE Transactions on Microwave Theory and Techniques*, vol. 40, pp. 846–850, may 1992.

[51] M. Reddy, R. Yu, H. Kroemer, M. Rodwell, S. Martin, R. Muller, and R. Smith, "Bias stabilization for Resonant Tunnel Diode oscillators," *IEEE Microwave and Guided Wave Letters*, vol. 5, pp. 219–221, jul 1995.

[52] J. Wallmark and A. Dansky, "Nonlinear biasing resistors for microwave tunnel-diode oscillators (correspondence)," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 11, pp. 260–262, jul 1963.

[53] M. Reddy, *Schottky–collector Resonant Tunnelling Diodes for sub–Millimeter–Wave applications*. PhD thesis, Department of Electrical and Computer Engineering, University of California, Santa Barbara, 1998.

[54] P. Chahal, F. Morris, and G. Frazier, "50 GHz resonant tunneling diode relaxation oscillator," in *62nd Device Research Conference*, vol. 1, pp. 241–242, jun. 2004.

[55] M. Reddy, S. Martin, A. Molnar, R. Muller, R. Smith, P. Siegel, M. Mondry, M. Rodwell, H. Kroemer, and J. Allen, S.J., "Monolithic Schottky-collector Resonant Tunnel Diode oscillator arrays to 650 GHz," *IEEE Electron Device Letters*, vol. 18, pp. 218–221, may 1997.

[56] L. Wang, *Reliable design of tunnel diode and resonant tunnelling diode based microwave sources*. PhD thesis, School of Engineering, UNIVERSITY OF GLASGOW, 2011.

[57] G. Haddad, J. East, and H. Eisele, "Two-terminal active devices for terahertz sources.," *International Journal of High Speed Electronics Systems*, vol. 13, no. 2, p. 395, 2003.

[58] O. Boric-Lubecke, D.-S. Pan, and T. Itoh, "Dc instability of the series connection of tunneling diodes," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 44, pp. 936–943, jun 1996.

[59] O. Boric-Lubecke, D.-S. Pan, and T. Itoh, "Rf excitation of an oscillator with several tunneling devices in series," *Microwave and Guided Wave Letters, IEEE*, vol. 4, pp. 364–366, nov. 1994.

[60] K. Stephan, S.-C. Wong, E. Brown, K. Molvar, A. Calawa, and M. Manfra, "5mw parallel-connected resonant-tunnelling diode oscillator," *Electronics Letters*, vol. 28, pp. 1411–1412, july 1992.

[61] H. Cantu and W. Truscott, "Injection-locking and power combining with double barrier resonant tunnelling diodes," *Electronics Letters*, vol. 37, pp. 1264–1265, sept. 2001.

[62] L. Wang, J. Figueiredo, C. Ironside, and E. Wasige, "Dc characterization of tunnel diodes under stable non-oscillatory circuit conditions," *Electron Devices, IEEE Transactions on*, vol. 58, pp. 343–347, feb. 2011.

[63] B. Sheinman, E. Wasige, M. Rudolph, R. Doerner, V. Sidorov, S. Cohen, and D. Ritter, "A peeling algorithm for extraction of the hbt small-signal equivalent circuit," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 50, pp. 2804–2810, dec 2002.

[64] S. Taking, D. MacFarlane, and E. Wasige, "Aln/gan mos-hemts with thermally grown passivation," *Electron Devices, IEEE Transactions on*, vol. 58, pp. 1418–1424, may 2011.

[65] Rohm and Haas Company, Marlborough, MA, USA, *Microposit S1800 G2 series photoresists*, 2008.

[66] R. Oxland, S. Chang, X. Li, S. Wang, G. Radhakrishnan, W. Priyantha, M. van Dal, C. Hsieh, G. Vellianitis, G. Doornbos, K. Bhuwalka, B. Duriez, I. Thayne, R. Droopad, M. Passlack, C. Diaz, and Y. Sun, "An ultralow-resistance ultrashallow metallic source/drain contact scheme for iii–v NMOS," *IEEE Electron Device Letters*, vol. 33, pp. 501–503, april 2012.

[67] G. Reeves and H. Harrison, "Obtaining the specific contact resistance from transmission line model measurements," *Electron Device Letters, IEEE*, vol. 3, pp. 111–113, may 1982.

[68] H. Berger, "Contact resistance on diffused resistors," in *1969 IEEE International Solid–State Cicruits Conference.*, vol. XII, pp. 160–161, feb 1969.

[69] M. Kim, C.-Y. Kim, and Y.-S. Kwon, "InP/InGaAs heterojunction bipolar transistors with low-resistance contact on heavily doped InP emitter layer," *Applied Physics Letters*, vol. 84, no. 15, pp. 2934–2936, 2004.

[70] A. Katz, B. E. Weir, and W. C. Dautremont-Smith, "Au/pt/ti contacts to p-in[sub 0.53]ga[sub 0.47]as and n-InP layers formed by a single metallization common step and rapid thermal processing," *Journal of Applied Physics*, vol. 68, no. 3, pp. 1123–1128, 1990.

[71] U. Singisetti, A. Crook, E. Lind, J. Zimmerman, M. Wistey, A. Gossard, and M. Rodwell, "Ultra-low resistance ohmic contacts to InGaAs/InP," in *65th Annual Device Research Conference*, pp. 149–150, june 2007.

[72] J.-L. Lee, Y.-T. Kim, and J. Y. Lee, "Microstructural evidence on direct contact of Au/Ge/Ni/Au ohmic metals to InGaAs channel in pseudomorphic high electron mobility transistor with undoped cap layer," *Applied Physics Letters*, vol. 73, no. 12, pp. 1670–1672, 1998.

[73] D. Caffin, C. Besombes, J. F. Bresse, P. Legay, G. L. Roux, G. Patriarche, and P. Launay, "Base metallization stability in InP/InGaAs heterojunction bipolar transistors and its influence on leakage currents," *Journal of Vacuum Science Technology B: Microelectronics and Nanometer Structures*, vol. 15, no. 4, pp. 854–861, 1997.

[74] E. Chor, R. Malik, R. Hamm, and R. Ryan, "Metallurgical stability of ohmic contacts on thin base InP/InGaAs/InP HBTs," *IEEE Electron Device Letters*, vol. 17, pp. 62–64, feb. 1996.

[75] W. Chong, E. Chor, C. Heng, and S. Chua, "Pd/Ti/Au–based ohmic contacts to p– and n–doped InGaAs," in *IEEE International Symposium on Compound Semiconductors*, pp. 171–174, sep. 1998.

[76]

[77] V. Craig, G. Burns, R. Oxland, and E. Wasige, "InP hbt technology for high frequency applications," in *High Frequency Postgraduate Student Colloquium, 2005*, pp. 47–49, sept. 2005.

[78] M. Rodwell, M. Le, and B. Brar, "InP bipolar ics: Scaling roadmaps, frequency limits, manufacturable technologies," *Proceedings of the IEEE*, vol. 96, pp. 271–286, feb. 2008.

[79] M. Rodwell, E. Lobisser, M. Wistey, V. Jain, A. Baraskar, E. Lind, J. Koo, Z. Griffith, J. Hacker, M. Urteaga, D. Mensa, R. Pierson, and B. Brar, "THz bipolar transistor circuits: technical feasibility, technology development, integrated circuit results," in *IEEE Compound Semiconductor Integrated Circuits Symposium, 2008*, pp. 1–3, oct. 2008.

[80] W. Liu, *Handbook of III–V Heterojunction Bipolar Transistors*. Chichester–Wiley, New York, 1998.

[81] R. L. Anderson, "Germanium-gallium arsenide heterojunctions," *IBM J. Res. Dev.*, vol. 4, pp. 283–287, July 1960.

[82] R. K. Oxland, *An HBT magnetic sensor with integrated 3–Dimesional magnetic structures*. PhD thesis, Faculty of Engineering, University of Glasgow, 2009.

[83] S. Topaloglu, *Process Technology for High Speed InP Based Heterojunction Bipolar Transistors*. PhD thesis, Fakult{at für Ingenieurwissenschaften der Universität Duisburg-Essen— Faculty of Engineering, University of Duisburg–Essen, 2006.

[84] A. Hanson, S. Stockman, and G. Stillman, "InP/$In_{0.53}$/$G_{0.47}$/As heterojunction bipolar transistors with a carbon–doped base grown by MOCVD," *IEEE Electron Device Letters, IEEE*, vol. 13, pp. 504–506, oct. 1992.

[85] R. N. Nottenburg, Y. K. Chen, T. Tanbun-Ek, R. A. Logan, and D. A. Humphrey, "High performance InP/InGaAs heterostructure bipolar transistors grown by metalorganic vapor phase epitaxy," *Applied Physics Letters*, vol. 55, no. 2, pp. 171–172, 1989.

[86] Agilent, *Topic 5: S–parameter simulation and optimization,(version 1.0)*, 2009.

[87] L. Degachi and F. Ghannouchi, "An augmented small–signal HBT model with its analytical based parameter extraction technique," *IEEE Transactions on Electron Devices*, vol. 55, pp. 968–972, april 2008.

[88] L. Degachi and F. Ghannouchi, "Systematic and rigorous extraction method of HBT small–signal model parameters," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, pp. 682–688, feb. 2006.

[89] S. Bousnina, P. Mandeville, A. Kouki, R. Surridge, and F. Ghannouchi, "Direct parameter–extraction method for HBT small-signal model," *IEEE*

Transactions on Microwave Theory and Techniques, vol. 50, pp. 529–536, Feb. 2002.

[90] S. Bousnina, C. Falt, P. Mandeville, A. Kouki, and F. Ghannouchi, "An accurate on-wafer deembedding technique with application to HBT devices characterization," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, pp. 420–424, Feb. 2002.

[91] M. Sotoodeh, L. Sozzi, A. Vinay, A. Khalid, Z. Hu, A. Rezazadeh, and R. Menozzi, "Stepping toward standard methods of small–signal parameter extraction for HBTs," *IEEE Transactions on Electron Devices*, vol. 47, pp. 1139 – 1151, June 2000.

[92] E. Wasige, "Extracting the HBT series parasitics under active bias conditions," *36th European Microwave Conference*, pp. 1267–1270, sept. 2006.

[93] D. Costa, W. Liu, and J. Harris, J.S., "Direct extraction of the AlGaAs/GaAs heterojunction bipolar transistor small–signal equivalent circuit," *IEEE Transactions on Electron Devices*, vol. 38, pp. 2018–2024, sept. 1991.

[94] Agilent, *Genesys 2008.01 documentation set*, 2008.

[95] A. Jarndal and G. Kompa, "A new small-signal modeling approach applied to GaN devices," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, pp. 3440–3448, nov. 2005.

[96] G. Dambrine, A. Cappy, F. Heliodore, and E. Playez, "A new method for determining the FET small–signal equivalent circuit," *IEEE Transactions on Microwave Theory and Techniques*, vol. 36, pp. 1151 – 1159, jul. 1988.

[97] S.-C. Huang, W.-B. Tang, and Y.-M. Hsin, "High–Frequency noise modeling of InGaP/GaAs HBT with base-contact capacitance and AC current crowding effect," *IEEE Electron Device Letters*, vol. 30, no. 11, pp. 1125–1127, 2009.

[98] W.-B. Tang, C.-M. Wang, and Y.-M. Hsin, "A new extraction technique for the complete Small–Signal Equivalent–Circuit model of InGaP/GaAs HBT including base contact impedance and AC current crowding effect," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, pp. 3641–3647, oct. 2006.

[99] J. H. Davies, *The Physics of low–dimensional semiconductors – An introduction.* Cambride University Press, 1st ed., 1998.

[100] P. Harrison, *Quantum Wells, Wires and Dots.* Wiley–Interscience, 2nd ed., 2005.

[101] T. J. Slight, *Integration of a Resonant Tunneling Diode and an optical communications laser.* PhD thesis, Department of Electronics and Electrical Engineering, University of Glasgow, 2006.

[102] J. M. L. Figueiredo, *Optoelectronic properties of Resonant Tunnelling Diodes.* PhD thesis, Departamento de Fisica, Faculdade de Ciencias da Universidade

do Porto— Department of Physics, Faculty of Sciences, University of Porto, 2000.

[103] D. R. Chowdhury, *Experimental study and modelling of AC characteristics of Resonant Tunneling Diodes*. PhD thesis, Fachbereich Elektrotechnik und Informationstechnik, Technischen Universität Darmstadt— Department of Electrical Engineering and Information Technology, Technical University of Darmstadt, 2008.

[104] H. Eisele and G. Haddad, "Two–terminal millimeter–wave sources," *IEEE Transactions on Microwave Theory and Techniques*, vol. 46, pp. 739–746, jun. 1998.