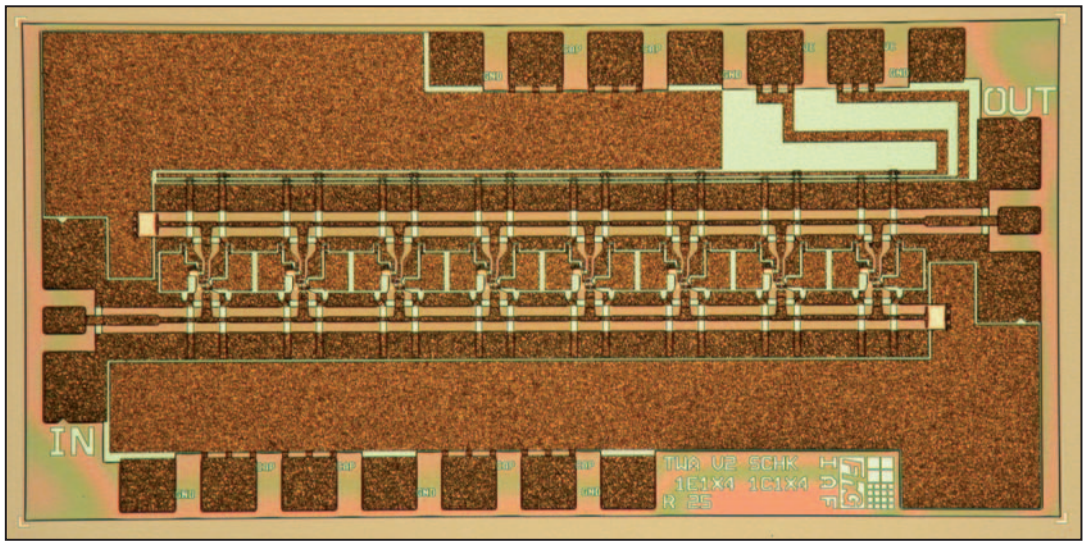
Karl Schneider

Broadband Amplifiers for High Data Rates using InP/InGaAs Double Heterojunction Bipolar Transistors



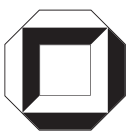
universitätsverlag karlsruhe

Karl Schneider

Broadband Amplifiers for High Data Rates using InP/InGaAs Double Heterojunction Bipolar Transistors

Broadband Amplifiers for High Data Rates using InP/InGaAs Double Heterojunction Bipolar Transistors

von
Karl Schneider



universitätsverlag karlsruhe

Dissertation, Universität Karlsruhe (TH)
Fakultät für Elektrotechnik und Informationstechnik, 2006

Impressum

Universitätsverlag Karlsruhe
c/o Universitätsbibliothek
Straße am Forum 2
D-76131 Karlsruhe
www.uvka.de

Universitätsverlag Karlsruhe 2006
Print on Demand

ISBN 3-86644-021-9

Broadband Amplifiers for High Data Rates using InP/InGaAs Double Heterojunction Bipolar Transistors

Zur Erlangung des akademischen Grades eines

DOKTOR-INGENIEURS

von der Fakultät für
Elektrotechnik und Informationstechnik
der Universität Fridericiana Karlsruhe
genehmigte

DISSERTATION

von

Dipl.-Ing. Karl Schneider

aus Hannover

Tag der mündlichen Prüfung:	16. Januar 2006
Hauptreferent:	Prof. Dr. rer. nat. G. Weimann
Korreferent:	Prof. Dr.-Ing. W. Wiesbeck

Contents

Kurzfassung **1**

Abstract **2**

1 Introduction **3**
1.1 Motivation . . . 3
1.1.1 Scope of the Work . . . 5
1.2 State of the Art . . . 6
1.2.1 Broadband Amplifiers . . . 6
1.2.2 Devices . . . 7

2 InP-based Double Heterojunction Bipolar Transistor **11**
2.1 Operating Principle . . . 12
2.2 Technology . . . 16
2.2.1 Active Devices . . . 16
2.2.2 Passive Structures . . . 18
2.3 Layer Structure Optimization . . . 19
2.3.1 Initial Layer Structure (A) . . . 19
2.3.2 Improved Layer Structure (B) . . . 24
2.3.3 Optimized Layer Structure (C) . . . 28
2.3.4 Summary . . . 33
2.4 Transistor Layout Optimization . . . 34
2.4.1 Emitter Width Design . . . 36
2.4.2 Base Contact Width Design . . . 38
2.4.3 Collector Contact Design . . . 40
2.4.4 Emitter Length Design . . . 40
2.4.5 Summary . . . 43

3 Transistor Models **45**
3.1 Large-Signal Models . . . 46
3.1.1 Gummel-Poon Model . . . 46
3.1.2 Alternative Models . . . 47
3.1.3 UCSD Model . . . 47

3.2 UCSD Model Extraction . . . 49
3.2.1 Diode Parameters . . . 49
3.2.2 Resistances . . . 51
3.2.3 Junction Capacitances . . . 55
3.2.4 Transit and Delay Times . . . 56
3.3 Small-Signal Model . . . 60
3.3.1 Model Extraction . . . 60
3.3.2 Investigation of Transistors . . . 63
3.4 Summary . . . 68

4 Methods of Broadband Amplifier Characterization 69
4.1 Figures of Merit . . . 69
4.1.1 Bandwidth . . . 69
4.1.2 Group Delay . . . 71
4.1.3 Output Power . . . 71
4.2 Small-Signal-Characterization (S-Parameter) . . . 72
4.3 Large-Signal-Characterization . . . 73
4.3.1 Output Power . . . 73
4.3.2 Signal Distortion (Eye-Diagram) . . . 74

5 Realization of Compact Lumped Amplifiers 77
5.1 Design Considerations . . . 77
5.2 Realized Amplifiers . . . 79
5.2.1 Comparison of Lumped Amplifier Approaches . . . 79
5.2.2 High-Gain Amplifier for 40 Gbit/s . . . 82
5.2.3 Low-Power Amplifier for 80 Gbit/s . . . 85
5.3 Output Power Limit . . . 87

6 Distributed Amplifiers for 80 Gbit/s 91
6.1 Design Considerations . . . 92
6.1.1 Gain Cell Design Concepts . . . 93
6.1.2 Attenuation Compensation . . . 93
6.1.3 Coplanar and Microstrip Transmission Lines . . . 96
6.1.4 Transmission Line Termination . . . 97
6.1.5 Linear and Matrix Design . . . 98
6.1.6 Single-ended and Differential Design . . . 98
6.2 Realized Amplifiers . . . 99
6.2.1 Low-Power 100 GHz Bandwidth Amplifier . . . 99
6.2.2 Tunable Amplifier for Loss Compensation . . . 103
6.2.3 High-Power Amplifier for 80 Gbit/s . . . 108
6.3 Summary . . . 114

7 Conclusion and Outlook 115

A Extracted Large-Signal Model Parameters **117**

Bibliography **119**

List of Abbreviations **129**

Acknowledgments **131**

Curriculum Vitae **132**

Kurzfassung

Diese Arbeit beschreibt die Entwicklung von elektrischen Breitbandverstärkern, die für den Einsatz als Modulatortreiber in optischen Nachrichtenübertragungssystemen, die bei einer Datenrate von 80 Gbit/s arbeiten, geeignet sind. Diese Systeme ermöglichen die Datenübertragung bei sehr hohen Bitraten, die die Anforderungen der stetig wachsenden Informationstechnologien, zu denen der Mobilfunk und das Internet gehören, erfüllen können. Die Realisierung der Verstärker wird in einem dreistufigen Entwicklungsprozess erreicht.

Als erstes werden die geometrischen Dimensionen, d.h. die vertikale Schichtstruktur und die lateralen Abmessungen der Einzeltransistoren der verwendeten InP-basierten Doppel-Hetero-Bipolartransistortechnologie optimiert. Dieses ist ein wichtiger Beitrag zur Entwicklung einer Transistortechnologie, die eine Kombination von sehr hoher Geschwindigkeit und hoher Durchbruchspannung bietet, die für die Herstellung von Modulatortreibern für 80 Gbit/s erforderlich ist und die von Silizium-basierten Technologien nicht geleistet werden kann. Experimentelle Untersuchungen verschiedener Transistoren zeigen, dass die vertikalen und laterale Abmessungen entscheidenden Einfluss auf das Transistorverhalten haben. Nach drei Iterationen haben die Transistoren eine Güte erreicht, die die Entwicklung von integrierten Schaltungen für 80 Gbit/s erlaubt.

Im zweiten Entwicklungsschritt werden geeignete Transistormodelle bereitgestellt. Aus mehreren publizierten Modellierungsansätzen, wird der aussichtsreichste übernommen und für die beabsichtigte Anwendung optimiert. Die Gültigkeit der Modelle wird durch den Vergleich von Messung und Simulation einzelner Transistoren und komplexer Schaltungen verifiziert. Die erzeugten Modelle sind nicht nur für die Groß- und Kleinsignalsimulation von analogen Schaltungen, sondern auch für die Entwicklung der digitalen Schaltungen eines Nachrichtenübertragungssystems geeignet.

Schließlich werden mit Hilfe der optimierten Transistoren und Modelle verschiedene Verstärker auf Grundlage unterschiedlicher Konzepte entworfen, hergestellt und mit S-Parameter-, Leistungs- und Augendiagrammmessungen untersucht. Die konzentrierten Verstärker weisen Vorteile wie geringe Leistungsaufnahme, kleine Abmessungen und damit geringe Herstellungskosten auf. Ihre Ausgangsleistung ist jedoch sehr begrenzt. Im Gegensatz dazu bieten die verteilten Verstärker mehr Ausgangsleistung und höhere Bandbreite. Ein auf Leistung optimierter verteilter Verstärker erweist sich als besonders geeignet für die Entwicklung von Modulatortreibern für 80 Gbit/s und gehört zu besten Ergebnissen, die in diesem Bereich bisher veröffentlicht wurden.

Abstract

This work describes the development process of electrical broadband amplifiers, which are suitable as modulator drivers in electrical time division multiplex (ETDM) systems, operating at 80 Gbit/s. Such systems are promising candidates for the next generation of high bit rate data transmission systems, which can satisfy the demands of the continuously growing information technologies, including mobile communications and internet. The realization is accomplished in three major development steps.

First, the vertical dimensions, i.e. the epitaxial layer structure, as well as the lateral dimensions, i.e. the transistor layout, of InP-based Double Heterojunction Bipolar Transistors (DHBT) are optimized. This is an important contribution in the development of an InP-technology, which offers devices with a combination of high speed and breakdown voltage, suitable for modulator driver development at 80 Gbit/s and not achievable with silicon based technologies. Various direct current (DC) and high frequency (HF) measurements of different DHBTs are conducted and the results are related to the devices' geometries. It is shown, that optimized vertical and horizontal dimensions are crucial for realizing high performance transistors. The optimized devices feature state-of-the-art performance allowing the development of integrated circuits for 80 Gbit/s.

Second, transistor models are obtained from extracted device data. From several published models, the most suitable approach is adopted and subsequently refined and optimized to make a compromise between accuracy and complexity. The validity of these models is confirmed by comparing not only measurement and simulation results of transistors, but also of complex circuits. In addition, these models are not only used for small-signal and large-signal simulations in analog circuits, but are equally suitable for developing digital circuits of an ETDM system.

Third, using the optimized transistors and models, several amplifiers representing different design concepts are realized and experimentally evaluated using S-parameter, power and eye diagram measurements. Lumped amplifiers show a low power consumption at 80 Gbit/s operation. An additional advantage is their small chip size and subsequent low costs. However, their output power capabilities are very limited. In contrast, distributed amplifiers are more complex, but offer extended bandwidth and superior output power levels. A state-of-the-art distributed amplifier featuring eight gain cells, each using emitter followers and a cascode configuration at its input and output, respectively, shows superior performance with respect to modulator driver application at 80 Gbit/s. These results are a significant contribution towards the realization of ETDM systems operating at 80 Gbit/s.

Chapter 1

Introduction

1.1 Motivation

By the end of the 20th century and the beginning of the 21st century the telecommunication and information technology started to change society strongly. Today, all areas of society, politics, and economy make use of the possibilities of these technologies and as a consequence the influence on economic processes, for example, is enormous. Also the private user takes more and more advantage of the possibilities of mobile communication and internet. Consequently, the requirements on the telecommunication technologies have changed dramatically, as they have come to play a key role in all social areas and in the

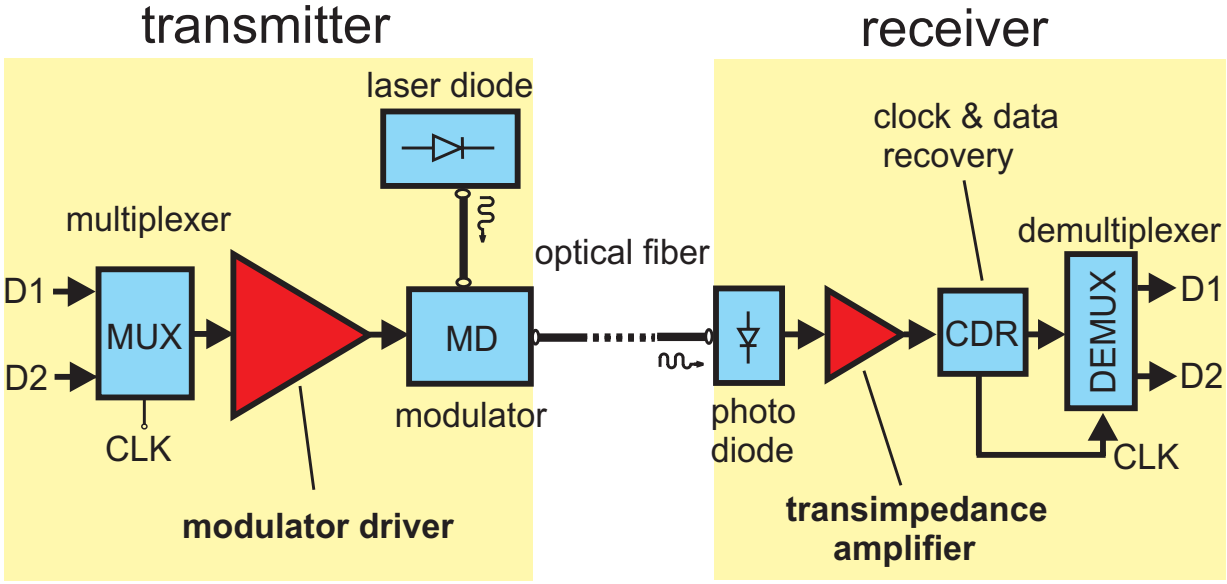
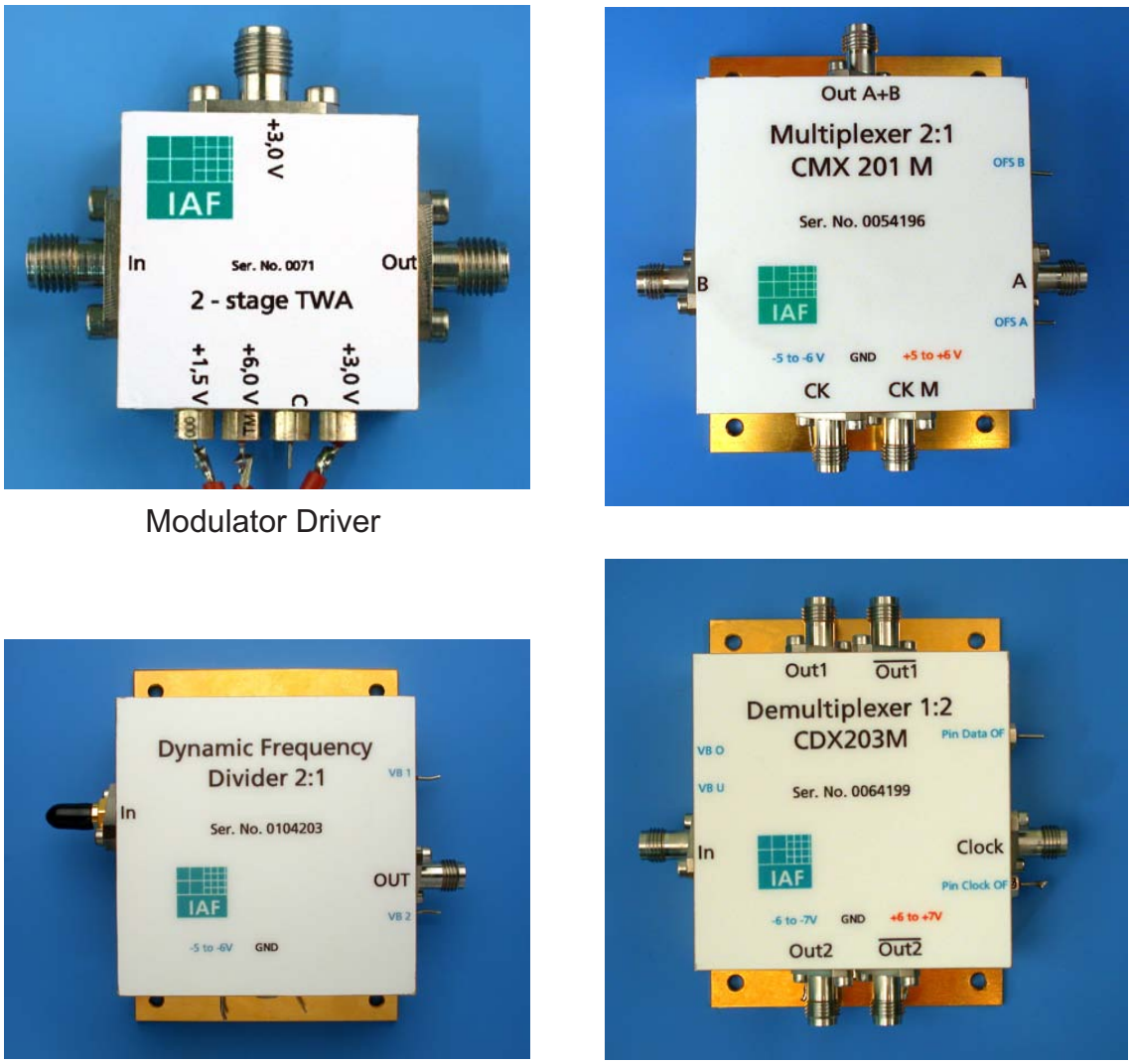


Figure 1.1: Schematic of ETDM system. Two data channels (D1, D2) are multiplexed into a single data stream and transmitted over an optical fiber. At the receiver side the data channels D1 and D2 are recovered.



Modulator Driver

Figure 1.2: Electrical key components for 40 Gbit/s ETDM system. Chips and modules were developed and manufactured at Fraunhofer IAF.

lives of people.

In today's metro and core networks 10 Gbit/s transmission systems are established. Until the year 2001 the speed of development of optical transmission systems was enormous, from below 1 Gbit/s to 2.5 Gbit/s and today's 10 Gbit/s. At that time, all major system vendors promised to provide systems with a per channel data rate of 40 Gbit/s soon. Figure 1.1 illustrates the architecture of an electrical time division multiplex (ETDM) optical transmission system with the necessary electrical and electro-optical components. At Fraunhofer Institute of Applied Solid-State Physics (IAF) major electrical components for data transmission systems operating at 40 Gbit/s were developed by the year 2002 (s. Fig. 1.2). With the downturn of the communication industry the commercial interest in higher data rates nearly stopped. Suddenly, not the best performing system bridging

a distance of several thousands of kilometers was of interest, but a cheap system just satisfying the short-term requirements.

Recently, the situation has become more relaxed again. Some operators now ask for the option to deploy 40 Gbit/s systems, although so far no larger installation of this bit rate has been performed. To be prepared for the next step, the development of optical data transmission systems at even higher bit rates was started. In 2002 as part of the effort, Fraunhofer IAF undertook the task of developing all electrical components needed for an 80 Gbit/s ETDM optical data link.

1.1.1 Scope of the Work

The scope of this thesis is the development of broadband amplifiers suitable as transimpedance amplifiers and modulator drivers in ETDM systems with a data rate of 80 Gbit/s and beyond (Fig. 1.1). This task required not only the design of the electrical circuits, but a great amount of preliminary development steps, which are a prerequisite for circuit design. The necessary work packages, that were solved in the course of this work, are described in the following paragraphs.

First of all a new InP-based heterojunction bipolar transistor (HBT) technology had to be developed at Fraunhofer IAF, which provided transistor devices with excellent high-frequency performance and high breakdown voltages. Although Fraunhofer IAF has a long tradition of providing state-of-the-art technologies, the existing technologies did either not provide enough breakdown voltage or speed with respect to modulator driver development. This work contributed an experimental investigation and evaluation of different epitaxial layer structures. Furthermore, the high-frequency performance of transistors depending on device geometry was measured and evaluated. The efforts resulted in high-performance transistor devices with optimized vertical and lateral dimensions in the new technology. The optimization process is described in Chapter 2.

The link between technology, i.e., the single physical devices, and the design of complex circuits is a transistor model. From several proposed InP-HBT models, the most promising model approach was adopted. Starting from a rudimentary extraction procedure, the extraction of the model was refined and optimized to make a compromise between complexity and accuracy. This resulted in transistor models, which were successfully applied in the design of circuits. Furthermore, models were used to provide an insight into several transistors revealing limitations of the devices. This knowledge supported the optimization of the devices (Chapter 2). The underlying model considerations, the extraction procedure and modeling results are presented in Chapter 3.

Before circuits can be successfully designed, figures of merit (FOMs) and requirements have to be defined. In addition, measurement procedures have to be determined, which allow to evaluate the amplifiers with respect to the FOMs. Chapter 4 summarizes the key requirements for amplifiers used at bit rates of 80 Gbit/s and beyond and gives an overview of the measurement techniques that were used to test the amplifiers.

With optimized transistor devices and suitable models at hand, the development of amplifiers according to the requirements is possible. For amplifiers two common design

approaches exist: the lumped amplifier and the distributed amplifier. For both approaches several concepts were realized, evaluated and optimized with respect to different design goals like bandwidth, output power or both. The most challenging part was the realization of an amplifier concept, which features enough bandwidth as well as enough output power for the realization of modulator drivers. The results of the design efforts are described in Chapter 5 for lumped amplifiers and in Chapter 6 for distributed amplifiers.

Relationship to other Work at Fraunhofer IAF

As stated above, the development efforts at Fraunhofer IAF are aiming towards the realization of not only amplifiers but all electrical components needed for an ETDM system operating at 80 Gbit/s (s. Fig. 1.1). Consequently, some results of this thesis contributed to the realization of other system components as well. The optimized transistor devices (Chapter 2) were the building blocks not only in analog amplifier, but also in digital circuit design. The transistor model extraction (Chapter 3) was optimized for the simulation of analog as well as digital circuits. Among these related circuits are a 80 Gbit/s 2:1 multiplexer (MUX) [74], a low phase noise voltage controlled oscillator (VCO) operating up to 89 GHz [73], as well as a 80 Gbit/s clock and data recovery circuit (CDR) with a 1:2 demultiplexer [72]. In summary, the combined efforts at IAF have provided a transistor technology and all circuit concepts needed for an 80 Gbit/s ETDM system in a period of less than four years.

1.2 State of the Art

1.2.1 Broadband Amplifiers

Bandwidth

High-speed optical data communication systems require broadband amplifier circuits with high gain. Distributed amplifiers (DA) are ideal candidates for enhancing bandwidth, because the input and output artificial transmission lines have high cut-off frequencies. Using high electron mobility transistors (HEMT), excellent circuit performance with bandwidth values exceeding 110 GHz has been demonstrated [75], [77], [53]. Similarly, several distributed amplifiers using hetero-bipolar transistors (HBTs) have been published with bandwidth values of up to 100 GHz [4], [8], [95]. Some of these results are summarized in Fig. 1.3. Even bandwidth values exceeding 150 GHz have been reported using InP HEMTs [2], but the applied design approach limited the lower bandwidth boundary to more than 1 GHz, which is not sufficient for the transmission of 80 Gbit/s digital signals.

Output Power

The lower bandwidth performance of HBT DAs is generally considered to be due to the lower input resistance of HBTs, which results in higher losses in the input artificial trans-

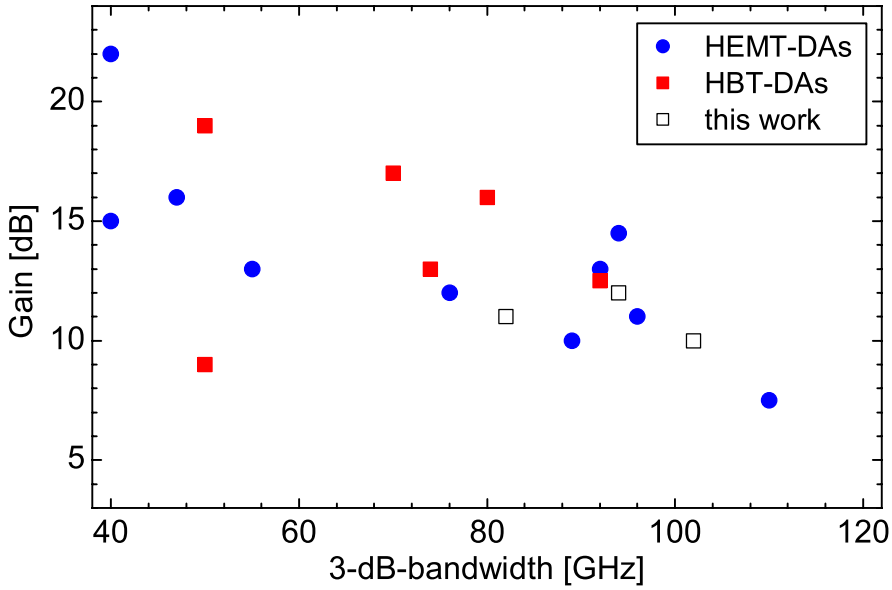


Figure 1.3: Summary of reported distributed amplifier results.

mission line of the amplifier circuit [61]. This limits the number of gain cells and degrades bandwidth. However, HBTs are still attractive candidates for the realization of modulator drivers for high data rates > 80 Gbit/s, because high output power is required to drive either electro-absorption or Mach-Zehnder modulators [8]. This is especially true for InP-based DHBTs (s. Sec. 1.2.2). Differential modulator drivers for 40 Gbit/s systems using such a technology have shown 25 dBm output power [9]. Only very recently, a 80 Gbit/s eye diagram was demonstrated with a reasonable voltage swing of 2.7 V_{pp} [99].

In this work, amplifiers featuring a 3-dB bandwidth exceeding 80 GHz and a gain of more than 10 dB were realized (s. Fig. 1.3). Additionally, investigations of power performance show, that the final amplifier concept of this work is suitable for modulator drivers. The presentation and explanation of the critical development steps to achieve these state-of-the-art results is part of this thesis.

1.2.2 Devices

Transistor's figures of merit (FOM), which are important in the development of circuits for high bit rates, are device speed, breakdown voltage and noise.

Device Speed

The high-frequency characteristic of a transistor is usually characterized by the FOMs cut-off frequency of the short-circuit current gain (f_T) and maximum frequency of oscil-

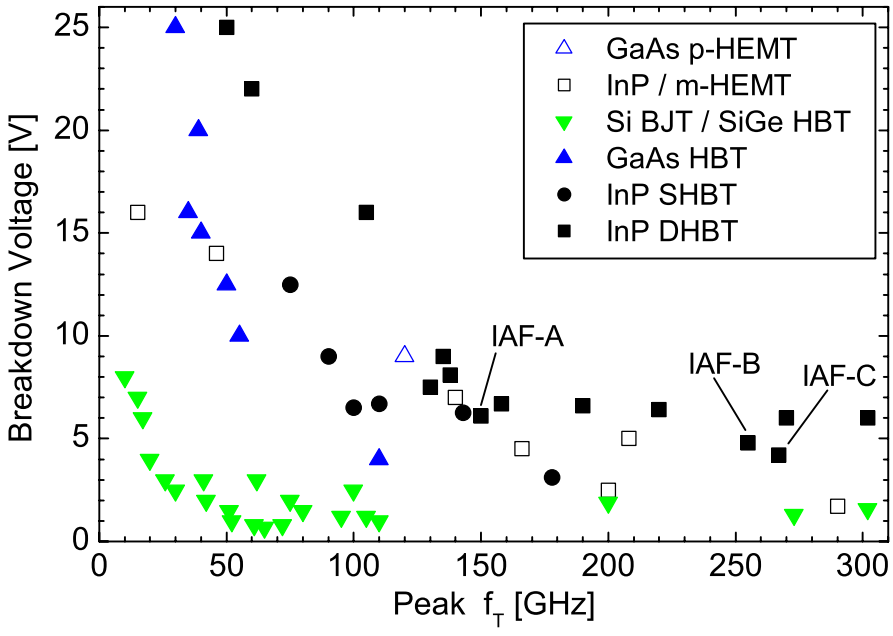


Figure 1.4: Breakdown voltage versus maximum current gain cutoff frequency (f_T) for representative Si, GaAs, and InP device technologies according to references [79], [30], [80], [98], and [93]. The data points IAF-A/B/C correspond to three development steps described in Sec. 2.3.

lation f_{max} (s. Sec. 2.1). They have a strong influence on the switching times in digital circuits as well as the bandwidth, which can be achieved in amplifier design. A commonly cited statement is that f_T is more important for digital circuits, while for analog applications f_{max} is more significant [31]. In this study, we intended to develop a device technology that would be used for both kind of circuits. The optimization process to achieve $f_T \approx f_{max}$ is described in Chap. 2.

The successful realization of 40 Gbit/s circuits required a technology with high-frequency parameters in excess of 100 GHz. Consequently, for the realization of 80 Gbit/s circuits high-frequency parameter values exceeding 200 GHz are assumed to be necessary. Fig. 1.4 gives an overview of different kinds of technologies that can fulfill this requirement, with regards to f_T.

Breakdown Voltage

The breakdown voltage of a device determines the upper limit of the voltage swing that a device can deliver to a load. In the case of broadband amplifiers used as a modulator driver, this is the maximum available voltage swing, that can drive a modulator (s. Sec. 5.3). Depending on the modulator concept voltage swings between 3 and 7 V are needed to obtain an acceptable extinction ratio at the optical output of the modulator. Therefore, the device technology has to provide a breakdown voltage of at least 3 V.

However, a fundamental limitation of transistors is the trade-off between breakdown voltage and the cut-off frequency for short-circuit current gain f_T. For this reason, Fig. 1.4 shows the breakdown voltage of various technologies as a function of high-frequency performance. It is clear, that InP-based technologies show outstanding characteristics in comparison to GaAs- and Si-based technologies.

InP Double Heterojunction Bipolar Transistor

For the realization of electrical components for 80 Gbit/s an InP Double Heterojuction Bipolar Transistor (DHBT) technology was developed at Fraunhofer IAF. The physical properties of InP related materials allow to develop devices, which offer a unique combination of high speed and high breakdown voltages. Based on reported data, an initial epitaxial layer structure (A) was grown and devices were processed. Subsequently, the devices were investigated and weaknesses of the layer structure were identified. Improved layer structures were designed, which resulted in an optimized layer structure (C) after two iterations (B,C) as illustrated in Fig 1.4. The optimization process is described in Sec. 2.3.

When compared with HEMTs, the InP-based HBT offers the advantage of high transconductance per area allowing high integration density, higher homogeneity of turn-on voltage across the waver and reduced low-frequency noise [8]. The two former points are beneficial in the design of digital circuits. The latter point allows the realization of very low phase noise VCOs , which are important building blocks of the CDR (Fig. 1.1).

Although the highest bandwidths in distributed amplifier design were achieved using HEMTs (s. Fig. 1.3), the bipolar technology is still suitable for amplifier development. Improved design techniques such as an RC emitter degeneration network (s. Sec. 6.1), can yield HBT distributed amplifiers which have comparable gain-bandwidth products as HEMT versions, as shown in reference [7].

Chapter 2

InP-based Double Heterojunction Bipolar Transistor

The heterojunction bipolar transistor (HBT) was proposed by William Shockley (1948), who received together with Walter Brattain and John Bardeen the nobel prize in physics for the invention of the bipolar transistor in 1956. However, it was Herbert Kroemer [63], who is credited with developing the detailed theory of HBTs and received the nobel prize in physics for his work on semiconductor heterostructures in 2000.

The name "transistor" was created from the words "*trans*fer resi*stor*", as the transistor operation was discovered while experimenting with transfer resistances between semiconductor layers. In contrast to field effect transistors, bipolar transitors consist of semiconductor layers of opposite doping-types. Heterojunction bipolar transistors incorporate one and Double Heterojunction Bipolar Transistors (DHBT) incorporate two semiconductor heterojunctions, which enhances their performance in comparison to homojunction bipolar transistors [64].

III-V semiconductor devices offer improved characteristics in comparison to silicon technologies. In the case of InP-based DHBTs, the electron mobility in the base layer is one order of magnitude larger. Moreover, the InP wide-bandgap collector features a higher breakdown voltage. Therefore, operation at higher power levels as well as higher frequencies in comparison to silicon based devices is possible. This makes InP-DHBTs suitable for developing circuits such as modulator drivers for high frequencies at medium or high power levels.

This chapter explains the operating principle of DHBTs, gives an overview of the process and explains how the layer structure and the transistor layout were optimized. As a result, transistor devices featuring f_{T} and f_{max} values in excess of 280 GHz were realized.

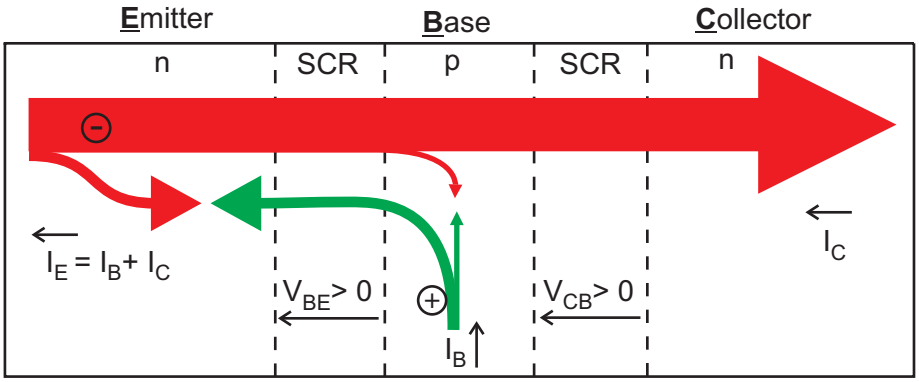


Figure 2.1: Flow of electrons and holes in a BJT operating in forward active mode. Electrons enter the base layer from the emitter. The vast majority reaches the collector. A very small fraction recombines in the base layer. The hole diffusion current recombines with electrons in the emitter. In first order approximation the recombination of electrons in the base can be neglected as well as generation and recombination of carriers in the space charge regions (SCR).

2.1 Operating Principle

Bipolar Junction Transistor

A bipolar junction transitor (BJT) consists of three main layers, the emitter, the base and the collector. The emitter and the collector are doped of the same type, whereas the base layer has the opposite type of doping. These three layers form two p-n junctions with the base between the other two layers (s. Fig. 2.1). For high-speed applications, n-p-n structures are preferably used due to the higher mobility of electrons in comparison to holes in the base (s. Section 2.3.1).

In normal operation condition, the base-emitter junction is forward biased ($V_{be} > 0$) and the base-collector junction is reverse biased ($V_{bc} < 0$). This is called forward-active mode of operation. In first order approximation, a current flows at the base emitter junction, consisting of two carrier diffusion currents I_{nD} and I_{pD}. I_{nD} is the carrier diffusion current of electrons injected into the base and I_{pD} is the carrier diffusion current of holes injected into the emitter. Fig. 2.1 illustrates the current flow in a BJT.

I_{nD} is affected by the base-collector junction. The electrons in the base diffuse from the base-emitter junction to the base-collector junction. Because the base thickness X_B is chosen to be much smaller than the diffusion length for electrons L_n, most electrons reach the base collector space charge region before they can recombine with holes in the base. The electric field in the space charge region pulls those electrons towards the collector. Thus, the majority of electrons injected from the emitter into the base reach the collector. As a result, the collector current I_C is approximately equal to the diffusion current I_{nD} of a short diode,

$$I_C \approx I_{nD} = \frac{qA_{BE}D_n}{X_B}\frac{n_{iB}^2}{N_B}\left(\mathrm{e}^{\frac{qV_{be}}{kT}} - 1\right), \tag{2.1}$$

where q is the electric charge, A_{BE} is the base-emitter junction area, D_n is the minority electron diffusion coefficient, n_{iB} is the intrinsic carrier concentration of the base layer, and N_B is the base doping level [68].

The hole diffusion current I_{pD} consists of holes that originate in the base. As only very few electrons recombine in the base, the base current is approximated to be equal to the diffusion current of holes in the emitter,

$$I_B \approx I_{pD} = \frac{qA_{BE}D_p}{L_p}\frac{n_{iE}^2}{N_E}\left(\mathrm{e}^{\frac{qV_{be}}{kT}} - 1\right), \tag{2.2}$$

where D_p is the minority hole diffusion coefficient, L_p is the diffusion length for holes in the emitter, n_{iE} is the intrinsic carrier concentration of the emitter, and N_E is the emitter doping level [68].

The direct-current (dc) gain β of a bipolar transistor is defined as:

$$\beta_{DC} = \frac{I_C}{I_B} \approx \frac{D_n L_p N_E}{D_p X_B N_B} \tag{2.3}$$

From equation 2.3 we infer the following: in BJTs, high gain requires low base and high emitter doping levels. However, low doping levels in the base can lead to several undesired phenomena in transitor operation, like Early effect [29] and Webster effect (high injection) [104]. HBTs can reduce these problems.

Heterojunction Bipolar Transistor

Heterostructure p-n junctions offer greater design freedom than homostructure p-n junctions. With the evolution of more and more advanced growth technologies such as MBE (Molecular Beam Epitaxy) and MOCVD (Metal Organic Chemical Vapor Deposition), a change in semiconductor material composition, i.e., bandgap, is not significantly harder to achieve than a change in doping level. In HBTs energy gap variations in addition to electric fields act as driving forces on electrons and holes to control their distribution and flow. By selecting the appropriate bandgap and doping levels, the forces acting on electrons and holes can be separately controlled to a certain degree.

Figure 2.2 shows the band diagram of a InP-DHBT in forward active mode. The base-emitter junction is a heterojunction. As a result the current gain equation has to be modified. The intrinsic carrier concentration in the emitter is smaller than in the base. If we assume that the density of states in both semiconductors is the same, the relation between n_{iB} and n_{iE} depends only on the bandgap difference ΔE_G [68] and is:

$$n_{iB} = n_{iE}\mathrm{e}^{\frac{\Delta E_G}{kT}}$$

Therefore, the current gain of an HBT is approximated to be:

$$\beta_{DC} = \frac{I_C}{I_B} \approx \frac{D_n L_p N_E}{D_p X_B N_B}\mathrm{e}^{\frac{\Delta E_G}{kT}} \tag{2.4}$$

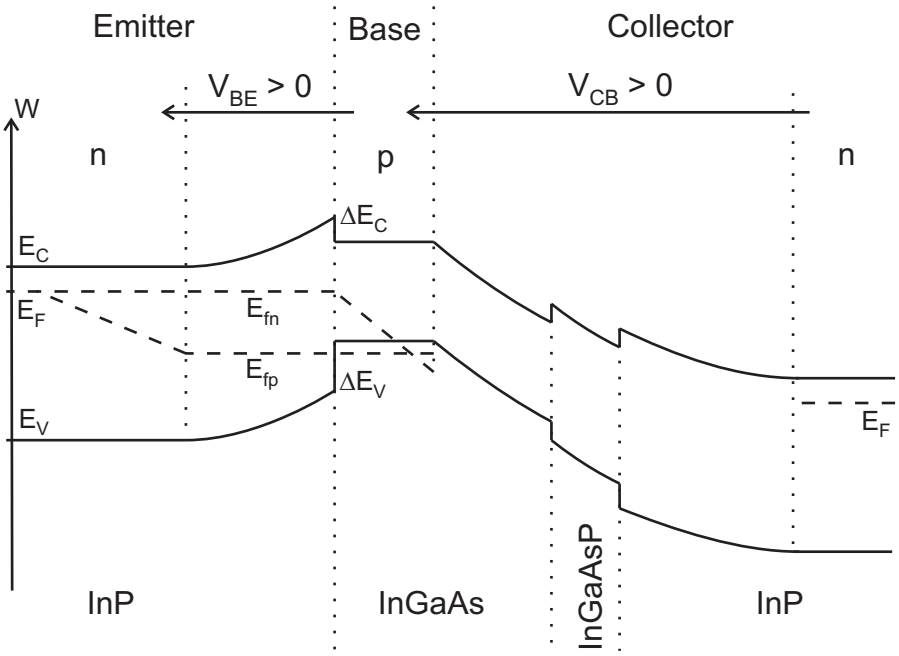


Figure 2.2: Band diagram of a DHBT in forward active mode. The bandgaps E_G of InP and InGaAs are 1.35 and 0.75 eV, respectively. The conduction and valence band discontinuities at the emitter base junction have values of $\Delta E_C = 0.23$ eV and $\Delta E_V = 0.37$ eV. The step graded collector supports uninhibited electron transport and is explained in Sec. 2.3.1.

Equation 2.4 shows the key advantage of HBTs over BJTs. Because usually $\Delta E_G \gg kT$, HBTs feature high values for β almost regardless of the base doping level. This offers the opportunity to increase the base doping level, considerably. This has several positive consequences for the transistor performance: Webster and Early effect, which are likely to occur at low base doping levels, play a negligible role in HBTs. Additionally, the base resistance is reduced allowing a reduction of base thickness X_B which leads to even higher gain and improves device speed as will be explained in the following paragraphs.

High-Frequency Performance

For an estimation of the high-frequency (hf) performance, the small-signal equivalent circuit is used (Fig. 2.4). It is motivated by the schematic of an HBT shown in Fig. 2.3: all semiconductor layers exhibit a resistance, although the resistance of the subcollector layer adjacent to the substrate is insignificant. The regions above and underneath the base layer are depleted and constitute parallel plate capacitors. The base-emitter and base-collector diodes are modeled using RC-circuits. The current source with gain α models the transfer of electrons, which do not recombine in the base, from the base to the collector. At high frequencies the distributed nature of the base-collector junction is revealed. Hence the total collector capacitance C_{bc} is split in parts, C_{jc} and C_{ex}. Outside the internal HBT, resistors represent layer, contact, and wiring resistances, and capacitors model any capacitive

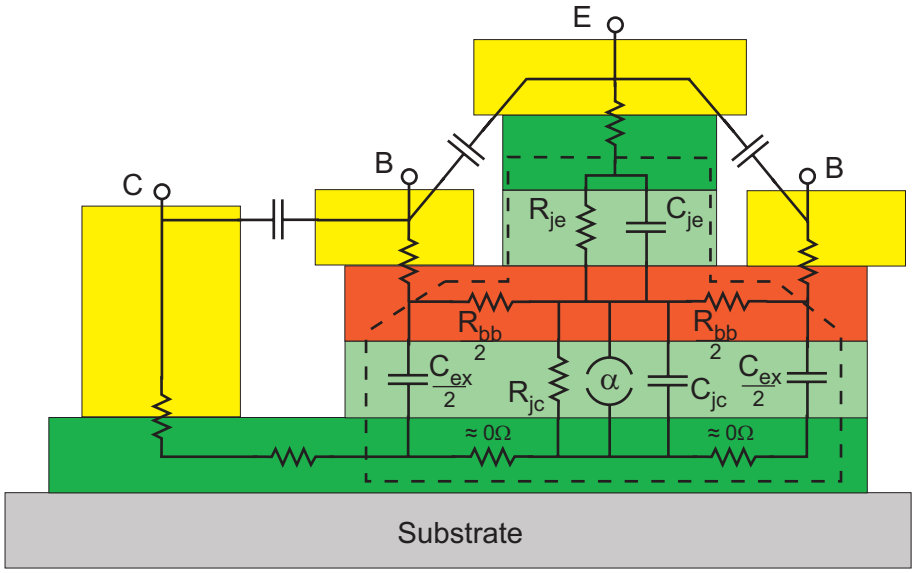


Figure 2.3: Schematic of HBT structure and electrical components associated with its small-signal behavior. The internal HBT is surrounded by a dashed line. The base layer is in the middle and the emitter and collector layers are above and underneath, respectively. Parasitic inductances have been omitted for clarity. Figure 2.4 shows the resulting small-signal equivalent circuit.

coupling between contact pads.

From the equivalent circuit (Fig. 2.4), two important figures of merit, f_{T} and $f_{\max}$ can be derived. The former parameter is called cut-off frequency. It is defined as the frequency at which the magnitude of the two-port parameter h_{21} decreases to unity. h_{21} is the small-signal current gain with base and collector under short-circuit condition. A detailed derivation is given in [67] and its result is:

$$f_{\mathrm{T}} = \frac{1}{2\pi\tau_{ce}} \quad ; \quad \tau_{ce} = \underbrace{R_{je}(C_{je}+C_{bc})}_{\tau_e} + \underbrace{\tau_b + \tau_c}_{\tau_T} + C_{bc}(R_E + R_C) \tag{2.5}$$

The second parameter $f_{\max}$ is called maximum frequency of oscillation. At this frequency, the unilateral power gain of the transistor rolls off to unity. This frequency marks the boundary between active and passive network. The approximation of $f_{\max}$ is also given in [67]:

$$f_{\max} = \sqrt{\frac{f_{\mathrm{T}}}{8\pi(R_{bb}+R_B)C_{jc}}} \tag{2.6}$$

The equations for f_{T} and $f_{\max}$ reveal another advantage of HBTs. Due to the excellent electron mobility in the base, the base transit time τ_b is very low and supports high values for f_{T}. The possibility to dope the base aggressively gives low values for the base layer resistance R_{bb} which support high values for $f_{\max}$ ($\geq f_{\mathrm{T}}$).

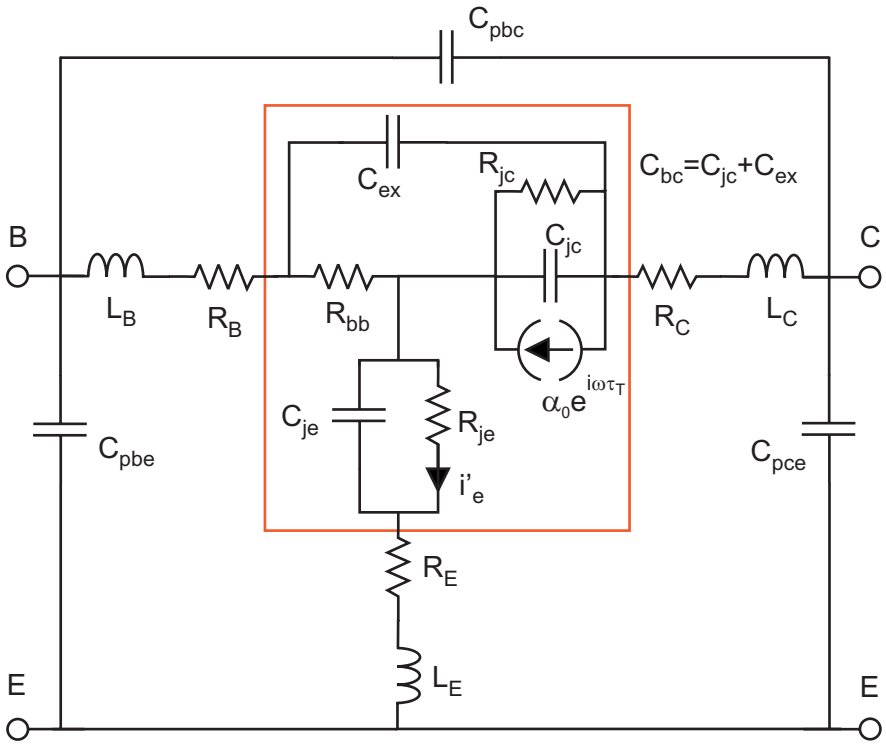


Figure 2.4: Small-signal equivalent circuit of HBT. The internal device is within the boundary. It is motivated by the schematic of the HBT (Fig. 2.3). It is surrounded by parasitic environment consisting of capacitors, inductors and resistors. These parasitics account for contact and wiring resistances, capacitive coupling, and inductive behavior of connecting wires. More details regarding this small-signal model are given in Sec. 3.3.

The dc formulas and the description of the transitor operation in this section have been simplified. A discussion of more physical effects influencing transistor performance will follow in Sections 2.3 and 2.4. How these effects influence the mathematical description of the device will be dealt with in Chapter 3.

2.2 Technology

2.2.1 Active Devices

At Fraunhofer IAF the fabrication process of InP/InGaAs DHBTs and MMICs on 3-inch InP substrates is based on a standard triple mesa process [48]. This technology uses self-aligned base-emitter contacts and selective wet chemical etching. Scanning Electron Microscope (SEM) pictures of processed HBTs are shown in Figures 2.5 and 2.6. Benzocyclobutene (BCB) is used for device passivation and planarization. The collector and the base are connected through via holes in the BCB. A schematic displaying all resulting elements of the process is depicted in Figure 2.7.

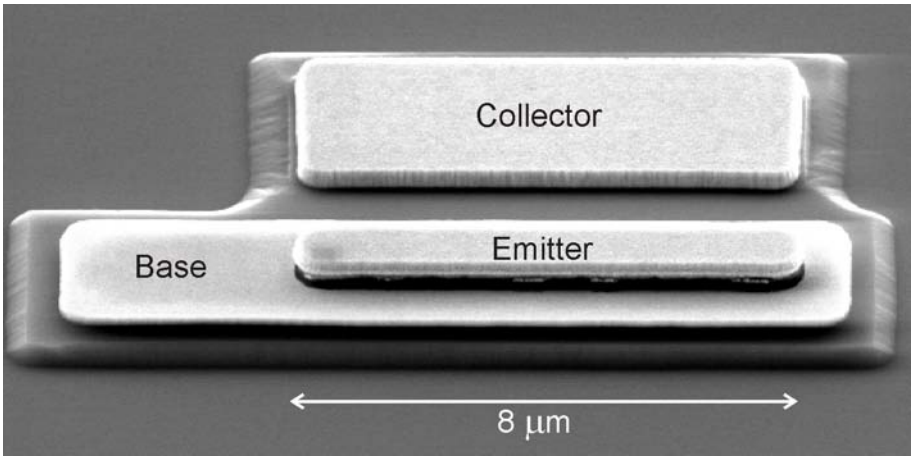


Figure 2.5: SEM top view of an HBT. A cross-section of an identical device is shown in Fig. 2.6. A schematic of the transistor layout is shown in Fig. 2.25.

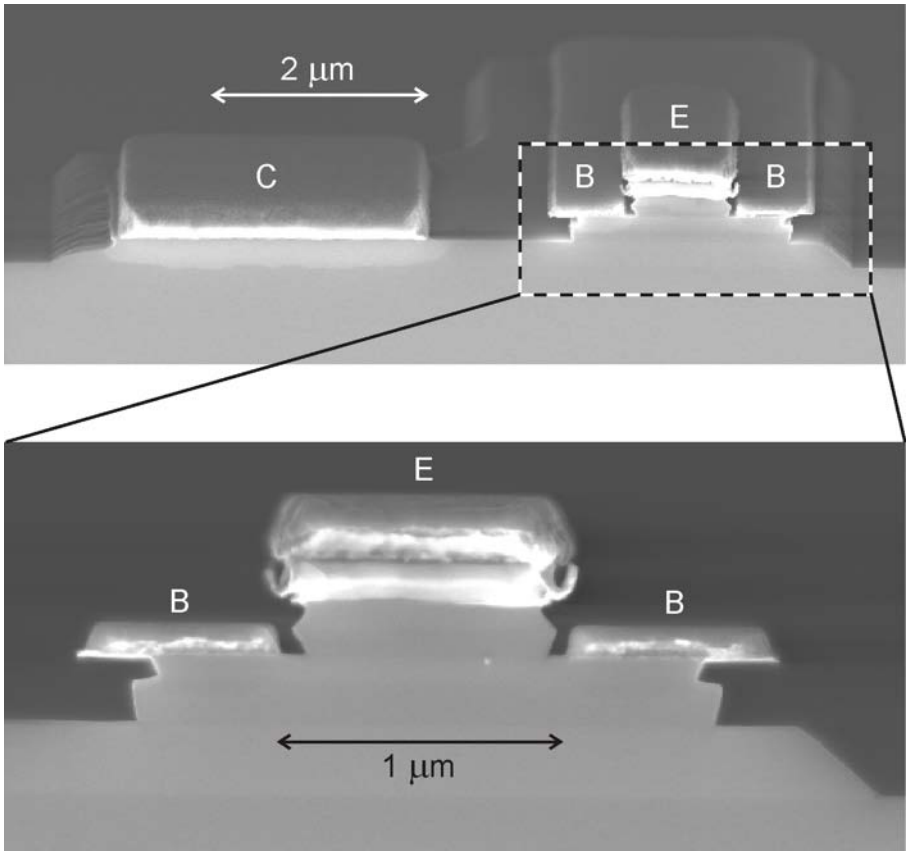


Figure 2.6: SEM pictures of HBT cross-section.

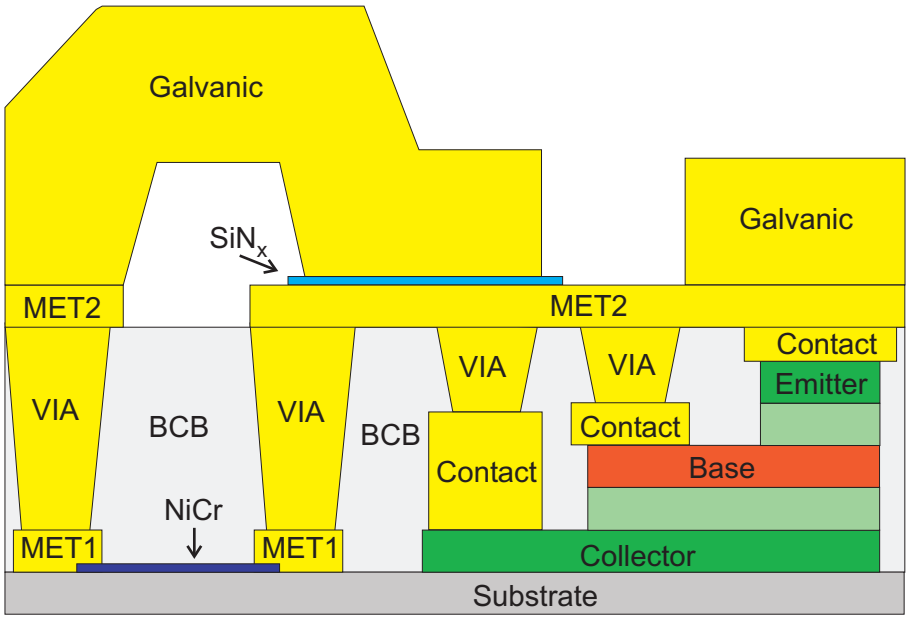


Figure 2.7: Schematic of the process showing elements available for circuit design including NiCr resistors and silicon nitride MIM capacitors.

The process includes thin film MIM capacitors (23 fF/μm^2), NiCr resistors (50 $\Omega/\square$), and three levels of Au-based interconnect metals. The NiCr resistors are deposited on the substrate and are connected using MET1 before BCB coating. The MIM capacitors are realized using a thin layer of silicon nitride between the layers MET2 and Galvanic metal. The Galvanic metal can also be used to realize airbridges. Therefore, three levels of metalization can be used for interconnects.

2.2.2 Passive Structures

Coplanar waveguides (CPW) are used as transmission lines. As shown in Fig. 2.8, a coplanar waveguide consists of three conductors: one signal strip in the center and two ground conductors. The CPW technology is used, because it does not need a backside metallization, which is required for microstrip lines, and reduces the processing efforts.

The ground-to-ground spacing (d) of the coplanar lines in this work is 50 μm. By varying the width of the center conductor (w) the impedance of the transmission line can be adjusted. Formulas describing the impedance level are given in [42]. In the amplifier designs in Chapter 5 and 6 the center conductors have a width of 17 and 7 μm, corresponding to a line impedance of 50 and 69 Ω.

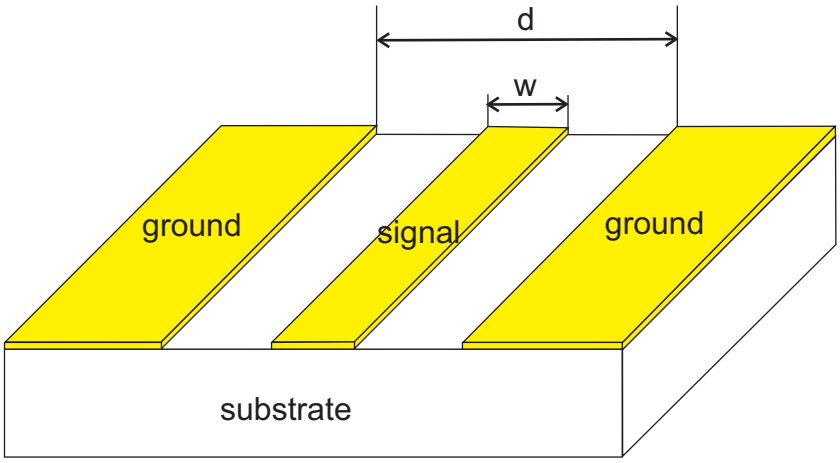


Figure 2.8: Schematic of a coplanar waveguide cross-section.

2.3 Layer Structure Optimization

An optimized epitaxial layer structure is a prerequisite for high-performance transistors, as it determines the intrinsic device performance. The layer structure and the transistor layout (s. Sec. 2.4) determine the values of the small-signal equivalent circuit elements in Fig. 2.4. As explained above, two important figures of merit f_{T} and f_{max} can be calculated from these element values (s. Eqs. 2.5 and 2.6). Furthermore, the direct-current (dc) gain and power performance of the transistors depend on the layer structure.

This section describes how the experimental investigation of three different layer structures standing for three development steps contributed to the optimization process. First, the results for an initial layer structure are presented. Features and deficiencies of the structure are discussed. Following this, the results of two subsequent structures are presented showing how intended improvements could be realized by changing the layer structures resulting in cut-off frequencies (f_{T}, f_{max}) exceeding 250 GHz. Finally, the results of this section are summarized.

2.3.1 Initial Layer Structure (A)

Emitter Layer

Considering the high-frequency performance of HBTs (s. Eq. 2.5), the contribution of the emitter to the total delay time is:

$$\tau_e = R_{je}(C_{je} + C_{bc})$$

In forward active mode the base-collector capacitance C_{bc} is much smaller than C_{je} and can be neglected (s. Figs. 2.4 and 3.17). Hence, especially at low collector currents, where the differential resistance R_{je} of the emitter base diode has a considerable value, the base

Layer Structure A

Layer	Thickness (nm)	Semiconductor	Doping level (cm^{-3})
Emitter	70	InP	
Base	40	$In_{0.53}Ga_{0.47}As$: C	5.0×10^{19}
Collector	50	$In_{0.53}Ga_{0.47}As$	
Collector	20	$In_{0.85}Ga_{0.15}As_{0.32}P_{0.68}$	
Collector	20	InP: Si	2.0×10^{17}
Collector	150	InP	

Figure 2.9: Schematic of layer structure A.

emitter capacitance C_{je} has a major impact on f_T (Eq. 2.5). Thus, for high-frequency performance it seems advisable to keep the emitter junction capacitance low by implementing a long emitter depletion layer. It has been shown experimentally that a lightly doped emitter with a relatively long emitter space charge region can result in higher values of f_T in comparison to emitters with high doping and a short space charge region [19]. The extreme case of this concept is an undoped emitter [45], where the thickness of the emitter layer alone determines the length of the space charge region and controlls the base emitter junction capacitance. In our first structure (A) shown in Fig. 2.9 the latter concept is implemented with an undoped emitter of 70 nm length in between the highly doped base and cap layers.

Base Layer

The base layer affects the transit time of minority carriers in the base τ_b and the base layer resistance R_{bb}. τ_b is a critical parameter for high-frequency performance (s. Eq. 2.5), because the only transport mechanism in the base is diffusion, unless a grading has been implemented. The base layer resistance R_{bb} has an effect on the relation between f_T and f_{max} (s. Eq. 2.6).

A reduction of the transit time of minority carriers in the base increases f_T. The mobility of electrons has been reported to be larger than 3000 cm^2/Vs in heavily p-type doped ($N_A = 3.1 \times 10^{19}$ cm^{-3}) InGaAs base layers of InP-HBTs [13]. Therefore, the base layer in InP-DHBTs is generally chosen to be of p-type [68]. Furthermore, to achieve a small value for τ_b the base should be as thin as possible. A short transit time also has a positive effect on the current gain as the probability of recombination for each minority carrier crossing the base is small. However, the base layer resistance sets a lower limit for the base thickness.

To keep the base layer sheet resistance R_{bb} small, the base doping should be very high. The structures investigated were doped with carbon, which is the preferred p-type dopant for the base in this material system [10], [11], [28]. The doping level in this work

is 5 $\times$ 10^{19} cm^{-3}. However, even at such doping level, the InGaAs resistivity is noticeable. Therefore, the thickness of the base layer was chosen to be 40 nm to effect a compromise between short transit time τ_b and low base layer sheet resistance R_{bb} (Fig. 2.9).

Collector Layer

The collector layer has a major influence on both high-frequency parameters f_{T} and f_{max}. From Eq. 2.5 the delay time associated with the collector is:

$$\tau = \tau_c + C_{bc}(R_E + R_C + R_{je})$$

The first parameter τ_c describes the time electrons need to travel through the collector. The second term is the charging time constant of the base collector capacitance C_{bc}. The resistances R_E and R_C reside outside the internal HBT and are determined by the metal-semiconductor contacts, the subcollector layer resistance and the resistance of the external wiring. R_{je} depends on the emitter design and is small at normal operation in forward active mode. The maximum frequency of oscillation depends on f_{T} and C_{jc} (s. Eq. 2.6). Therefore, regarding high-frequency performance the collector thickness is a compromise between short delay time τ_c (thin collector) and low capacitance values for C_{bc} and C_{jc} (thick collector).

Due to their large bandgap, InP layers in the collector of DHBTs can improve the breakdown voltage, but they also tend to induce current blocking at the base-collector junction if an InGaAs base is used. If the collector design is poor, electrons propagating through the space charge region can get caught at the barrier and rejected into the base. This reduces the dc current gain and causes additional storage of carriers in the collector deteriorating the dc and high-frequency performance. A number of grading schemes have been developed to overcome blocking at the heterojunction [28].

Among various approaches, a thin n-doped InGaAs layer can be used to lower the conduction band spike [41]. Another concept employs a superlattice structure consisting of alternating InP/InGaAs layers with varying thickness helping to overcome the potential barrier in the conduction band [40]. Alternatively, a quaternary layer of InGaAsP with an intermediate band gap between that of InGaAs and InP can reduce the height of the potential spike at the B-C junction [33]. In addition, a thin highly n-doped InP layer in the collector can alleviate the current blocking effect [65]. This doping raises the Fermi level at the peak of the barrier or, in other words, pulls down the effective barrier.

In structure A, the collector incorporates an InGaAsP step grading and a thin n-doped InP layer adjacent to it (s. Fig. 2.9). The total thickness of the collector layer is 240 nm. A schematic band diagram of a DHBT with step graded collector is shown in Fig. 2.2.

Experimental Results

The characteristics of the layer structure were investigated by measuring an HBT with an emitter area of 10 μm^2. The resulting output characteristic and Gummel plot are shown in Figs. 2.10 and 2.11, respectively.

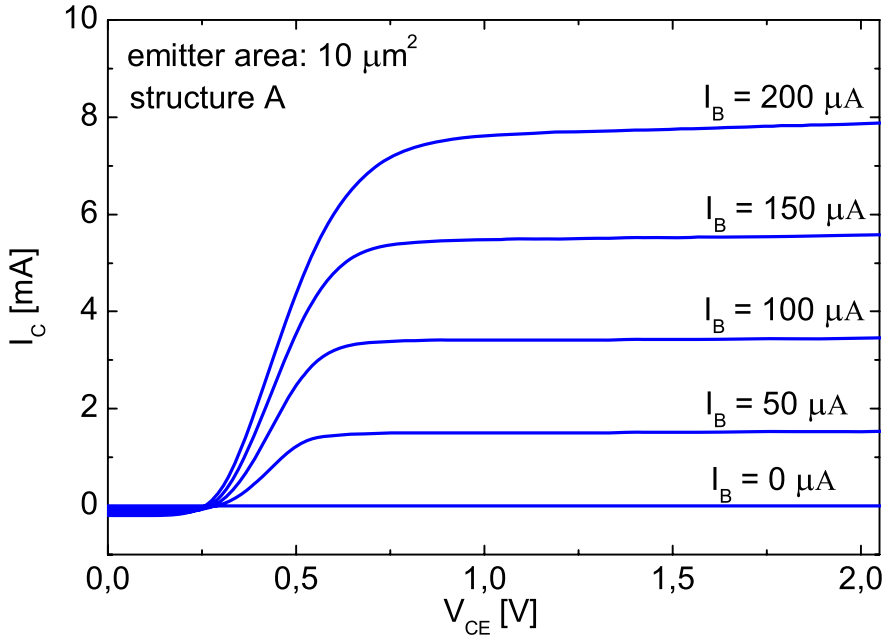


Figure 2.10: Output characteristic of a transistor from layer structure A.

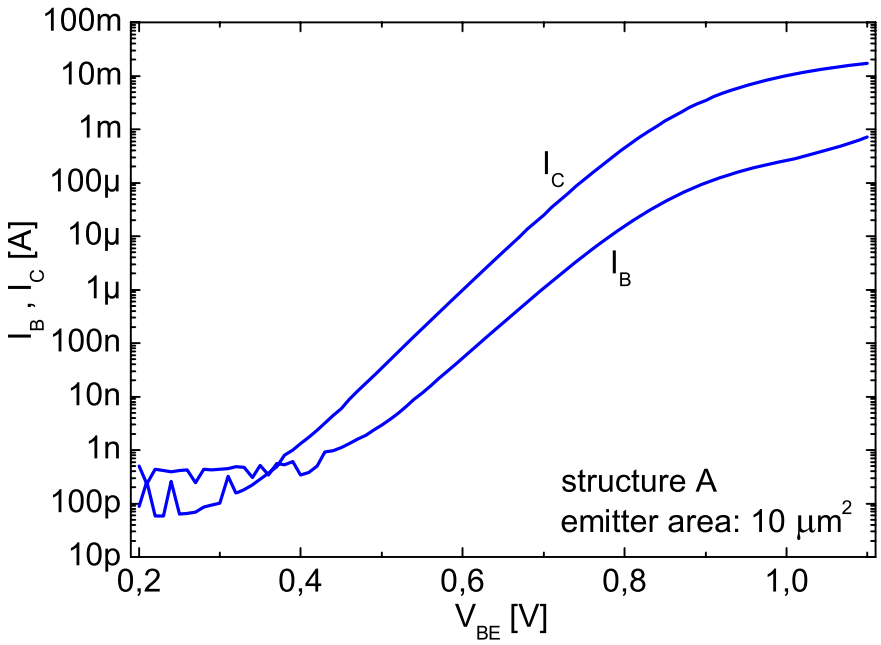


Figure 2.11: Measured Gummel plot of a transistor with an emitter area of 10 μm^2 from layer structure A.

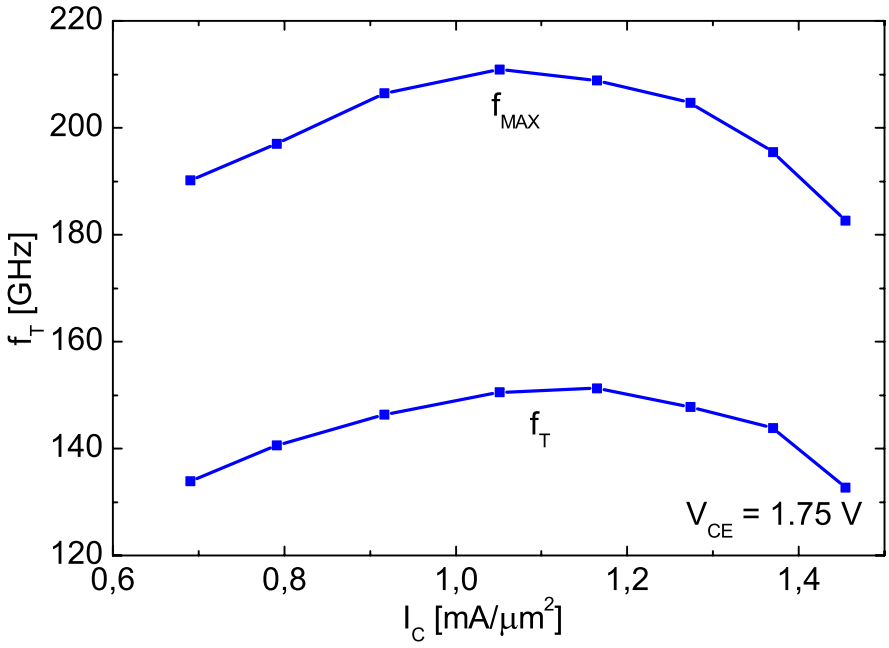


Figure 2.12: High-frequency parameters as a function of current density of an HBT with an emitter area of 10 μm^2 from layer structure A (s. Fig. 2.9).

The turn on voltage in the output characteristic is approximately 0.3 V and the saturation voltage (knee voltage) 0.7 V. The current gain extracted from the Gummel plot is 39. The Gummel plot also shows, that the leakage currents are below 1 nA and the ideality factors for base and collector current extracted at 1 μA are 1.3 and 1.2, respectively.

For high-frequency characterization, scattering parameters were measured using an Agilent 8510XF measurement system from 0.25 to 120 GHz (s. Sec. 4.2). From the generated data the short-circuit current gain H_{21} and the unilateral gain U were calculated and the high-frequency parameters f_T and f_{max} extracted. The measurements were conducted with a collector-to-emitter voltage of 1.75 V, which typically resulted in optimized high-frequency performance (compare Fig. 3.11). The results are shown as a function of current density in Fig. 2.12. The highest speed is reached at a collector current density of 1.1 mA/μm^2 reaching f_T and f_{max} values of 150 and 210 GHz, respectively.

Discussion

The results demonstrate, that this first structure displays desirable features needed for the design of high-speed circuits. The maximum frequency of oscillation f_{max} exceeds 200 GHz and the current gain has an acceptable value of 39. From the process point of view, the low leakage currents in the Gummel plot present a good result.

Nevertheless there are several points, which need to be improved. First, f_T should also exceed 200 GHz. This is desirable, because the process is intended for digital as well as analog circuits targeting high bit rates. The switching time in digital circuits strongly

Layer Structure B

Layer	Thickness (nm)	Semiconductor	Doping level (cm^{-3})
Emitter	40	InP: Si	6.0×10^{17}
Base	30	$In_{0.53...0.45}Ga_{0.47...0.55}As$: C	5.0×10^{19}
Collector	30	$In_{0.53}Ga_{0.47}As$	
Collector	20	$In_{0.75}Ga_{0.25}As_{0.54}P_{0.46}$	
Collector	100	InP: Si	2.0×10^{16}

Figure 2.13: Schematic of layer structure B.

depends on f_T and the bandwidth of distributed amplifiers depends on f_{max} [1]. As a consequence, the goal of the optimization process is to achieve $f_{max} \approx f_T$. Second, the turn-on voltage and the saturation voltage ought to be decreased to enable operation at higher current densities without increasing the power consumption. Higher current densities would reduce the differential resistance R_{je} of the emitter-base junction (s. Eq. 3.6) and support higher values for the high-frequency parameters (s. Eqs. 2.5 and 2.6). Third, a larger current gain would be beneficial for the design of high gain amplifiers. These improvements can be achieved by changing the layer structure. An improved layer structure is presented in the following section.

2.3.2 Improved Layer Structure (B)

Emitter Layer

As discussed in the last section, a reduction of the turn-on voltage is desirable. This can be achieved by changing the thickness and doping level of the emitter layer. In abrupt InP/InGaAs emitter base junctions, the electron transport through the interface, where a discontinuity in energy levels occurs (s. Fig. 2.2), is controlled by thermionic emission over and tunneling transmission through the spike [69]. However, in an undoped emitter, the spike barrier is wide enough, owing to the absence of n-type dopants, to prevent effective tunneling [45]. Without tunneling a higher base-emitter voltage is needed to generate the same emitter current resulting in a higher turn-on voltage [68]. Therefore, a higher doping level and thinner emitter layer than in structure A are expected to reduce the turn-on voltage. For layer structure B the emitter layer was chosen to have a thickness of 40 nm and a doping level of 6×10^{17} cm^{-3} (s. Fig. 2.13).

Base Layer

A key to improving the current gain and the cut-off frequency f_T is a change in base layer design. In order to improve the current gain, an In grading is introduced in the base, which increases the band gap in the base gradually from the collector side to the emitter

side. The indium mole fraction x is decreased from the lattice-matched value (0.53) at the base-collector interface down to $x = 0.53 - \delta x$ at the base-emitter interface. The induced internal built-in field has been shown to be very effective for reducing the transit time, i.e., increasing f_T, and reducing the recombination probability of the electrons in the base, i.e., increasing current gain [66], [10]. The indium grading δx is maximized, while making the base as thin as possible and accommodating the strain to prevent dislocations due to the lattice mismatch at the base emitter junction. In layer structure B (Fig. 2.13), a grading of $\delta x = 0.09$ and a layer thickness of 30 nm are implemented.

The reduction in base thickness in comparison to structure A changes the ratio between transit time and layer resistance in favor of f_T. This is acceptable since f_{max} is much higher than f_T in that structure, which suggests that R_{bb} is very small (s. Eq. 2.6).

Collector Layer

In structure A the high knee voltage and the low current densities are likely due to limitations in the collector layer design. The high saturation voltage seems to be caused by a high resistance at low collector-to-emitter voltages. At low bias voltages the collector space charge region might not be fully depleted and cause high layer resistances in non-depleted regions. This can be due to the large thickness of the collector, which results in relatively low field intensities. Therefore, the collector for structure B was designed to be shorter.

A shorter collector will also influence the high-frequency parameters. The value for f_T will increase as a result of a shorter collector delay time τ_c. On the other hand, the collector capacitance C_{jc} will increase, which reduces f_{max}. However, this reduction might get overcompensated by the increase in f_T especially due to the changed base layer design.

A shorter collector should also enable higher current densities. The limitations in structure A are possibly due to the Kirk effect [55], [106]. At high current densities, the electrons entering the collector from the base bear enough charge to neutralize the electric field inside the collector. This prevents efficient field transport mechanisms and limits device speed.

A shorter collector is also likely to result in a lower breakdown voltage. However, the breakdown voltage of layer structure A is 6.1 V and a reduction is acceptable. A comparison of the breakdown voltages of all layer structures investigated, is presented in Sec. 2.3.4.

Considering the arguments above, the all over collector length was reduced from 240 nm in structure A to 150 nm in structure B. As no current blocking could be observed in structure A, the moderately doped InP layer adjacent to the InGaAsP step grading was replaced by a lightly doped InP collector layer (compare Figs. 2.9 and 2.13).

Experimental Results

For characterization of structure B, the same device type, as the one investigated on the previous structure, was measured. A comparison of the dc characteristics between devices on layer structure B and A are given in Figs. 2.14 and 2.15. The turn-on voltage decreased to 0.2 V. The saturation voltage increased to 0.8 V at high currents. The dc current gain

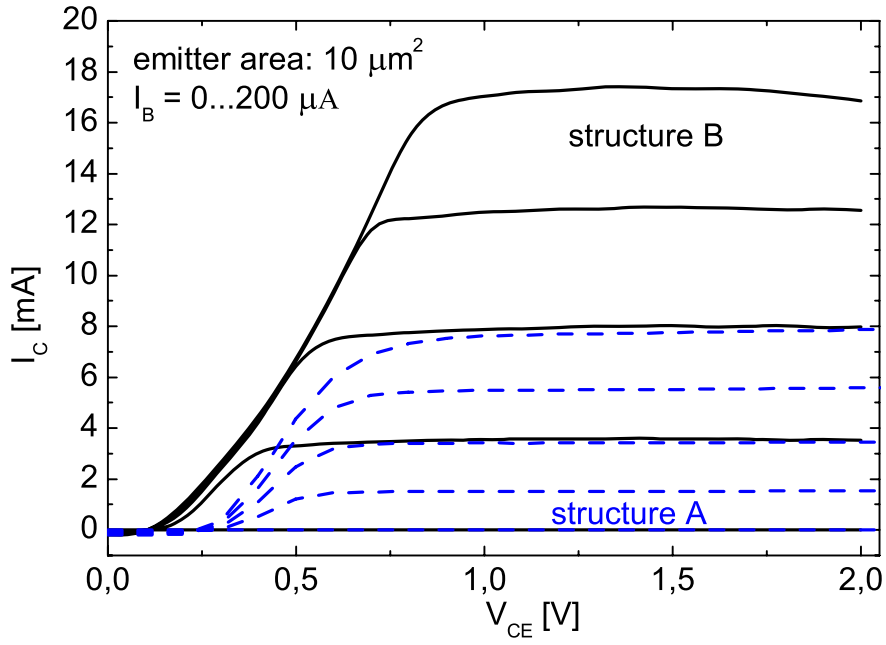


Figure 2.14: Comparison of output characteristics of transistors from layer structure B (solid lines) and layer structure A (dashed lines).

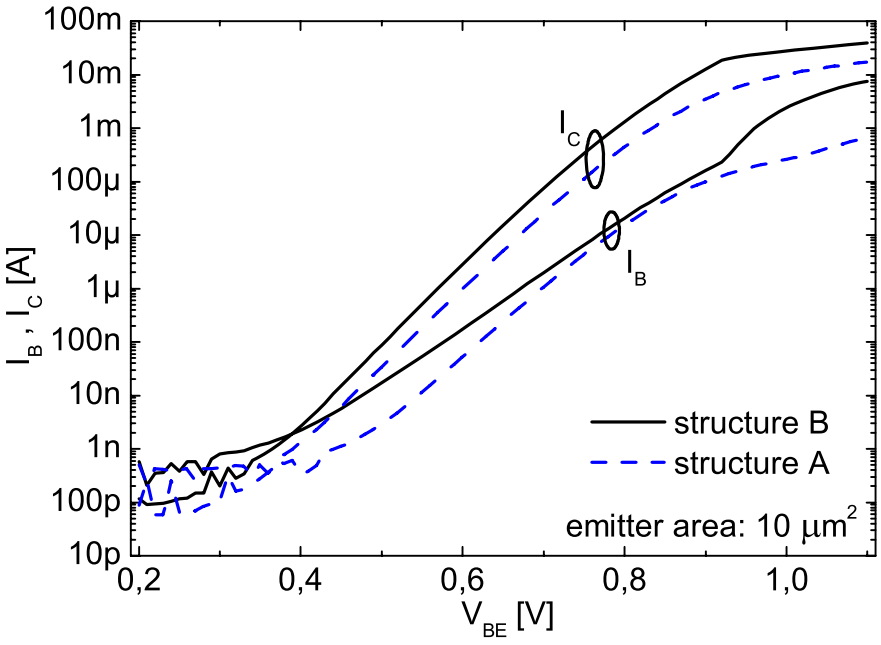


Figure 2.15: Comparison of Gummel plots of transistors with an emitter area of 10 μm^2 from layer structure B (solid lines) and layer structure A (dashed lines).

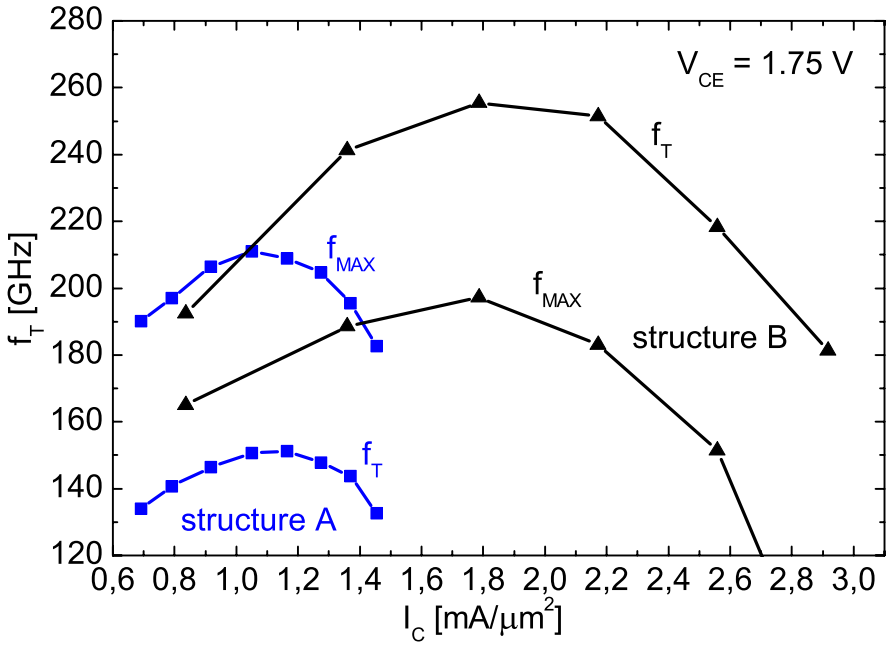


Figure 2.16: Comparison between layer structures B and A using high-frequency parameters of an HBT with an emitter area of 10 μm^2.

is more than twice as high as before, reaching a value of 80. The Gummel plot shows, that the leakage currents are still below 1 nA, and the ideality factors for base and collector current extracted at a current of 1μA are 1.5 and 1.1, respectively. In contrast to BJTs, DHBTs regularly feature different ideality factors for base and collector current, which depend in a complicated fashion on the design of the heterojunctions [68].

The measured high-frequency characteristics are shown in Fig. 2.16. As before, the measurements were conducted with a collector-to-emitter voltage of 1.75 V. The results show a cut-off frequency f_T of more than 250 GHz and a maximum frequency of oscillation f_max of almost 200 GHz at a current density of 1.8 mA/μm^2.

Discussion

The measurement results demonstrate several improvements in comparison with the previous layer structure. The turn-on voltage could be reduced and the current gain is doubled. The former result enables lower operating voltages and the latter enhances the capability to realize high gain amplifiers. These improvements are directly related to the change in emitter and base layer design. The graded base is responsible for the dramatic increase in β and f_T. At the same time, higher current densities are possible, suggesting, that the Kirk effect has been reduced.

However, there are still two parameters, which have to be improved. First, the saturation voltage at high currents is 0.8 V, which counteracts the low turn-on voltage. Second, the maximum frequency of oscillation, which did not rise at all, should be increased to

Layer Structure C

Layer	Thickness (nm)	Semiconductor	Doping level (cm^{-3})
Emitter	40	InP: Si	6.0×10^{17}
Base	30	$In_{0.53...0.45}Ga_{0.47...0.55}As$: C	5.0×10^{19}
Collector	50	$In_{0.53}Ga_{0.47}As$	
Collector	20	$In_{0.75}Ga_{0.25}As_{0.54}P_{0.46}$	
Collector	20	InP: Si	2.0×10^{17}
Collector	70	InP: Si	2.0×10^{16}

Figure 2.17: Schematic of layer structure C.

roughly match the high value of f_T. The limiting factor for both parameters is the effect of current blocking at the base-collector heterojunction. This effect does not only show in the output characteristic. The Gummel plot for structure B (Fig. 2.15) displays a kink for collector and base currents at 0.9 V, which is characteristic for this effect. The high-frequency performance is limited, because electrons accumulate at the heterojunction and increase the charge storage in the collector. The current blocking effect and means to eliminate it will be discussed in the following section.

2.3.3 Optimized Layer Structure (C)

The results of layer structure B show, that the new emitter and base layer designs led to some of the desired improvements. The remaining deficiencies are mostly due to the collector design. Therefore, emitter and base were not changed in the final structure (C).

Collector Layer

The discussion of structure B identified the phenomenon of current blocking as a problem, which is due to the conduction band potential barriers in the collector. It has been shown that a potential barrier in the collector can cause a high saturation voltage in the output characteristic (compare Fig. 2.14) and a low current density at the onset of current gain compression [41]. Therefore, a reduction of this effect would lower the saturation voltage and enable transistor operation at higher current densities and higher speed.

In addition to the step graded collector, which is already implemented in layer structures A and B, sophisticated doping profiles can reduce current blocking. If the doping level and the resulting electric field is appropriate, blocking can be overcome [68]. However, if the field intensity is insufficient, current blocking cannot be fully suppressed [65], which is probably the case in structure B. On the other hand, if the electric field strength is too high, the collector cannot be fully depleted at a low collector bias, which is likely the reason why the turn-on characteristic in layer structure A is poor.

Consequently, a collector layer structure was chosen, which is a compromise between layer A and B. The total length of the collector was hardly changed to keep high values for f_T. However, a thin moderately doped InP layer was introduced adjacent to the InGaAsP step grading. The layer is 20 nm thick and is also used in layer structure A. The new layer structure (C) is shown in Fig. 2.17.

Experimental Results

The output characteristic of a device with an emitter area of 10 μm^2 is shown in Fig. 2.18. The result does not show the blocking effect observed in layer structure B and the saturation voltage is 0.7 V. The turn-on voltage is 0.2 V. A comparison of the output characteristics of all three layer structures at $I_B = 200\mu A$ is given in Fig. 2.19. Figs. 2.20 and 2.21 show the Gummel plot and the current gain, respectively. The Gummel plot is compared with the results from structure B. The data show that the blocking effect is much weaker in structure C. The leakage currents in structure C are still below 1 nA and the ideality factors for base diode and collector diode extracted from the Gummel plot at a current of 1 μA are 1.4 and 1.1, respectively. The comparison of the current gains for all three structures shows, that the maximum gain is 90 in the optimized structure (C), higher than in the other structures.

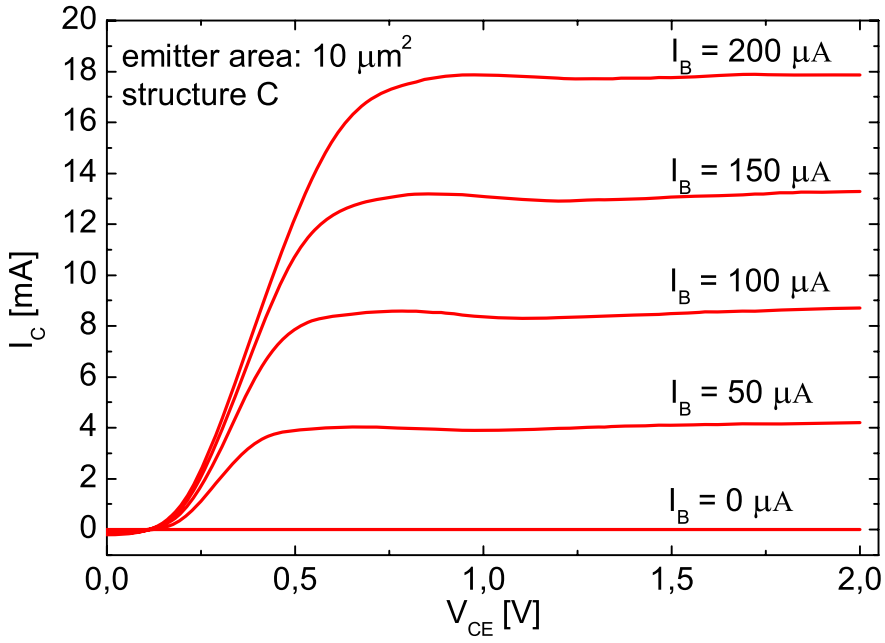


Figure 2.18: Output characteristic of a transistor from layer structure C.

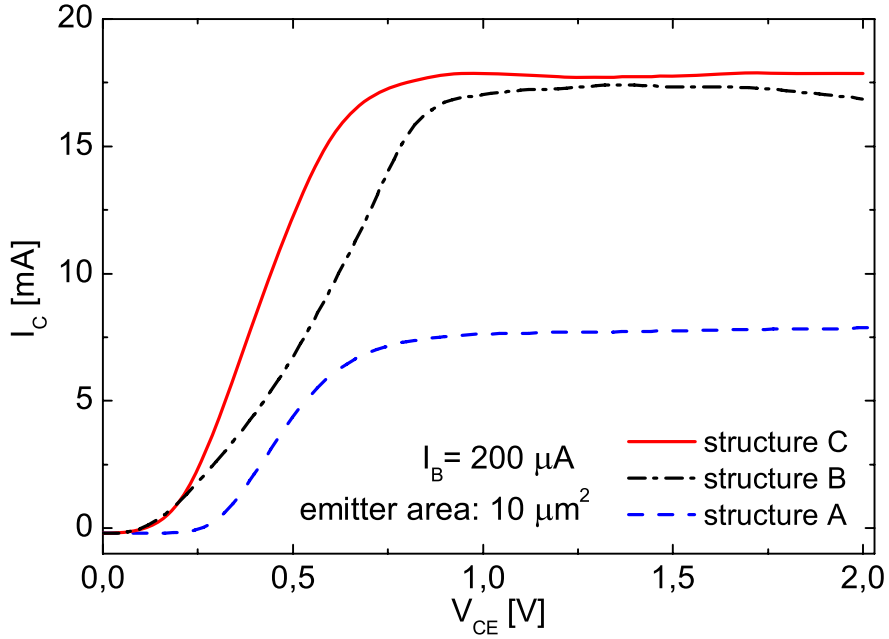


Figure 2.19: Comparison of output characteristics from different layer structures at a base current of 200 μA. Layer structure C displays the lowest turn-on and saturation voltage as well as the highest current gain.

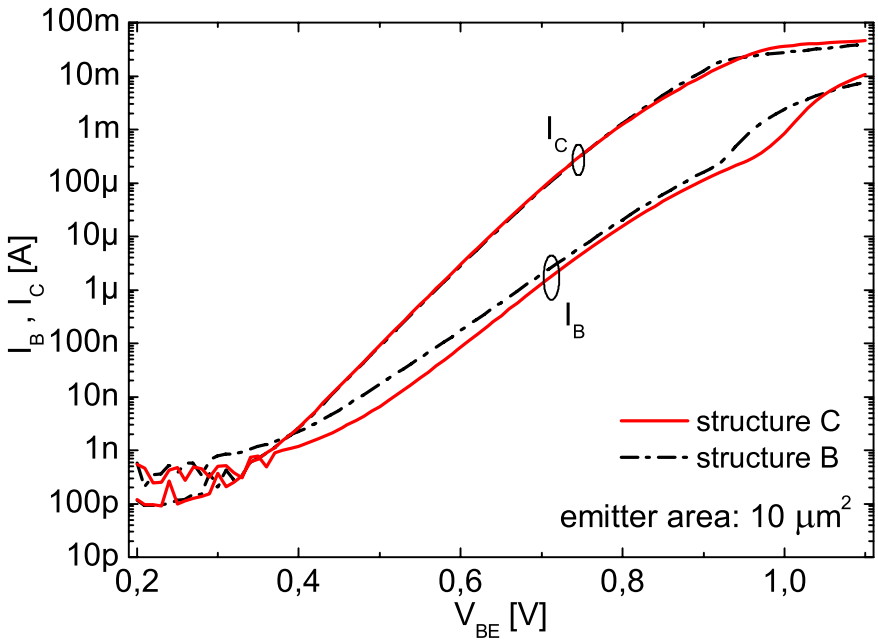


Figure 2.20: Comparison of the Gummel plots from layer structure B and C. The current blocking in structure C is only weak, as the kink at 1 V is moderate.

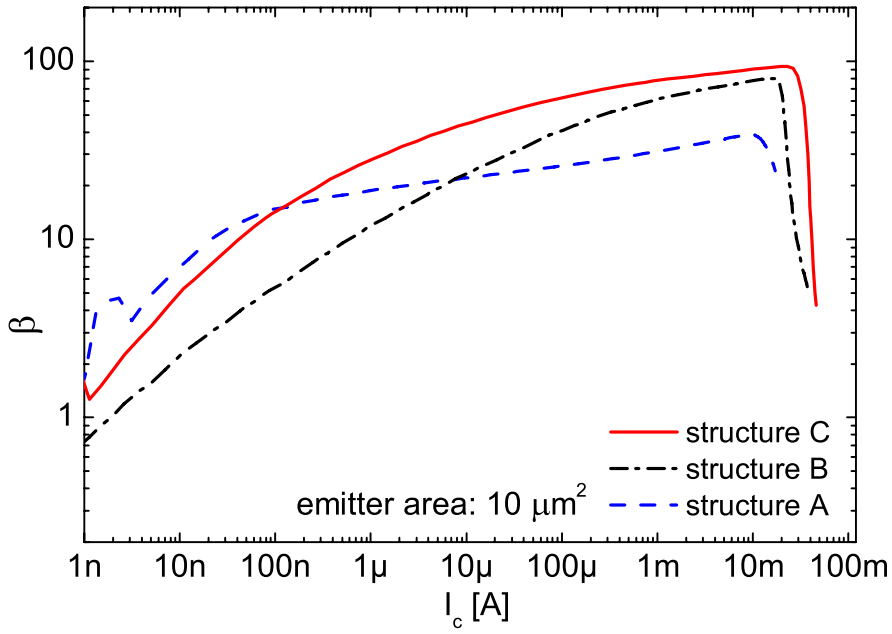


Figure 2.21: Comparison of the measured current gains from all three structures. The current gain of structure C reaches 90 at its maximum and the current compression sets in at higher currents than for the other structures.

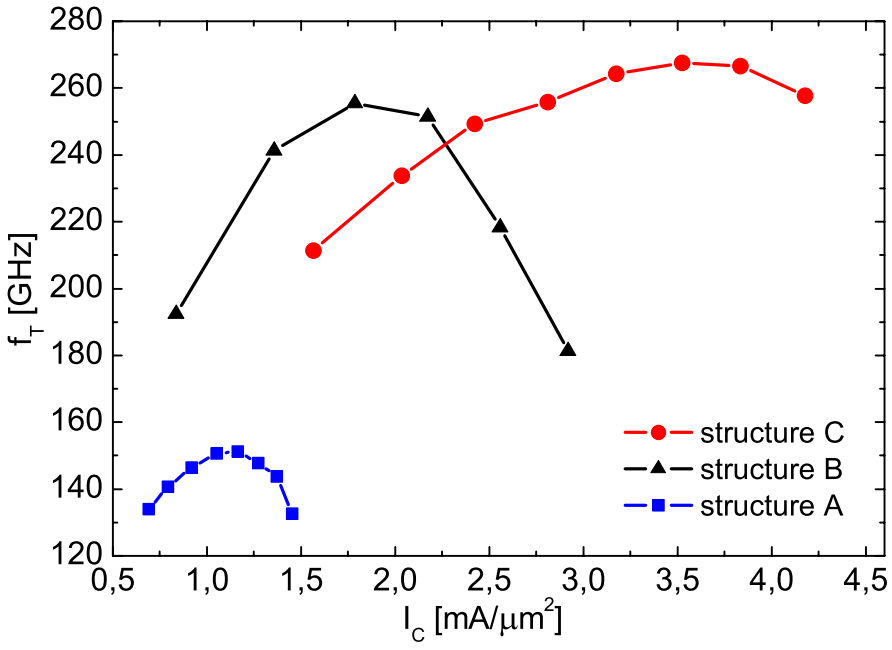


Figure 2.22: Comparison of cut-off frequencies of transistors from three different layer structures.

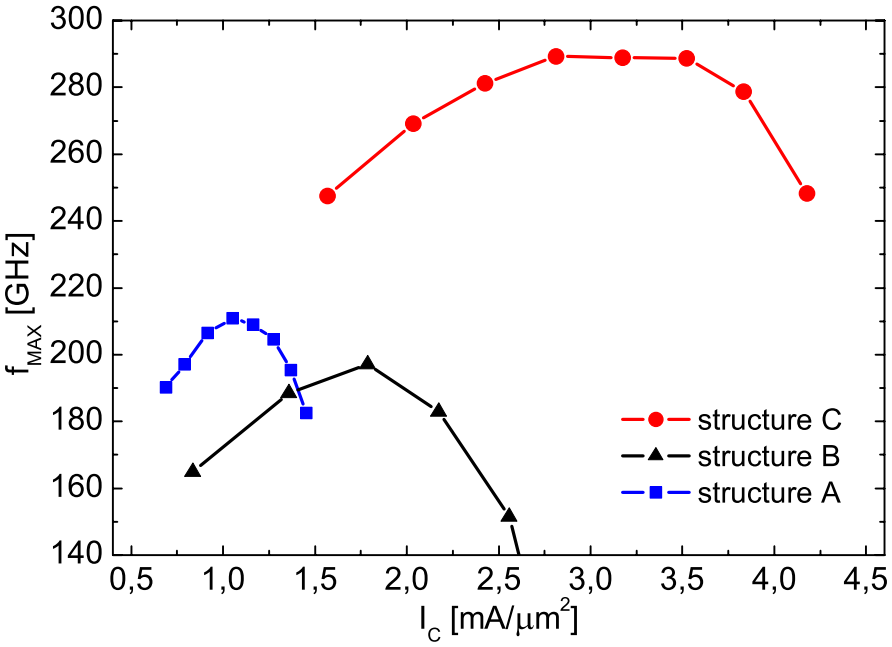


Figure 2.23: Comparison of $f_{\max}$ of transistors from three different layer structrures.

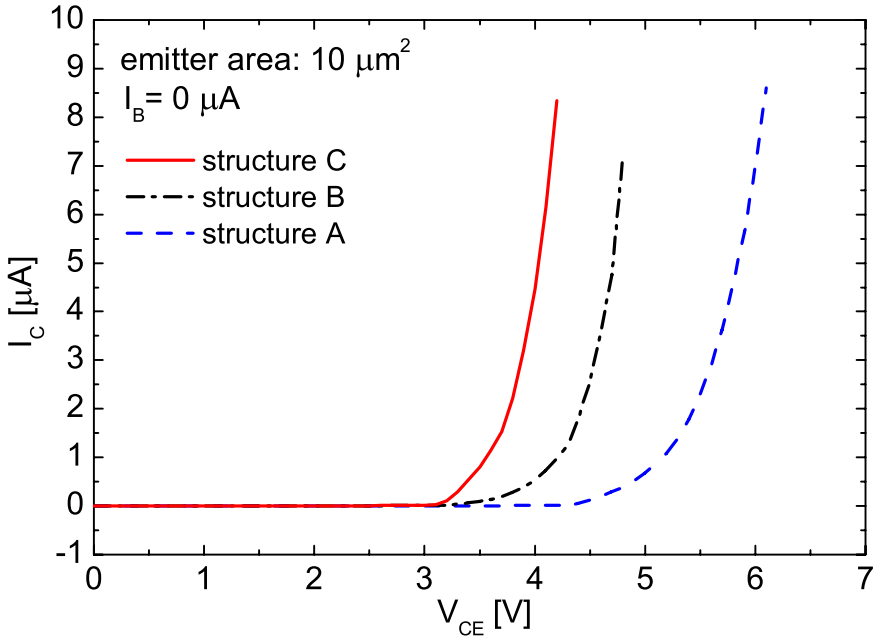


Figure 2.24: Comparison of collector-to-emitter breakdown voltage of the three structures. The devices were measured with a collector current compliance of 10 μA. The resulting breakdown voltages are 6.1 V for structure A, 4.8 V for structure B, and 4.2 V for structure C.

The results of the RF measurements are shown as a function of current density in Figs. 2.22 and 2.23. The devices on structure C reach the highest speed at a collector current density of 3.0...3.5 mA/μm^2, reaching f_{T} and f_{max} values of 267 and 288 GHz, respectively. The comparison with layer structures A and B reveals, that the current density as well as the speed is gradually increased from structure A to C.

Discussion

The change in collector design resulted in significant improvements. The saturation voltage was decreased and the maximum frequency of oscillation was shifted upwards, significantly. Both improvements were realized by alleviating the current blocking effect. The electrons can cross the heterojunction barrier at low operating voltages, already, which has a positive effect on the saturation voltage. Also, the current gain compression is moved to higher current densities, enabling higher operating speed due to a lower value of the differential resistance R_{je} at the emitter base junction (Eq. 2.5).

The elimination of current blocking prevents the accumulation of electrons at the barrier in the collector. This reduces the small-signal value for the base-collector capacitance significantly and is considered the reason for the tremendous increase in f_{max} when comparing structures C and B (Fig. 2.23).

Structure C is optimized for dc and high-frequency operation. However, a fundamental limitation of transistors is the trade-off between breakdown voltage and f_{T} [79]. Accordingly, conservative measurements of the collector-to-emitter breakdown voltage BV_{CE0} with a current compliance of 1 μA/μm^2 show a drop from 6.1 V for the initial structure (A) to 4.2 V for layer structure C (s. Fig. 2.24). From Table 2.1 it can be seen, that the breakdown voltage decreases as the high-frequency parameters f_{T} and f_{max} increase. In Fig. 1.4 these results are compared to other semiconductor technologies. Apparently, the tradeoff between high-frequency performance and breakdown voltage is a fundamental rule for HBTs and HEMTs for all listed semiconductor technologies. The smaller and faster the devices get, the larger are the electrical field intensities inside the devices, and the smaller is the breakdown voltage.

2.3.4 Summary

The optimization process of the layer structure with respect to power and high speed is summarized in Table 2.1. It shows the major changes that were made in the layer structure during the design process. Direct-current and high-frequency measurements were used to analyze each layer structure. The analysis revealed limiting factors, which were subsequently removed step by step. Correspondingly, the most important performance parameters are listed, too. The data demonstrate that a thin emitter, a compositionally graded base and a thin collector including a very thin moderately doped InP layer are important factors in the realization of high-speed DHBTs. The final optimized structure features high-frequency parameters in excess of 260 GHz, which enable the development of digital as well as analog circuits at very high speed. At the same time the breakdown voltage

Summary of Layer Structure Optimization Process

	structure A	structure B	structure C
emitter length (nm)	70	40	40
doping	no	yes	yes
base length (nm)	40	30	30
composition	uniform	graded	graded
collector length (nm)	240	150	160
"delta" doping	yes	no	yes
turn-on voltage (V)	0.3	0.2	0.2
saturation voltage (V)	0.7	0.8	0.7
current gain	39	80	90
f_{T} (GHz)	150	255	267
f_{max} (GHz)	210	195	288
BV_{CE0} (V)	6.1	4.8	4.2

Table 2.1: Comparison of the most important design and performance parameters of layer structures A, B, and C. More details of the layer structures are given in the schematics in Figs. 2.9, 2.13, and 2.17.

is 4.2 V enabling the development of modulator drivers at power levels not conceivable in silicon based technologies (s. Fig. 1.4).

The layer structure, i.e. the vertical dimensions of a device, have been optimized in this section. However, the lateral dimensions of a transistor are also important. The total area covered as well as the geometry of a transistor influence its high-frequency performance. For this reason the next section deals with the optimization of the transistor layout.

2.4 Transistor Layout Optimization

Transistor layout is a critical issue in the development of high-performance transistors. These dimensions reflect in the parameter values of the small-signal equivalent circuit in Figure 2.4. Hence, the layout influences the device speed.

The geometry dependence of device performance has been investigated theoretically for high-speed devices and a good overview of this topic is given in [89]. Also several experimental results have been published. Among them are investigations of geometrical optimization of InP DHBTs with regards to speed [27], [15]. Also, for other HBT technologies like SiGe and GaAs geometrical investigations have been conducted [112], [49], [103], [81]. In addition, experimental data regarding power performance and thermal resistance as functions of geometrical dimensions are available [109].

However, it is very difficult to calculate the optimum dimensions of InP DHBTs for a desired performance for the following reasons. First, the investigations show, that the optimization depends on the epitaxial layer structure. Second, besides layer structure

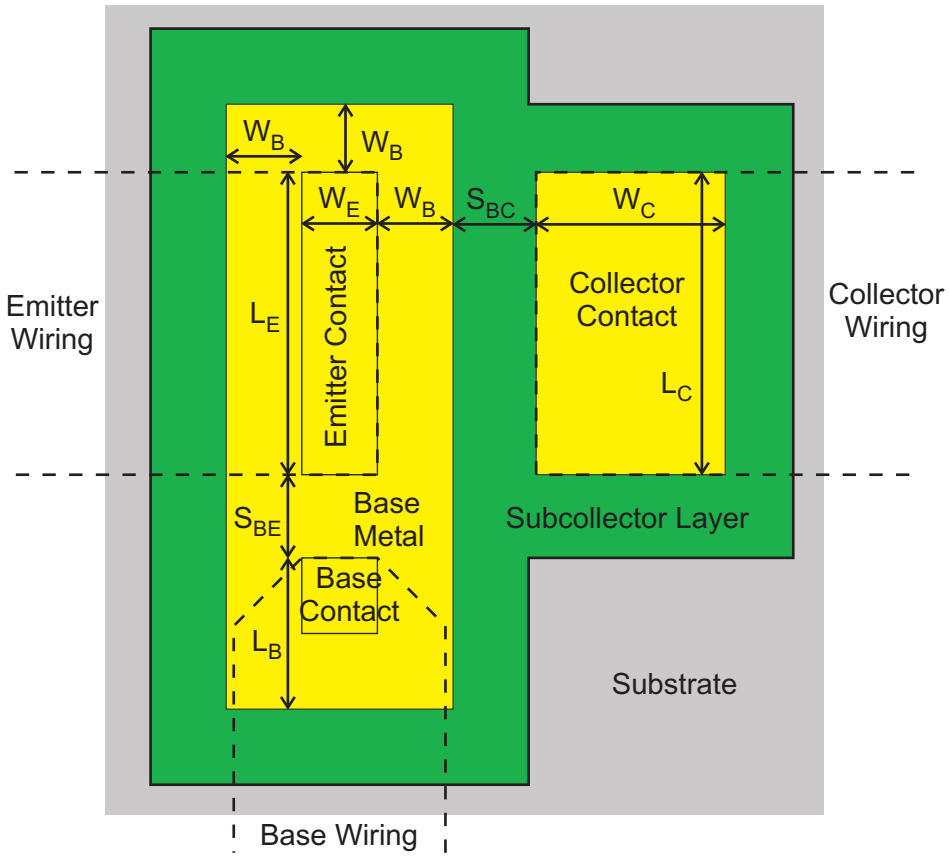


Figure 2.25: Schematic of HBT layout (top view).

and lateral dimension, process parameters like contact resistances and undercut etching can have a major impact on the results. In conclusion, the data from literature can give guidelines but not accurate numbers for the geometrical layout of InP DHBTs optimized for highest speed.

Therefore, an analysis of devices with varying lateral dimensions was conducted. The epitaxial layer structure in this investigation is the optimized layer structure discussed in the previous Section (Fig. 2.17). The optimized layout results in a device speed with parameter values of more than 280 GHz for f_{T} and f_{max}.

The layout under investigation is shown in Fig. 2.25. It consists of rectangular shapes. An advantage of such a rectangular layout is the fact that extracted models scale with transistor size to a certain degree (s. Sec. 3.3). In the following sections the dependence of the high-frequency parameters f_{T} and f_{max} on emitter size (L_E, W_E) and base contact width (W_B) is investigated. The results confirm that a proper layout is critical for high-frequency operation and they indicate, how the performance could be further improved. The results are summarized in the last section.

2.4.1 Emitter Width Design

The design of the emitter width W_E is guided by two major considerations. Current crowding effects as well as high values of the internal base resistance R_{bb} can degrade the transistor performance.

Current crowding effects result from a voltage drop over the finite resistance of the base layer caused by the base current flowing into the center of the base emitter junction. The effective base emitter voltage at the center of the junction is lower than at the edges. As a result the current density at the edges is higher than in the center and reduces the effective base emitter junction area. In other words, the base-emitter junction capacitance C_{je}, which is proportional to the area of the junction, is larger than it needs to be. From Equation 2.5 it can be seen, that this has negative consequences for f_{T}. Another technical drawback of a distributed nature of the emitter would be, that a more complicated transistor model than used in this work (Ch. 3) would be necessary.

The second straight forward consideration is the maximization of $f_{\max}$. For a fixed emitter area $f_{\max}$ will be largest if R_{bb} is minimized (Eq. 2.6), i.e. if the ratio of perimeter to area is maximized. In conclusion, the narrowest emitter should result in the best high-frequency performance.

Experimental Results

Three different transistors with an emitter width W_E of 2.0, 1.0, and 0.7 μm were analyzed. The devices investigated had an emitter length L_E of 4 μm and a base width W_B of 0.75 μm. The base and collector contact dimensions are described in Sections 2.4.2 and 2.4.3.

The results are shown as a function of current density in Figs. 2.26 and 2.27. The smaller W_E is, the larger the maximum values for f_{T} and $f_{\max}$ are. The cut-off frequency and the maximum frequency of oscillation increase from 214 to 235 GHz and from 210 to 280 GHz, respectively. Also, the current densities at which the maxima occur shift to higher values as the emitter width decreases: from 3.5 to 4.1 mA/μm^2 for f_{T} and from 3.0 to 3.7 mA/μm^2 for $f_{\max}$.

Discussion

The results show a strong dependence of the high-frequency performance on the emitter width W_E. Especially $f_{\max}$ increases significantly for the narrow emitters. This increase can be attributed to the reduction of the internal base resistance R_{bb} due to the change in perimeter to area ratio of the emitter (Eq. 2.6). The changes in cut-off frequency and current density probably result from a reduction of current crowding. The narrower the emitter the more uniform the current distribution in the emitter is. As discussed above, this leads to higher current densities and higher f_{T}. However, these effects have a weaker influence on device performance than the reduction of R_{bb}. The results confirm the assumption, that the emitter should be as narrow as possible.

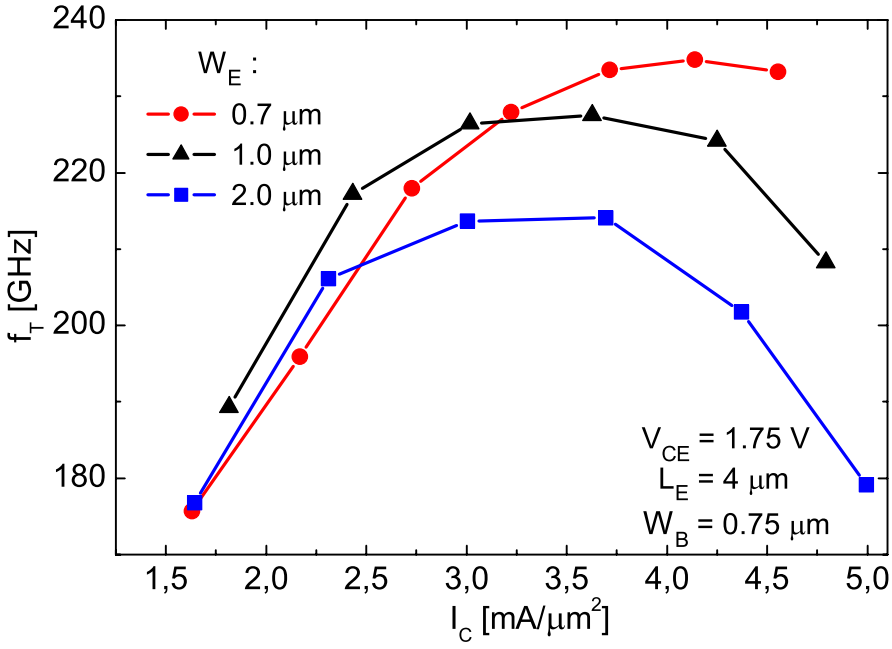


Figure 2.26: Measured cut-off frequencies for different emitter widths.

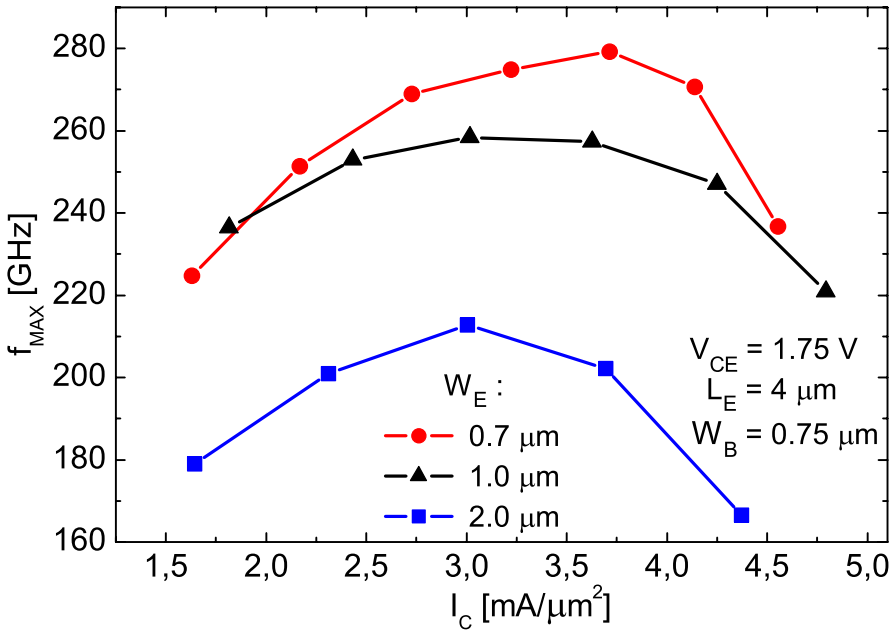


Figure 2.27: Measured maximum frequency of ozcillation $f_{\max}$ for different emitter widths.

2.4.2 Base Contact Width Design

Besides the emitter size, the base contact width W_B and the base contact pad (L_B, S_{BE}) determine the dimensions of the internal transistor. Its size is proportional to the collector capacitance C_{bc}, which affects the high-frequency parameters of the transistors (Eqs. 2.5 and 2.6). The smaller C_{bc} is, the higher the values for f_{T} and f_{max} will be. From this point of view, to achieve high cut-off frequencies, W_B needs to be as small as technologically possible.

Depending on the quality of the base metal semiconductor contact, the base contact width W_B can have a major impact on the base resistance R_B. The transfer length of a metal semiconductor contact is the length over which the current transfers from metal to semiconductor. In general, this length is longer for contacts displaying a high contact resistance and very short for contacts with negligible contact resistance [12]. Therefore, if the base contact width is smaller than the transfer length of the base metal contact, the parasitic base resistance R_B will increase. This would result in a decrease of f_{max}, even though f_{T} might rise. Hence, W_B should not be smaller than the transfer length of the contact.

The lengths L_B and S_{BE} are determined by the technology. L_B and S_{BE} add to the area of the base collector junction which is proportional to C_{bc}. So, they should be as small as possible, but large enough to ensure a reliable base contact to achieve high yield in processing. For the current process the values of L_B and S_{BE} are 2.0 and 1.5 μm, respectively. An SEM picture of a HBT showing the base metal is shown in Fig. 2.5.

Experimental Results

Three different transistors with a base contact width W_B of 1.0, 0.75, and 0.5 μm were investigated. The devices had an emitter length L_E and emitter width W_E of 8 and 1 μm, respectively. The collector contact dimensions are described in Section 2.4.3. All measurements were conducted with a collector-to-emitter voltage of 1.75 V.

In Figs. 2.28 and 2.29 the results of the high-frequency measurements are shown as a function of current density. Both parameters f_{T} and f_{max} show an approximately linear increase for decreasing values of W_B. The cut-off frequency and the maximum frequency of oscillation increase from 225 to 280 GHz and from 245 to 295 GHz, respectively. The current density at which the maxima occur is almost constant at ~ 3.7 mA/μm^2 for both f_{T} and f_{max}.

Discussion

The results document very clearly, how important the reduction of the parasitic base collector capacitance is. Without changing the intrinsic transistor the performance could be increased, remarkably. The increase is of about the same magnitude for f_{T} and f_{max} and confirms that the assumptions based on Eq. 2.6 predict the high-frequency performance, correctly. Additionally, the transfer length of the base metal semiconductor contact is probably shorter than 0.5 μm, because the increase in f_{max} appears to be linear and not

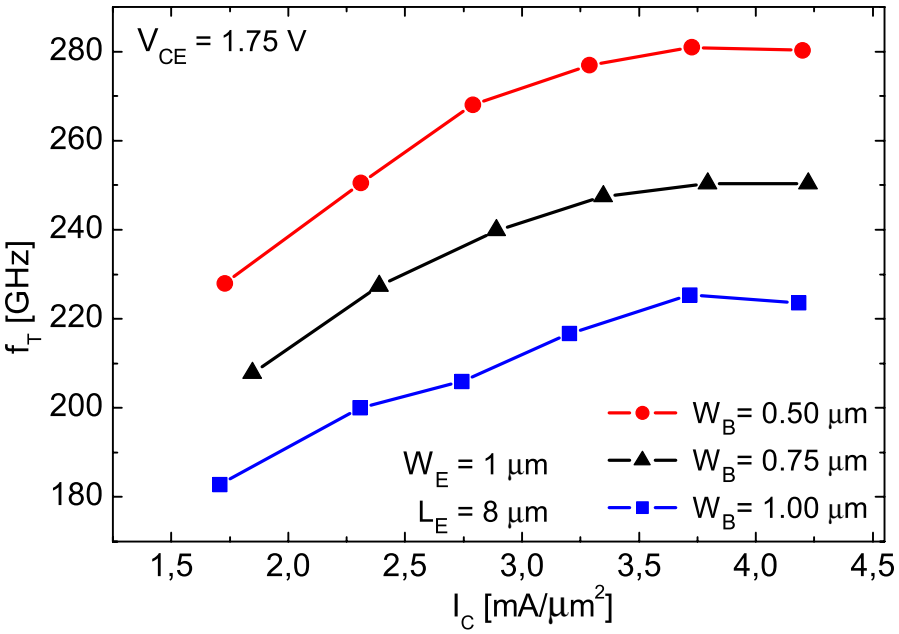


Figure 2.28: Measured cut-off frequency f_{T} for various base contact width values.

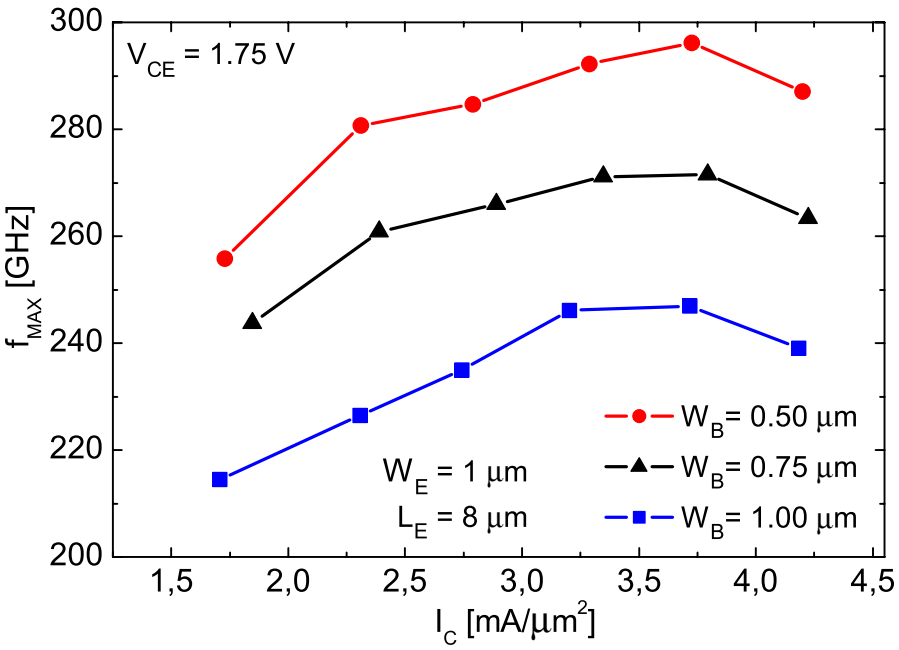


Figure 2.29: Measurements of maximum frequency of oscillation f_{max} for various base contact width values.

saturated for the smallest base width. Therefore, a more advanced process, which can realize shorter values for W_B, should enable a further increase in high-frequency parameters. The fact that the current density for all maxima is the same, supports the argument of the previous section, that the current density dependence is determined by the emitter layout, which is not changed in this section. In summary, the reduction of base contact width is very effective way of increasing the high-frequency performance of HBTs.

2.4.3 Collector Contact Design

The collector contact design is determined by three parameters S_{BC}, W_C, and L_C (Fig. 2.25). The contact is outside the active part of the device. The contact geometry affects mostly the extrinsic collector resistance R_C, which should be small to enable high-speed performance (s. Eq. 2.5).

Consequently, the spacing S_{BC} should be as small as possible and has been fixed to 1 μm in this technology. The collector is designed to have twice the width of the emitter ($W_C = 2W_E$). This is large enough to ensure low resistance values and small enough to enable compact circuit layouts. The collector length is set equal to the emitter length ($L_C = L_E$) to provide a uniform current distribution in the collector.

2.4.4 Emitter Length Design

From a theoretical point of view, the emitter length L_E is expected to have the following effects on the device performance. If the device is very small, the influence of the base contact pad (L_B, S_{BE}) will be significant (Fig. 2.25). The experimental results from Section 2.4.2 show clearly, that the parasitic base collector capacitance is one limiting factor for high-frequency performance. On the other hand, a very long emitter might reduce the device speed. Resistive and inductive effects in the narrow and long base contacts along the emitter would be conceivable. Hence, especially for the optimization of L_E an experimental investigation is inevitable.

Experimental Results

Seven different transistors with emitter lengths varying from 2 to 16 μm were measured. The investigated devices had an emitter width W_E of 1 μm and a base width W_B of 0.75 μm. The base and collector contact dimensions were described in Sections 2.4.2 and 2.4.3. All measurements were conducted with a collector-to-emitter voltage of 1.75 V.

The high-frequency results for four different emitter lengths are compared in Figs. 2.30 and 2.31. Even though the current densities at which the maxima occur are not the same, a clear tendency is not apparent. The maximum value of f_T starts at 225 GHz for $L_E = 4\,\mu$m and seems to saturate at 250 GHz for large emitter length values. The maximum frequency of oscillation seems to reach a maximum of 270 GHz for an intermediate emitter length.

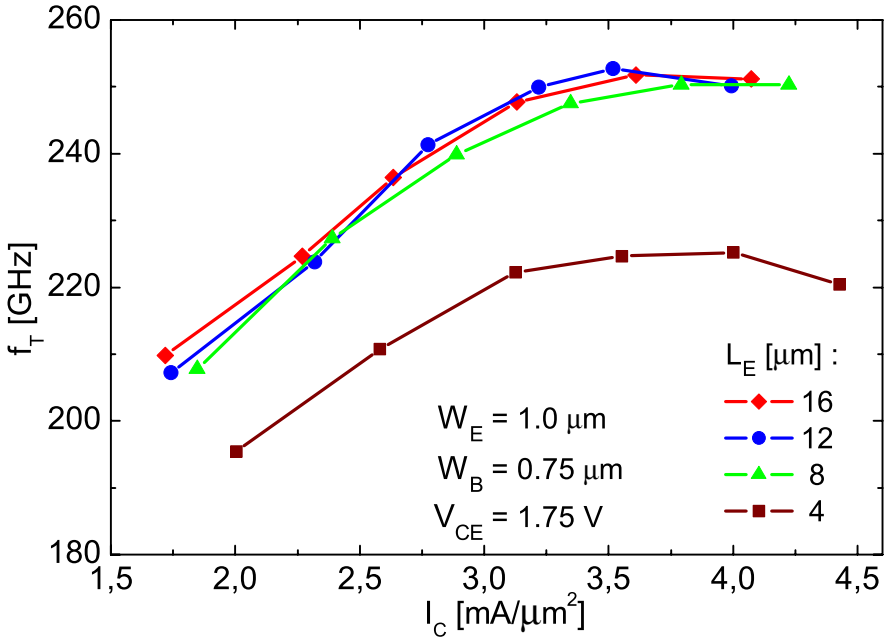


Figure 2.30: Dependence of f_{T} on emitter length.

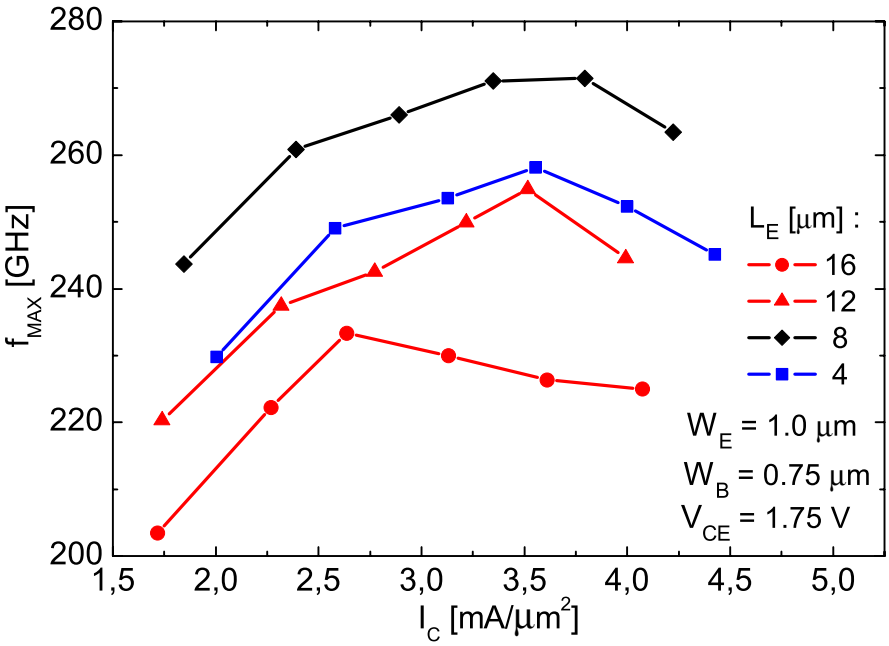


Figure 2.31: Dependence of f_{max} on emitter length.

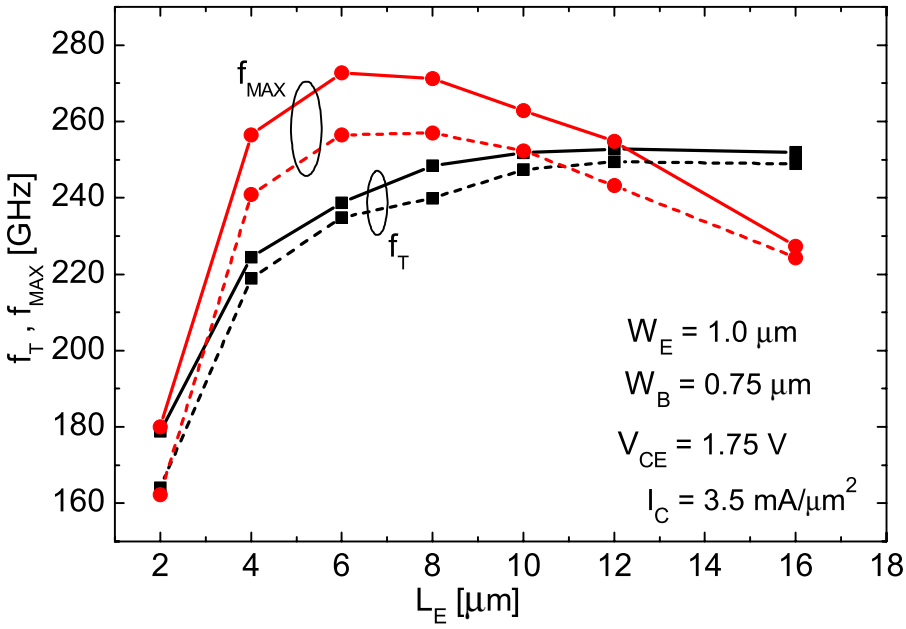


Figure 2.32: f_T and f_max vs. emitter length (solid lines: measurement, dashed lines: calculation).

To clarify this behavior, the high-frequency parameters for seven different devices at a current density of 3.5 mA/μm^2 are plotted as a function of emitter length L_E in Fig. 2.32. The cut-off frequency increases with emitter length, saturating at 250 GHz for long emitters ($\geq$ 8 μm). The maximum frequency of oscillation shows a maximum of more than 270 GHz at an emitter length of 6 μm with a sharp decrease for shorter emitters and a moderate decline for longer emitters.

For a more detailed analysis, small-signal models (s. Sec. 3.3) were extracted for each transistor at the operating points that correspond to the measured data points in Fig. 2.32. Using the extracted model parameters the high-frequency parameters f_T and f_max were calculated according to Equations 2.5 and 2.6 and included in Fig. 2.32. Calculated and measured data for f_T match closely and although the calculation of f_max differs up to 10 % from the measured data, the principal behavior is described correctly.

Discussion

The results of this section reveal significant information. The length of the intrinsic emitter seems to have a much stronger influence on f_max than on f_T. The decrease of both high-frequency parameters for short emitter length can be attributed to the collector-base capacitance of the base contact pad (s. Fig. 2.25), which has the same size for all devices. The shorter the emitter gets, the greater the influence of the additional parasitic capacitance is.

Focusing emitter lengths larger than 6 μm, the scaling behavior can be seen. The cut-off frequency f_T seems to be almost independent of the emitter length, while f_max

gradually decreases with emitter length. A possible reason for the latter effect are resistive and inductive effects in the narrow base contacts along the emitter (s. Fig. 2.25).

The scaling behavior of the high-frequency parameters can be explained using the small-signal parameters. A detailed discussion of these parameters and how they affect the results summarized in Fig. 2.32 is given in Sec. 3.3, which deals with the small-signal model. Apparently, the choice of the emitter length is a compromise of minimizing the influence of the parasitic base contact pad and the influence of resistive or inductive effects. The latter effects seem unavoidable in high-speed HBTs with long emitters, as a narrow base contact is a requirement for high speed (s. Sec. 2.4.2). The reduction of parasitic capacitances depends also on the process technology. Therefore, it is difficult to propose a general design rule for the emitter length, and the choice of transistor size depends on the circuit design requirements. For driver amplifiers critical considerations in addition to f_{T} and f_{max} are discussed in Sec. 6.1.

2.4.5 Summary

The experimental investigation of the transistor geometry (Fig. 2.25) produced the following results: as predicted by theoretical considerations, **the base contact width W_B and the emitter width W_E** in Fig. 2.25 **should be as small as possible**. However, while the influence of W_B on both frequency parameters is comparable, W_E has a much stronger effect on f_{max} than on f_{T}.

The choice of the emitter length L_E depends on the size of the base contact pad (L_B,S_{BE}) and the base width W_B. If the emitter is too short, both f_{max} and f_{T} degrade very strongly (s. Fig. 2.32) due to the parasitic capacitance of the base contact pad. For very long emitters the maximum frequency of oscillation declines due to inductive and resistive effects in the base contact.

The development of complex circuits requires high yield of devices. Depending on circuit concepts amplifiers can consist of more than 50 transistors. The failure of one device can lead to a malfunction of the whole circuit and a layout that can be manufactured reliably is crucial. Therefore, the geometrical dimensions of the devices used in amplifier design are chosen so, that the technological process provides a high yield on the device level. This is achieved by slightly relaxing the dimensions in comparison to the smallest achievable values used in this investigation. The devices intended for circuit design are identical to those, whose frequency performance is shown in Fig. 2.32. **The final choice of a particular transistor length L_E at a certain point in a circuit depends on the needed current, the speed requirements (f_{T}, f_{max}), and also on input and output impedance of the devices** (s. Sec. 6.1).

The discussion of this chapter shows, that the optimization of heterojunction bipolar transistors involves several issues, which are correlated. Process technology, the epitaxial layer structure and the geometrical dimensions have to be optimized. First, different epitaxial layers were investigated using high-frequency and dc measurements. From the results, improved layer structures could be designed. Second, the geometrical layout of the transistors was optimized by investigating different transistors manufactured on the

optimized layer structure. As a result, InP DHBTs with high-frequency parameter values of more than 280 GHz for both f_{T} and f_{max} could be realized, while maintaining a breakdown voltage of over 4 V.

Chapter 3

Transistor Models

The transistor model is an important link between the process technology and the complex circuit. On the one side it has to accurately describe the single device and can assist the optimization and investigation of the transistors in a particular technology. On the other side it is a necessary tool which makes the simulation and development of complex circuits possible.

In principle, the modeling task could be fulfilled by a 3D-simulation of the carrier and field distribution based on the physical parameters of each semiconductor layer and the topology of the transistor structure. Particularly with regard to circuit simulation, such an approach results in unacceptable long simulation times [85]. Therefore, a "compact" model is needed, which describes the terminal behavior rather than the internal device physics.

Background

Although this work deals with analog circuits, its interaction with related research topics has an effect on its methods. Concurrently, digital circuits for data rates up to 80 Gbit/s were developed at Fraunhofer IAF using the InP DHBT technology discussed in Chapter 2. Therefore, the modeling effort ought to provide a transistor model, which was equally suited for the development of analog as well as digital circuits. Low-power transimpedance amplifiers (TIA) and distributed amplifiers for higher output power levels like modulator drivers operate in the linear region of the devices. In contrast, digital circuits like multiplexers and demultiplexers are based on the nonlinear behavior of the transistors. Other building blocks of transmission systems like voltage controlled oscillators (VCO) and clock and data recovery circuits require an accurate description of both kinds of characteristics.

Application to Circuit Design

The model described in this chapter was applied in the development of the amplifier circuits presented in Chapters 5 and 6 [94], [95]. In addition, VCOs, multiplexers, and clock and data recovery circuits suited for 80 Gbit/s applications have been developed, using the

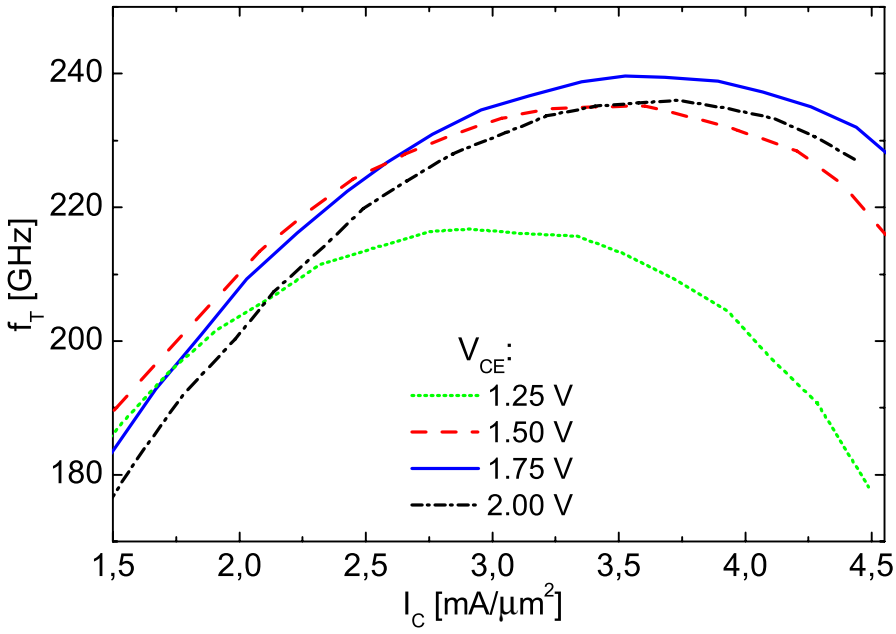


Figure 3.1: Measured cut-off frequency f_T of an InP-DHBT as a function of collector current density for different collector-to-emitter voltages.

same model [73],[74],[72]. This shows that the modeling strategy described in this chapter is a suitable approach for the development of high-speed circuits for data rates at 80 Gbit/s and beyond.

Content

In the following section different large-signal models are discussed. Following this, an extraction procedure for the UCSD model is presented. The procedure's starting point resulted from a collaboration with the University of California, San Diego, and it was adapted and refined to satisfy the needs of analog and digital circuit designs. In the last section, small-signal model results for transistor evaluation and optimization are presented.

3.1 Large-Signal Models

3.1.1 Gummel-Poon Model

The Gummel-Poon model was the first widely used large-signal model for bipolar transistors and is still considered a standard model for such devices. This model was proposed for silicon bipolar transistors by Gummel and Poon in 1969 [38]. It is available in virtually all circuit simulators. However, DHBTs based on III-V technology display characteristics which are not accounted for by this model.

Limitations

One of the salient features of InP- or GaAs-based HBTs, in contrast to Si-based bipolar transistors, is the strong bias dependence of the collector delay. This dependence is due to the velocity modulation effect [46], which describes a non-monotonic dependence of the electron velocity on the electric field. At low electric field all electrons reside in the so called Γ-valley of the valence band. If electrons gain enough energy they can scatter into a higher energy state in the L-valley. The effective electron mass in the L-valley is much higher than in the Γ-valley. In III-V semiconductors this reduces the electron velocity tremendously resulting in an increase of τ_c.

As the collector delay represents a significant portion of the total delay of an HBT, its variation is clearly visible in the f_T characteristics of the device [47]. The resulting non-monotonic dependence of f_T on bias can be seen in Fig. 3.1. The graph shows that the high-frequency performance has a maximum at a collector-to-emitter voltage of 1.75 V and a collector current density of 3.6 $\mathrm{mA}/\mu\mathrm{m}^2$. In contrast to this behavior, the Gummel-Poon model assumes that the high-frequency performance improves with higher operating voltages at any given collector current level [68]. In addition, the model presumes that the dc current gain is independent of collector current over a wide range, which is not true for HBTs. Therefore, the Gummel-Poon model is not well suited for modeling InP-DHBTs.

3.1.2 Alternative Models

To overcome these limitations the following alternative large-signal models for HBTs have been proposed: With a focus on the accurate description of the collector delay time, a new model was developed at the Ferdinand-Braun-Institut für Höchstfrequenztechnik [90]. Partially built on the Gummel-Poon model other advanced models carefully treating self-heating and impact ionization effects were published [34], [51]. Based on previous work by Camnitz [18] the University of California, San Diego (UCSD) developed another model in collaboration with several industrial partners. The UCSD model [101] is the first model, which is dedicated to DHBTs in compound semiconductor technologies and which has also been implemented in several commercially available simulators, among them Smart Spice by Silvaco. Meanwhile, the capabilities of the UCSD model were further extended [47]. This altered version has been implemented under the name AgilentHBT model in the simulation software ADS by Agilent. In the following, the original version as well as the latter version, are called UCSD model, because both models are almost identical, if a suitable subset of model parameters is used.

3.1.3 UCSD Model

In this work the UCSD model was used. It is available in two simulators, which were used for digital and analog circuit design, SmartSpice by Silvaco and ADS by Agilent. Other models described in Sec. 3.1.2 were not available in more than one simulation environment. Using commercial models helps to avoid convergence problems, which can occur if model

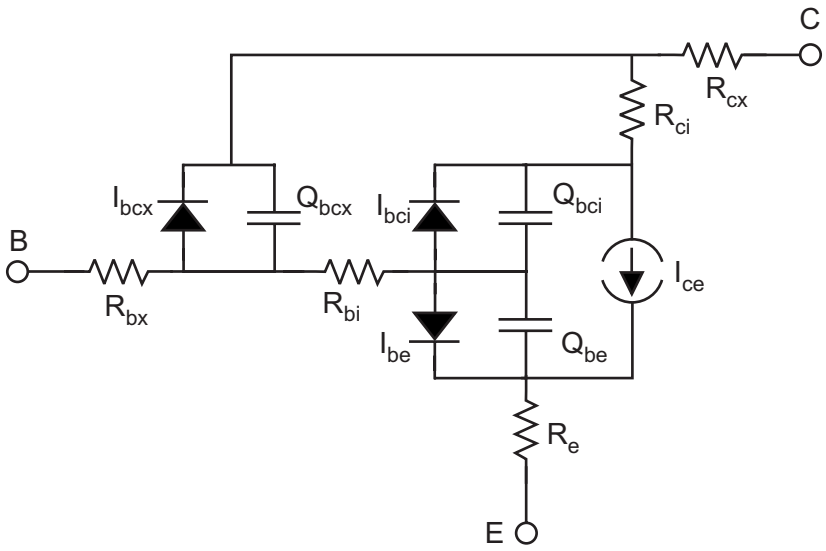


Figure 3.2: Equivalent circuit of the UCSD model displaying those components, that were utilized for modeling.

equations are not properly implemented. The model equations are freely available on the internet [101]. The most important model equations and parameters are described in Sec. 3.2.

Features

The model comprises features dedicated to DHBTs in III-V technology. A variety of transit time parameters are provided to model the velocity modulation, which was already discussed above (s. Sec. 3.1.1). The collector current equation of the standard Gummel-Poon model is replaced by a more complex formulation. It takes into account the potential spike, that can occur at the heterojunctions in HBTs, which causes a variation of current gain with current density. In addition, thermal effects including self-heating and the Kirk effect [55], which is responsible for the roll-off of the high-frequency parameters at high currents, can be modeled.

Topology

In contrast to the small-signal equivalent circuit (s. Fig. 2.4), the large-signal model uses a different topology for the intrinsic device (s. Fig. 3.2), which is called Π-topology. Large signal models often inherit their Π-topology for historical reasons from the Gummel-Poon model. However, there exists a one-to-one correspondence between the T-Model and Π-Model [100]. Therefore, either topology can be used to model a transistor and for both topologies an equally accurate model can be extracted [35]. The T-topology of the small-signal model is derived from the device geometry (s. Fig. 2.3).

3.2 UCSD Model Extraction

The extraction of a large-signal model was guided by the following considerations. First, an extracted model has to be suitable for both simulators, ADS (analog circuits) and SmartSpice (digital circuits). Second, only a subset of the vast number of more than 90 parameters should be used. This reduces the time that is needed for model extraction as well as the simulation time. On the other hand, the accuracy of the model requires a minimum amount of model parameters. Hence, a compromise between accuracy and complexity was made, which also allowed an easy transfer of models between simulator environments.

For model extraction, the following simplifications were made. Model features concerning thermal effects were switched off. The parasitic capacitances between the transistor ports and parasitic inductances at the transistor ports were neglected. Although these parasitics have finite values, their behavior was modeled by adjusting the other parameters. In addition many parameters, which did not significantly enhance the model accuracy, were disabled. A list of parameters that are being used and their extracted values are given in Appendix A.

The extraction of a model can follow different strategies. The model extraction is generally assisted by the evaluation of dc and high-frequency measurements in different operating points [97], [96]. The use of test structures in addition to transistor devices has also been reported [23]. Numerical optimization of the parameters can also be applied [35], [47].

A collaboration with the UCSD provided a preliminary extraction procedure for the InP-DHBT used in this work. Subsequently, the extraction procedure was refined and adjusted. In particular, test structures were designed to assist the extraction of base and collector resistances, which could not be extracted satisfactorily before. The deembedding procedure for high-frequency measurements was adopted from the small-signal model extraction (s. Sec. 3.3). Additionally, the transistors were stressed before extraction to produce devices, which resembled transistors in circuits under stress. The final extraction procedure is described for a device with an emitter size of 8×1 μm^2 in the following sections.

3.2.1 Diode Parameters

The ideal diode characteristics, which are extracted from the Gummel plot, predict a linear rise of diode currents, if plotted on a log scale as in Fig. 3.3. Therefore, the extraction was performed at low currents, where the linear behavior can be observed. At high currents, other effects, e.g. due to series resistances, influence the Gummel plot and deteriorate the ideal diode characteristic.

The reverse Gummel plot is shown in Fig. 3.4. The base current is accurately modeled for medium and high current densities. The simulation and measurement of the emitter current do not match. The main reason is the fact, that the UCSD model does not provide a diode parameter, which is exclusively dedicated to this current. The parameter determining this current is IS, which also determines the collector current in the forward

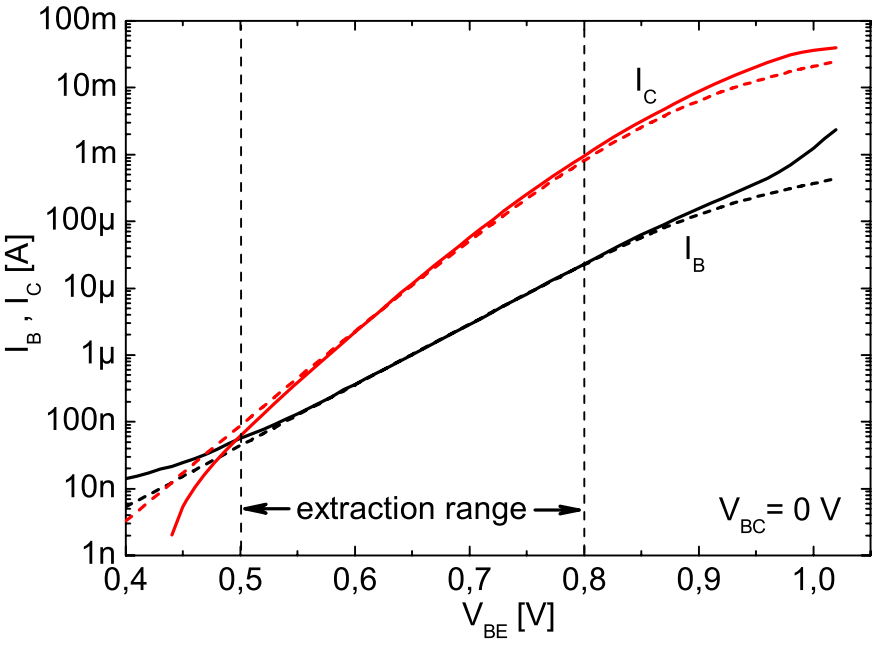


Figure 3.3: Gummel plot comparing measured (solid lines) and simulated data (dashed lines). The model parameters IS, NF, ISE, and NE are extracted at low current densities from this plot in the marked range.

Gummel plot. As the forward active mode is the regular mode of operation in circuits, IS was extracted from the forward Gummel plot.

One of the main features, which make the UCSD model distinct from the Gummel-Poon model, is the formulation of the equation for the collector-to-emitter current I_{CE} (s. Fig. 3.2). If Webster and Early effects are neglected (compare Sec. 2.1), the equation simplifies to:

$$I_{CE} = \frac{\exp\left(\frac{q \cdot V_{be}}{NF \cdot k \cdot T}\right) - \exp\left(\frac{q \cdot V_{bci}}{NR \cdot k \cdot T}\right)}{\frac{1}{ISA}\exp\left(\frac{q \cdot V_{be}}{NA \cdot k \cdot T}\right) + \frac{1}{ISB}\exp\left(\frac{q \cdot V_{bci}}{NB \cdot k \cdot T}\right)} \tag{3.1}$$

The four parameters in the denominator NA, ISA, NB, and ISB are new parameters which are intended to describe the influence of the semiconductor heterojunctions. V_{be} and V_{bci} are the potential differences between the internal nodes at the base-emitter and intrinsic base-collector diode. NF and NR are extracted from the Gummel diagrams (Figs. 3.3 and 3.4). The other parameters are the electric charge (q), Boltzmann's constant (k), and the temperature (T).

Theoretically, the parameters NA and ISA, which account for the influence of the base-collector heterojunction, could also be extracted from the forward Gummel plot. However, as can be seen from Fig. 3.3, the extracted model does not accurately describe the latter plot at high currents. This is due to the fact, that the base-collector bias condition at the Gummel plot ($V_{BC} = 0$ V) is different from the bias condition in forward active

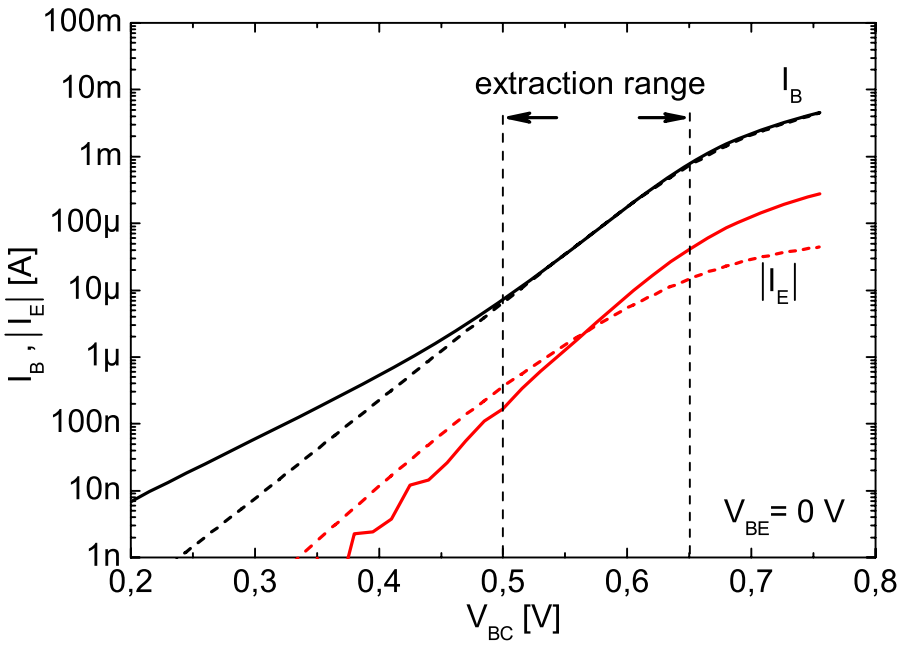


Figure 3.4: Reverse Gummel plot comparing measured (solid lines) and simulated data (dashed lines). The model parameters ISC, NC, and NR are extracted at medium current densities, where these parameters dominate the transistor behavior.

mode ($V_{BC} < 0V$) in Fig. 3.5. The use of more model parameters should reduce the mismatch in Fig. 3.3. However, as this large-signal model is intended for circuit design, the forward active mode (Fig. 3.5) needs to be described properly and the less accurate description of the Gummel plot is considered acceptable. In the output characteristic the gradual decrease of the measured collector current with increasing V_{CE} is regularly observed and is due to self-heating.

3.2.2 Resistances

Often, the extraction of resistance is achieved by measuring the S-parameters of the device in open collector mode or other special operating points. Using such purely S-parameter based methods, we did not achieve satisfactory extraction results. Therefore, we developed test structures, which allowed to get access to the resistance parameters by eliminating the influence of other parameters. Such methods have been successfully used before [23].

The test structures were designed to resemble the transistor as closely as possible. Still, the values that were measured were considered only close approximations and not accurate values. However, we found that these values enabled the extraction of a model. Nevertheless, as a final extraction step, numerical optimization using the Levenberg-Marquardt algorithm [82] was applied to the resistance parameters to improve the accuracy of the model. As a result, some of the resistance parameters changed moderately. The optimized values are listed in Tab. A.1.

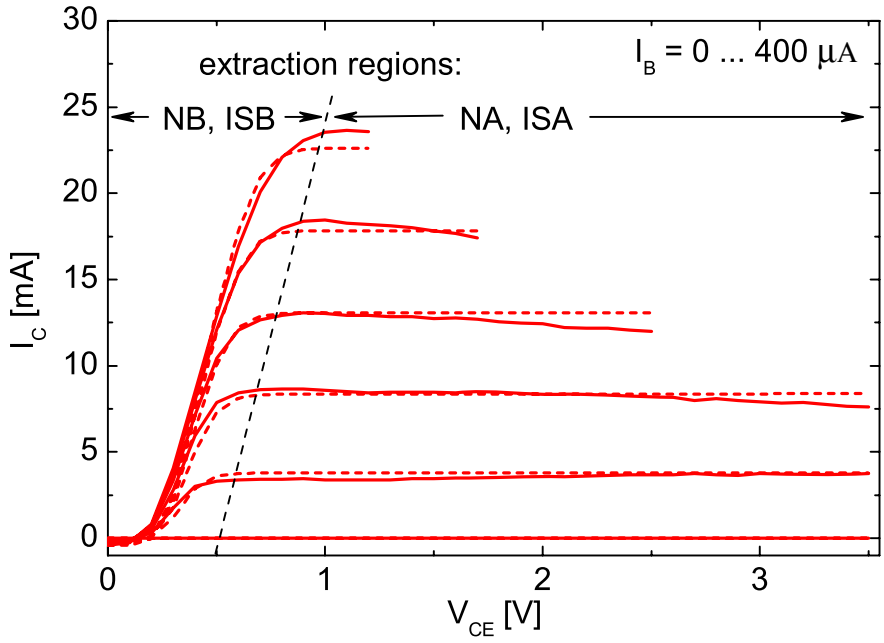


Figure 3.5: Comparison of measured (solid lines) and simulated data (dashed lines). The current source parameters NB, ISB, NA, and ISA are extracted in the marked regions.

Collector Resistance

The collector resistance comprises two components. The extrinsic resistance accounting all resistances outside the internal transistor and the intrinsic collector resistance associated with the subcollector layer resistance underneath the extrinsic base-collector junction (s. Fig. 2.3). In the large-signal model the parameters RCX and RCI present these resistances. They are extracted using a test structure shown in Fig. 3.6(a). By measuring such test structures of different dimensions the total resistance can be subdivided into an extrinsic and intrinsic part, which are defined as a subcollector layer sheet resistance $R_{C\square}$ and a contact resistance R_{Ccon}. Tab. 3.1 shows the measurement results and the derived resistances. In the derivation possible current spreading in the semiconductor layers was neglected, because the doping of the subcollector is very high and the anticipated resistance is small.

Using the estimated values for $R_{C\square}$ and R_{Ccon}, the resistance parameters RCX and RCI

measurement results			derived parameters	
y, z [μm]	$y=2$, $z=8$	$y=4$, $z=8$	$R_{C\square}$ [$\Omega/\square$]	20
$2\frac{R_{Ccon}}{z}+\frac{y}{z}R_{C\square}$ [Ω]	14	19	R_{Ccon} [$\Omega\mu$m]	36

Table 3.1: Results of dc resistance measurements of the collector resistance test structure in Fig. 3.6(a) and corresponding derived resistance parameters. The results are average values from six separate measurements.

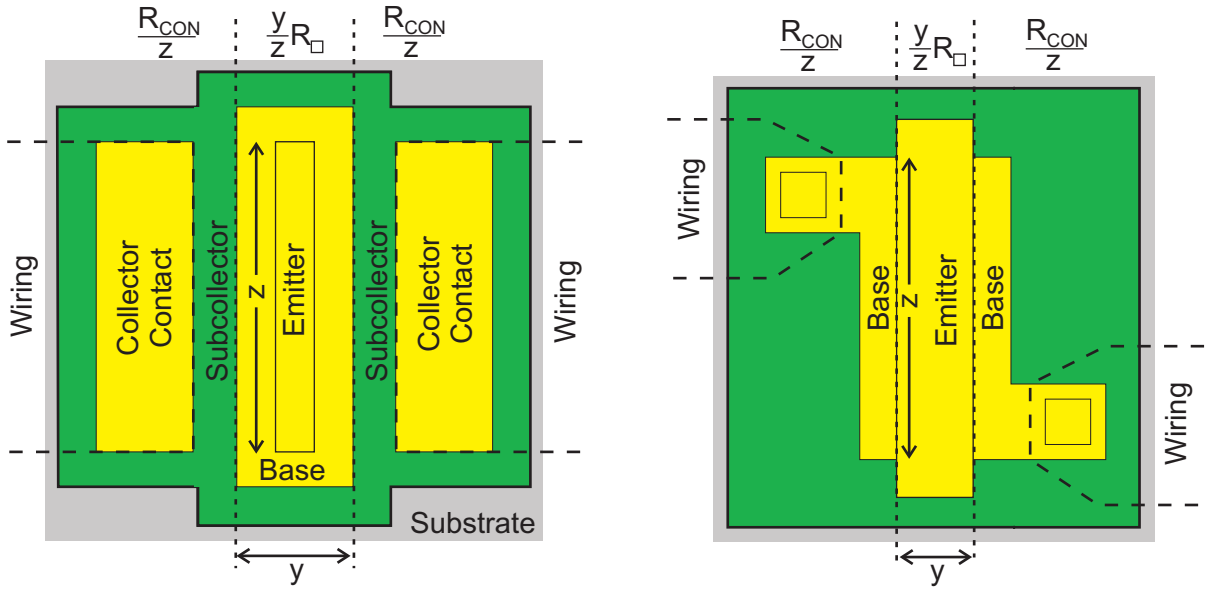


Figure 3.6: Schematic of test structures for extracting the resistance parameters RCX, RCI, RBX and RBI.

are calculated from the geometrical dimensions of the transistor:

$$RCX \approx \frac{R_{Ccon}}{L_E} \tag{3.2}$$

$$RCI \approx \frac{R_{C\square}}{2}\frac{W_E}{L_E} \tag{3.3}$$

W_E and L_E are the emitter width and length (s. Fig. 2.25).

Base Resistance

The base resistance is divided in two components: the intrinsic base resistance RBI and the extrinsic base resistance RBX. These parameters are extracted using a split base test structure (s. Fig. 3.6(b)). The emitter-base junction of such a structure is very similar to a regular transistor. However, the base contact is split into two parts by extending the length of the emitter. From measuring such test structures with different values for y and constant z, approximations of the base resistance parameters can be derived. Tab. 3.2 summarizes the results of several measurements and the inferred values for the sheet resistance of the base layer ($R_{B\square}$) and the metal-semiconductor contact resistance (R_{Bcon}). Any resistances of the base metal and wiring are neglected as other test structures revealed that their sheet resistance values are smaller than 0.2 $\Omega/\square$. Also, possible lateral current spreading effects at the edges of the metal-semiconductor contacts were neglected.

To calculate the intrinsic base resistance RBI, the spacial contribution of the base current has to be considered. Its current density decreases from the edges of the base-

measurement results			derived parameters	
y, z [μm]	$y=1$, $z=8$	$y=2$, $z=8$	$R_{B\square}$ [$\Omega/\square$]	880
$2\frac{R_{Bcon}}{z}+\frac{y}{z}R_{B\square}$ [Ω]	155	265	R_{Bcon} [$\Omega\mu$m]	180

Table 3.2: Results of dc resistance measurements of the base resistance test structure in Fig. 3.6(b) and corresponding derived resistance parameters. The results are average values from six separate measurements.

emitter interface reaching zero in the center. For a narrow emitter the intrinsic base resistance can be calculated using the following formula [68]:

$$RBI \approx \frac{R_{B\square}}{12}\frac{W_E}{L_E} \tag{3.4}$$

The extrinsic base resistance parameter RBX is estimated by dividing R_{Bcon} by the perimeter of the base-emitter junction:

$$RBX \approx \frac{R_{Bcon}}{2(W_E+L_E)} \tag{3.5}$$

Emitter Resistance

The emitter resistance parameter RE accounts for the emitter resistance R_E, which is the sum of metal contact resistance and epitaxial layer resistances in the emitter cap layers (s. Figs. 2.3 and 2.17). The present technology (self-aligned) does not allow to design a test structure to directly measure the emitter resistance. Hence, the emitter resistance is determined using a method proposed by Maas [70] and often applied in parameter extraction [92], [35], [105]. At low frequencies the real part of the small-signal two-port parameter Z_{21} of an HBT can be approximated by the sum of the emitter resistance R_E and emitter-base junction resistance R_{je}. The latter is inversely proportional to the emitter current I_E. Therefore:

$$\Re\{Z_{21}\} \approx R_E + R_{je} = R_E + \frac{1}{\delta I_E}, \tag{3.6}$$

where δ is a constant of proportionality.

By plotting $\Re\{Z_{21}\}$ versus $1/I_E$ and extrapolating the graph towards $1/I_E = 0$, the emitter resistance R_E can be extracted (s. Fig. 3.7) and its value is taken to be the parameter RE. The measurement at any frequency in the range from from 3 to 10 GHz resulted in comparable resistance values.

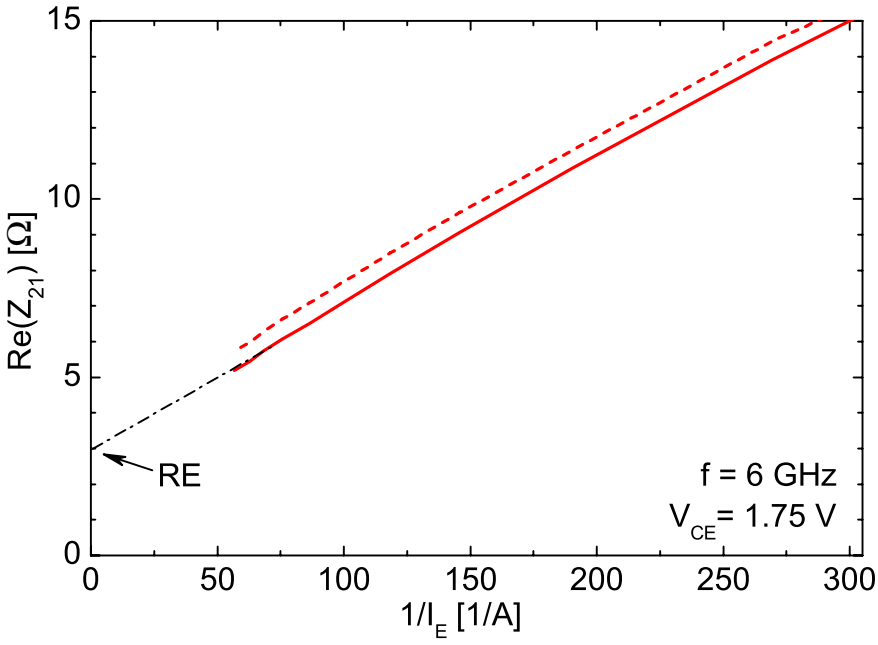


Figure 3.7: Graph showing the extraction of the emitter resistance by extrapolating the measured data (solid line) to $1/I_E = 0$. The simulation (dashed line) using the extracted model confirms that this measurement well approximates the parameter RE.

3.2.3 Junction Capacitances

The small-signal capacitance of an abrupt pn-junction (C_{pn}) as a function of applied reverse voltage (V_{bias}) can be calculated to be [35]:

$$C_{pn} = A_{pn} \frac{k}{\sqrt{\Phi + V_{bias}}} \tag{3.7}$$

where A_{pn} is the area of the junction, k is a function of the acceptor and donor doping levels, and Φ is the built-in voltage.

However, HBTs often feature complicated doping profiles (s. Sec. 2.3). Depending on the profile Eq. 3.7 does not accurately describe the capacitance characteristics. Consequently, the model equations offer several additional parameters, which can be used to achieve an accurate extraction.

An established method for the extraction of the junction capacitance are reverse bias measurements [20]. In reverse bias a minimal amount of current is flowing and the capacitive elements in Fig. 3.2 dominate the device's response. In this state the small-signal capacitance values can be directly calculated from S-parameters. After measuring S-parameters and converting them into Y-parameters the base-emitter and base-collector capacitances (C_{je}, C_{bc}) are given by the following equations:

$$C_{je} = \frac{\Im\{Y_{11} + Y_{12}\}}{2\pi f} \tag{3.8}$$

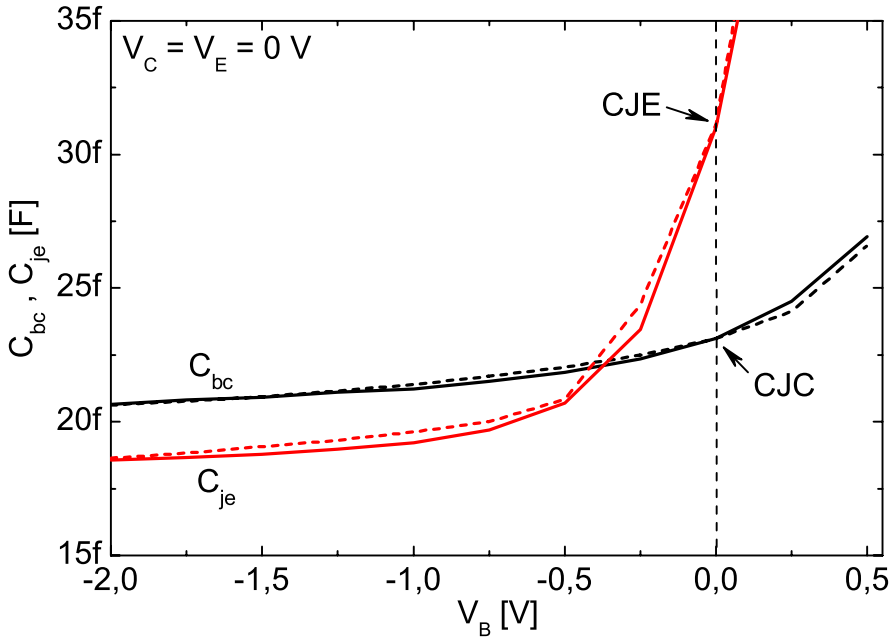


Figure 3.8: Comparison of measured (solid lines) and simulated data (dashed lines). The capacitance parameters CJC and CJE are extracted at zero bias. All other parameters related to the junction capacitances are extracted by fitting simulated to measured data.

$$C_{bc} = \frac{\Im\{Y_{12}\}}{2\pi f} \tag{3.9}$$

For model extraction the HBTs were measured under varying reverse bias and the associated junction capacitances were calculated according to Eqs. 3.8 and 3.9. To improve the accuracy of the results, the capacitance values were averaged over a frequency range of $f = 3...8$ GHz. The results are shown in Fig. 3.8. The capacitance values at zero bias correspond to the model parameters CJE and CJC. The other model parameters are fitting parameters, which were adjusted accordingly. As can be seen from the comparison of measured and simulated data, the parameter set of the large-signal model is well suited to describe the junction capacitance of the devices.

3.2.4 Transit and Delay Times

Finally, the delay time parameters are extracted. Due to the dependence of the collector transit time on collector current and collector-to-emitter voltage (s. Fig. 3.1) a large number of 25 model parameters is provided (s. Appendix A). This prohibits a direct extraction of each parameter and numerical optimization has to be used. However, an estimation of the most important delay time parameters is advisable. This speeds up the optimization process and increases the probability that an acceptable set of parameters is found, which accurately describes the transistor behavior.

Parameter Estimation

Three parameters were estimated: the base transit time τ_b, the collector transit time τ_c and the collector current at the onset of the Kirk effect I_K. The transit time τ_b is approximated using the base layer semiconductor parameters. According to [68], the transit time for electrons through a compositionally graded base (p-type) is given by the formulas

$$\tau_b = \frac{{W_B}^2}{2D_n} f(k) \; ; \; D_n = \mu_n \frac{kT}{q} \; ; f(\kappa) = \frac{2}{\kappa}(1 - \frac{1}{\kappa} + \frac{1}{\kappa} e^{-\kappa}) \; ; \; \kappa = \frac{\Delta E_G}{kT} \; , \tag{3.10}$$

where W_B is the thickness of the base layer, D_n is the diffusion constant for electrons, κ is the electric field factor, μ_n is the mobility of electrons, k is Boltzmann's constant, T is the temperature, q is the elementary charge, and ΔE_G is the change in bandgap due to the grading. W_B is 30 nm (Fig. 2.17). The electron mobility in highly p-type doped InGaAs-layers has been reported to be $\mu_n \approx 3000$ cm^2/(Vs) [13]. Although this result has not been confirmed by other authors, we are using this value for estimation. Considering the strain, which is introduced by the compositional grading specified in Sec. 2.3.2, and using reported experimental data [102] ΔE_G was estimated to be 50 meV. Applying all this information to equation above, it follows $\tau_b \approx 34$ fs, which was taken as starting point for model parameter TFB.

The transit time τ_c can be calculated using Eq. 2.5 and $R_{je} = \frac{1}{\delta I_E}$ (s. Eq. 3.6):

$$\tau_c = \frac{1}{2\pi f_{\mathrm{T}}} - \frac{1}{\delta I_E}(C_{je} + C_{bc}) - C_{bc}(R_E + R_C) - \tau_b \tag{3.11}$$

The resistance R_E equals the previously extracted parameter RE and R_C is approximated to be $RCX + RCI$. The factor $\frac{1}{\delta}$ is extracted from the slop of the graph in Fig. 3.7. The other parameters depend on the bias condition. For $V_{CE} = 1.75$ V and $I_E = 21$ mA the paramters f_{T} and C_{bc} were taken from Figs. 3.10 and 3.8, respectively. C_{je} was approximated to be equal to the small-signal parameter C_E (s. Fig. 3.17). This resulted in an estimation of $\tau_c \approx 205$ fs and was taken as the starting point for model parameter $TCMIN$.

The current I_K is taken from Fig. 3.10 and marks the current, at which the high-frequency parameter start to decrease (25 mA; starting point for model parameter $IKRK$).

Optimization

The extraction of the delay time parameters is performed by fitting the high-frequency parameters f_{T} and f_{max} over a wide range of operating points. The high-frequency parameters are obtained using a generally accepted method for high-frequency characterization of transistor devices [88]. First, the unilateral gain U and the two-port parameter H_{21} are calculated from the measured and simulated S-parameters. Then the curves are interpolated with a straight line at a slope of -20 dB/dec to the frequency points, where the parameters have decreased to unity (s. Fig. 3.9).

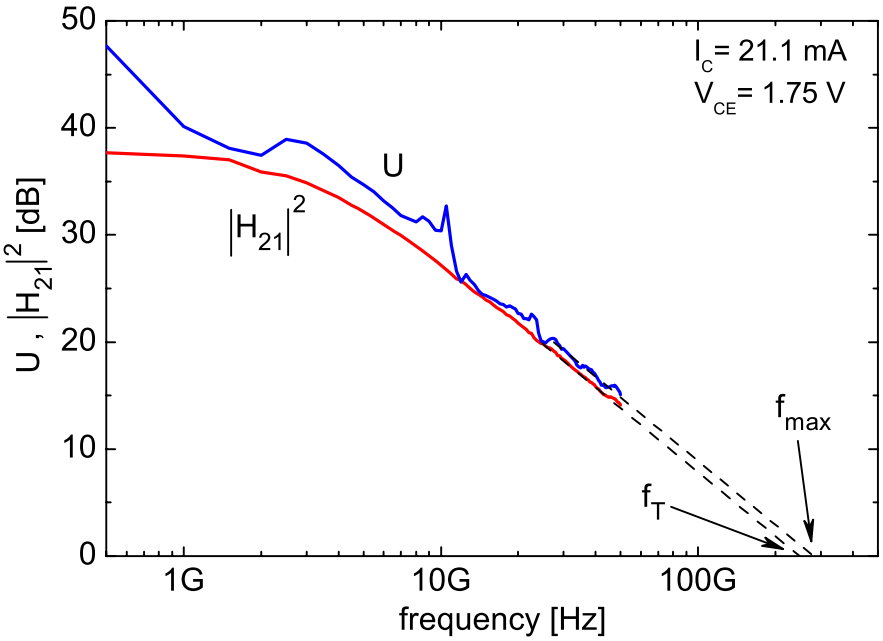


Figure 3.9: The graph shows how the high-frequency parameters f_T and f_max are obtained by extrapolating the small-signal current gain (H_{21}) and the unilateral gain (U).

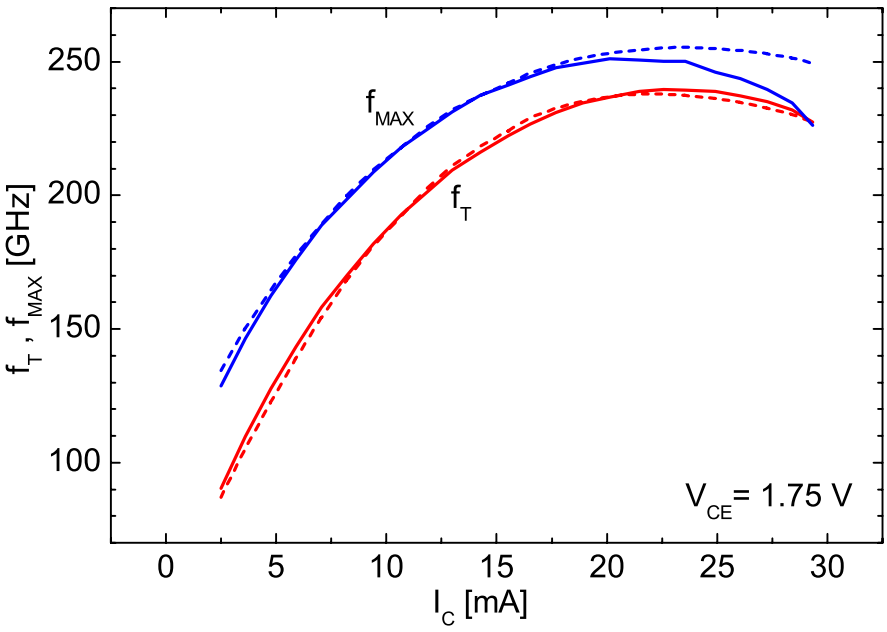


Figure 3.10: f_T and f_max as a function of collector current I_C. Comparison of measured (solid lines) and simulated data (dashed lines).

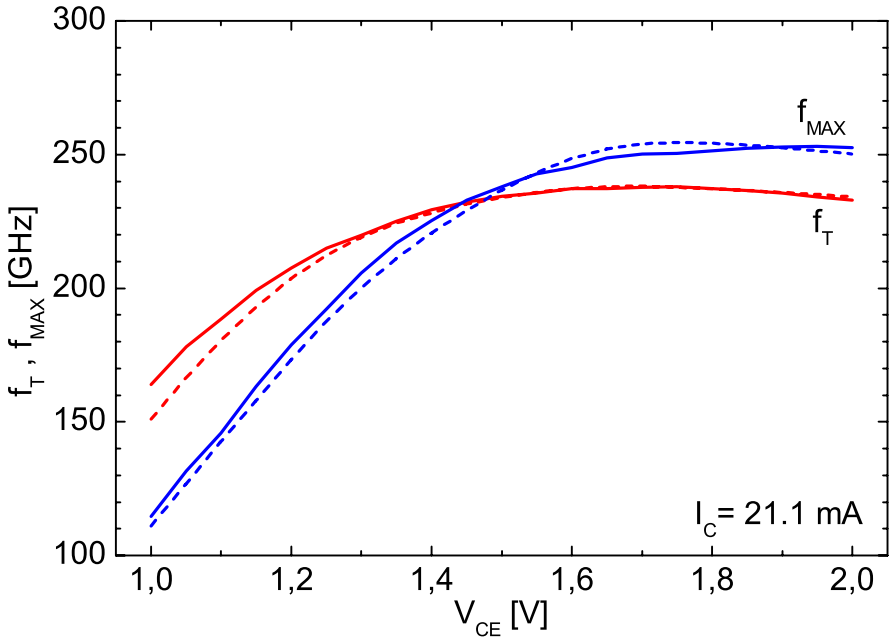


Figure 3.11: f_{T} and f_{max} as a function of collector-to-emitter voltage V_{CE}. Comparison of measured (solid lines) and simulated data (dashed lines).

The optimization process was started with the estimated parameter values from above and very few additional parameters. As expected the fit was very poor. Gradually, more and more parameters were included in the optimization and, simultaneously, the difference between measured and simulated data could be reduced. After allowing all delay parameters to be optimized, the curves shown in Figs. 3.10 and 3.11 were achieved. They demonstrate that the current dependence as well as the voltage dependence of the high-frequency parameters can be correctly modeled. The ability to model their non-monotonic dependence on V_{CE} (s. Fig. 3.11) is a feature of the UCSD model, which the Gummel-Poon model cannot provide.

Comparing the estimated and optimized values for TFB, $TCMIN$, and $IKRK$ suggests some interesting information about the model. The base transit time parameter and the Kirk current parameter have only moderately changed. It seems that these parameters have a physical meaning like the diode parameters which were extracted directly from measured data without optimization. In contrast, $TCMIN$ has dramatically changed, which indicates that it is a parameter, which is not directly related to the collector transit time. However, these observations are not supported by hard evidence and were not further investigated.

3.3 Small-Signal Model

The advantage of the large-signal model is its validity over a wide range of operating conditions described in the previous section and used in circuit simulation. A disadvantage of that model is its limited accuracy. In a special case described in this section, a so called small-signal model can be used. Among the advantages of this model are a reduced number of model parameters and a higher accuracy in comparison to the large-signal model.

Features

If the amplitudes of the signals at the input and output port of the transistor are small enough, the transistor behaves like a linear network. Such networks can be modeled using linear electrical elements (s. Fig. 2.4), are called small-signal models, and can be characterized using S-parameters (s. Sec. 4.2). Their extraction is fast due to a small number of model parameters. In addition, these models are very useful for evaluating and comparing different transistors. Their disadvantage is, that they can only be used, if the transistor is operated in small-signal mode. Consequently, small-signal models were used for circuit simulation in small-signal operation and for the evaluation of transistor properties.

3.3.1 Model Extraction

The transistors dimensions are very small and devices cannot be contacted directly by measurement equipment. Therefore, transistors intended for model extraction are embedded in an environment consisting of coplanar transmission lines, tapers and probe contact pads. Consequently, measured data always contains the influence of the transistor environment. In order to obtain data, which describes the transistor alone, the measured data have to be deembedded.

Deembedding

First, the transistor environment was analyzed using additional test structures and its A-parameters were calculated. Second, the measured S-parameters were converted to A-parameters and the the inverse of the environment's A-parameters were multiplied to both sides of the measured data. Converting the resulting A-parameters back to S-parameters yields a data set whose reference planes no longer reside at the contact pads, but are located right at the physical ports of the transistor. In this work only deembedded S-parameter data were used.

Extraction Method

The extraction of the small-signal equivalent circuit model can follow different strategies. Using network analysis the model parameters can be calculated to be functions of the measured S-parameter data [97], [17]. Another method is to apply numerical optimization techniques, which fit some or all parameters simultaneously, until the difference between

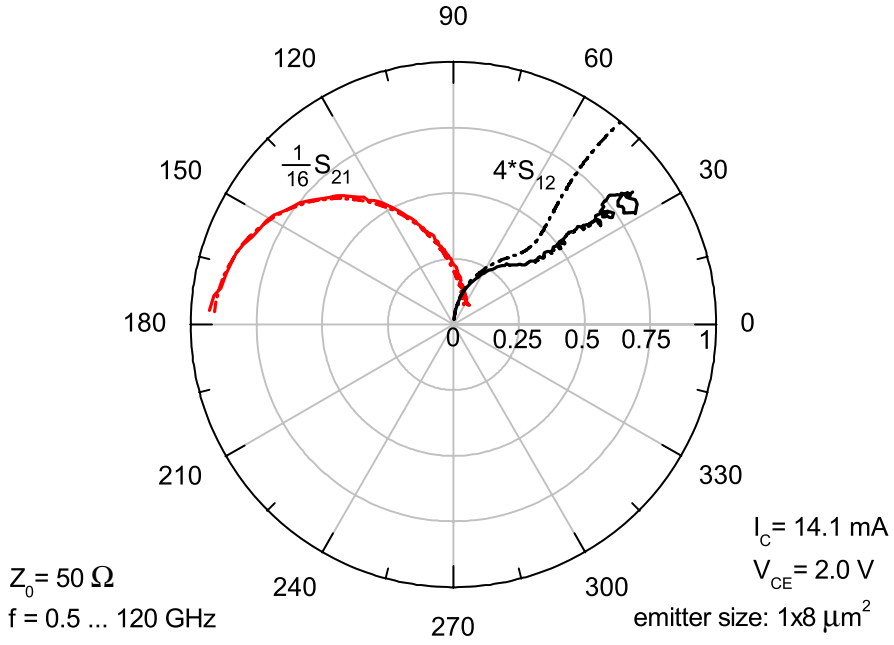


Figure 3.12: Measured (solid lines) and simulated (dashed lines) scattering parameters using the **small-signal model**.

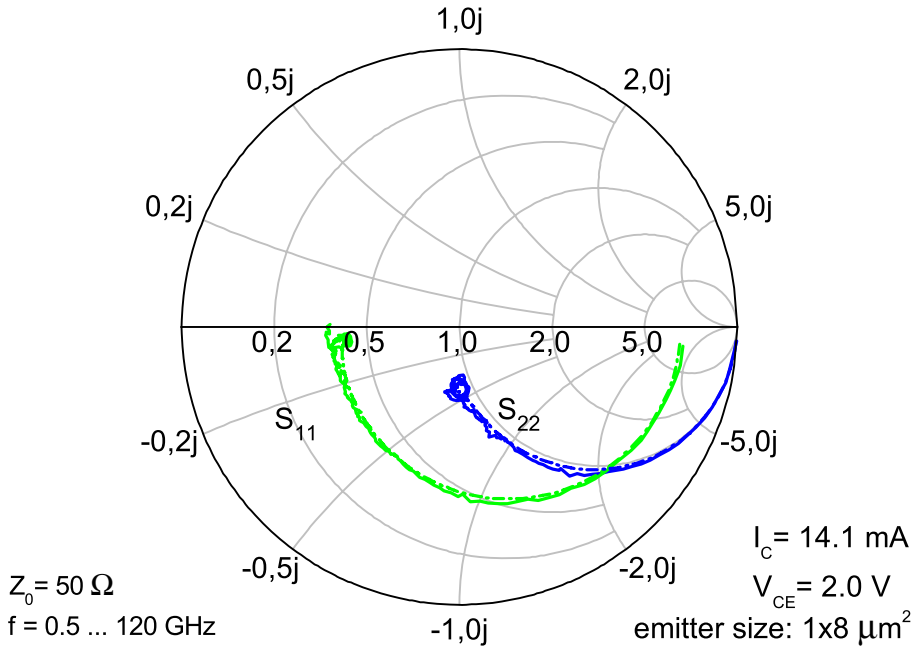


Figure 3.13: Measured (solid lines) and simulated (dashed lines) scattering parameters using the **small-signal model**.

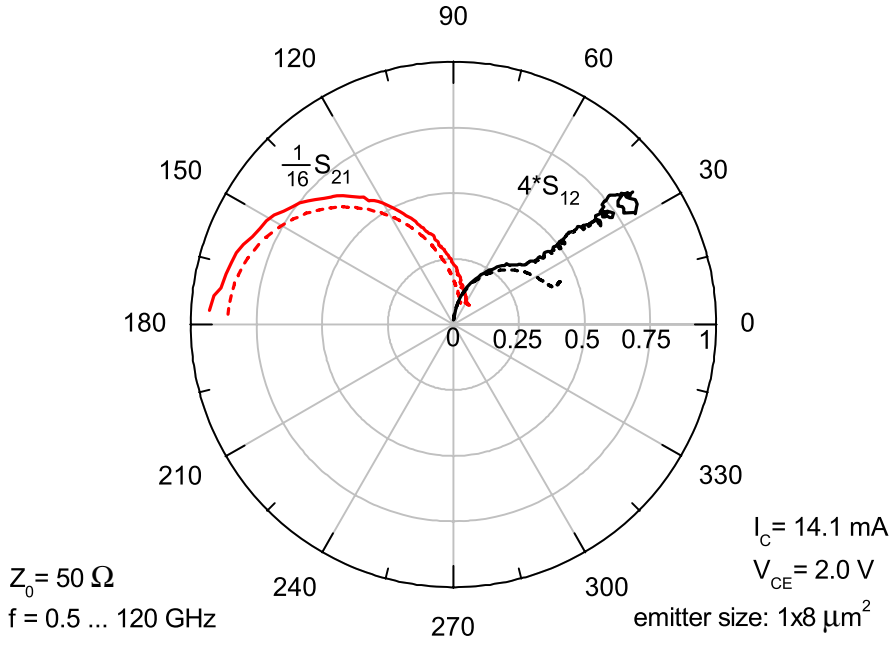


Figure 3.14: Measured (solid lines) and simulated (dashed lines) scattering parameters using the **large-signal model**.

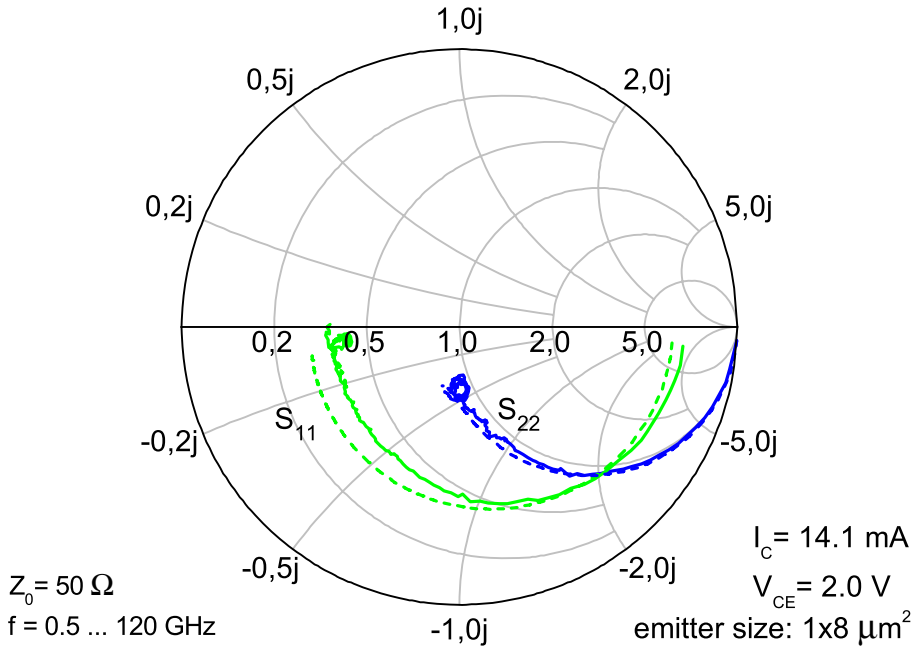


Figure 3.15: Measured (solid lines) and simulated (dashed lines) scattering parameters using the **large-signal model**.

modeled and simulated data is minimized [92], [91]. Using the former method, accurate fits could not be achieved. Therefore, numerical optimization was applied, which resulted in accurate models.

S-parameters of the small-signal model (s. Fig. 2.4) were calculated and numerically fitted to the deembedded measured S-parameter data using the downhill simplex method [82]. A comparison between measured and simulated data is shown in Figs. 3.12 and 3.13. Except for S_{12}, the simulated and measured data match very well. Even in the case of S_{12} the absolute error is very small as the scaling of its data is about 60 times larger in comparison to S_{21} (s. Fig. 3.12).

The fact that the large-signal model is not as accurate is illustrated in Figs. 3.14 and 3.15. These figures show a comparison of simulated and measured S-parameters using the large-signal model. The reduced accuracy in comparison to the small-signal model is apparent. However, the large-signal model is valid for large-signal and small-signal operation in any dc operating point. Therefore, the large-signal model was generally used for simulation of circuits. If a higher accuracy was desired and the circuit could be operated under small-signal condition, an additional small-signal model was extracted at the particular dc operating point and used for fine tuning.

3.3.2 Investigation of Transistors

The topology and the elements of the small-signal model, i.e. the model parameters, are motivated by the geometry of the DHBT and device physics (s. Fig. 2.3). Consequently, properly extracted parameters allow to evaluate single transistors and compare different transistors. In particular, limits with respect to high-frequency performance can be identified, as the high-frequency parameters f_{T} and f_{max} can be calculated from the small-signal parameters (s. Eqs. 2.5 and 2.6) with an error of not more than 10% (s. Fig. 2.32).

However, for the investigation of transistors an accurate model alone is not sufficient. During the extraction it was found that the system of 17 model parameters shown in Fig. 2.4 is under-determined. As a consequence an infinite number of solutions is possible. For example, the parasitic capacitance C_{pbc} can be absorbed in the extrinsic capacitance C_{ex} and vice versa, without changing the accuracy of the model. This prohibits a meaningful physical interpretation of the extracted parameters. Therefore, the number of parameters used in the extraction was gradually reduced. It was found, that by using only the internal device's parameters and setting R_{jc} to infinity an accurate model could be extracted. A similar result has been reported before [16]. A schematic of the resulting small-signal equivalent circuit is shown in Fig. 3.16. This procedure does not imply that there are no parasitic effects, which influence the behavior of the transistor. It merely shows, that the remaining elements of the equivalent circuit can absorb these effects during the fitting process. This fact explains, why the extracted model parameters are only correlated to the physical device parameters within limits. However, the correlation is still strong enough to provide useful interpretations of device behavior.

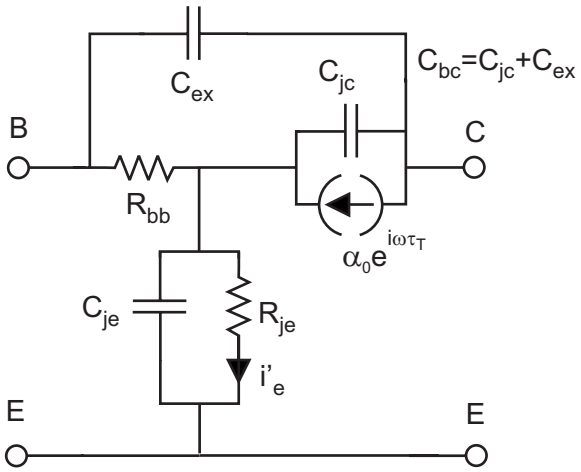


Figure 3.16: Equivalent circuit displaying elements that were needed to extract an accurate small-signal model.

Experimental Results

The transistors intended for circuit design were further investigated. These transistors have an emitter length varying from 2 to 16 μm. The other geometrical parameters are constant (base contact width $W_B = 0.75$ μm, emitter width $W_E = 1$ μm, s. Fig. 2.25). The devices provide a compromise between high speed and high yield on the device level. Their design is the result of an optimization of device geometry (s. Sec. 2.4).

Small-signal models for all devices were extracted at a collector-to-emitter voltage $V_{CE} = 1.75$ and a collector current density of 3.5 mA/μm^2. The model parameter values as a function of emitter length are shown in Figs. 3.17, 3.18, and 3.19. The larger the emitter length the larger the capacitances are and the smaller the resistance values get. While the parameters C_{ex} and C_{je} scale linearly, the intrinsic base-collector capacitance C_{jc} seems to scale quadratically. The resistance parameters R_{je} and R_{bb} scale linearly with $1/L_E$ with the exception of the shortest emitter length (2 μm). The transit time τ_T and α (Fig. 3.19) are approximately independent of emitter length for $L_E > 2$ μm. In the same figure two product terms $C_{je} \cdot R_{je}$ and $C_{jc} \cdot R_{bb}$ are plotted, of which the first one is constant and the second slightly increases with emitter length, if the shortest emitter length is neglected, which is affected by high value for R_{bb}.

The transistor with the shortest emitter length (2 μm) showed strong degradation and reliability problems during the measurement procedures. The origin of this degradation has been correlated to some limits in the present process technology such as device layout and layer structure composition. In effect, the aggressive down scaling of the emitter size likely results in an increase in the emitter cap to metal contact resistance. This causes the corresponding small-signal parameter values to deviate from the scaling behavior of the other devices. To overcome this problem, an incorporation of an InAs contact layer in the emitter cap has been shown to improve the device performance, as reported in [39].

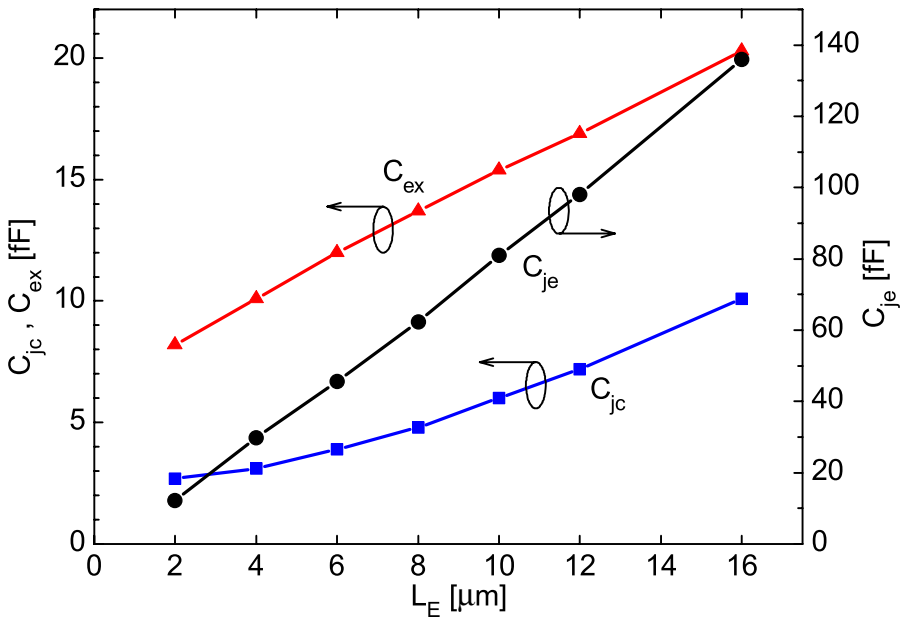


Figure 3.17: Graphical representation of capacitance parameter values versus emitter length. Except for C_{jc} the scaling behavior is linear. The non-linear scaling of C_{jc} is considered to be caused by parasitic effects.

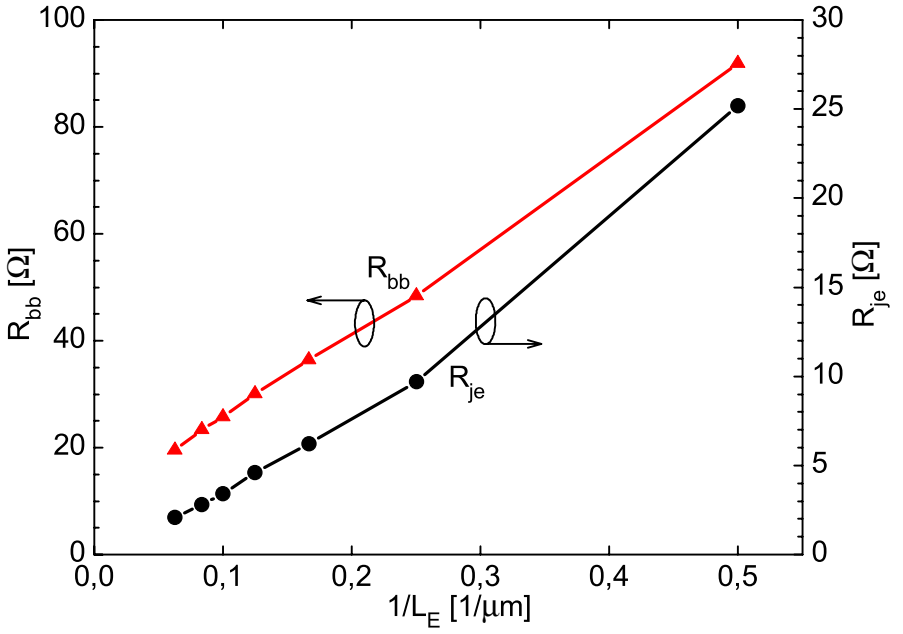


Figure 3.18: Resistance parameter values versus emitter length. Both parameters scale inversely proportional to emitter length. The parameter values for the shortest emitter length seem not to follow the scaling behavior as closely as the other data points. This is attributed to limitations in the present process technology with such small devices (compare Fig. 3.19).

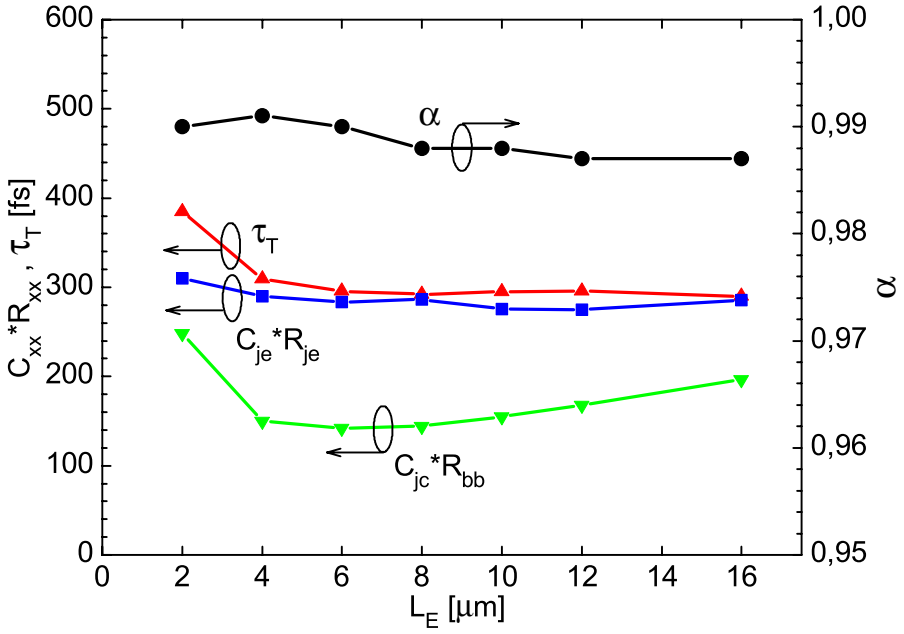


Figure 3.19: Parameter values of τ_T and α as well as the product terms $C_{je} \cdot R_{je}$ and $C_{jc} \cdot R_{bb}$ as a function of emitter length. The increase in τ_T and $C_{je} \cdot R_{je}$ for $L_E = 2$ μm is attributed to scaling limitations in the present process technology resulting in performance degradation.

Discussion

The results reveal some interesting information about the transistors. Especially the dependence of the high-frequency parameters f_T and f_max on emitter length can be understood. In the following discussion the smallest transistor with emitter length $L_E = 2$ μm is neglected as it showed a degradation which is due to the present process technology.

First of all, most parameters seem to scale with emitter size as expected from physical and geometrical considerations. The capacitances increase proportional to the increase in size and the resistances increase with the inverse of emitter length. The transit time τ_T and the base transport factor α are almost constant. This is plausible as these parameters depend on the vertical electron transport mechanisms, which are determined by the epitaxial layer structure and not by device geometry.

However, the intrinsic base-collector capacitance C_{jc} does not scale linearly with emitter length, but rather quadratically (s. Fig. 3.17). This behavior is attributed to parasitic effects. The topology of the extracted small-signal model (Fig. 3.16) does not provide additional parameters to account for parasitics, because they are not needed to extract models with high accuracy. Consequently, the influence of the parasitic effects is absorbed by the other parameters, which manifests in the non-linear dependence of C_{jc} on emitter length.

The cut-off frequency f_{T} and the maximum frequency of oscillation $f_{\max}$ of the discussed devices are shown in Fig. 2.32. The measured values are plotted as a function of emitter length. In addition, the small-signal parameters extracted in this section were used to calculate and plot f_{T} and $f_{\max}$ applying Equations 2.5 and 2.6. Measurement and calculation are correlated closely enough to enable an interpretation of the measurement results using the small-signal model parameters of this section.

The decrease of both high-frequency parameters for short emitter length can be attributed to the collector-base capacitance of the base contact pad (s. Fig. 2.25), which has the same size for all devices. This can be seen in Fig. 3.17, where a linear interpolation of the parameter C_{ex}, which is associated with the extrinsic collector base capacitance (s. Fig. 2.3), appears to result in a considerable value for a device of emitter length zero. The shorter the emitter gets, the greater the influence of the additional parasitic capacitance is. The effect, that a comparatively large extrinsic base-collector capacitance seriously reduces the high-frequency performance of a transistor, was already shown in Sec. 2.4.2. For this reason the base contact pad as well as the base contact width were chosen to be as small as possible in the present process technology. For emitter length exceeding 8 μm the influence of the base contact pad seems to become negligible.

Focusing on long emitter lengths, the scaling behavior of the device neglecting the base contact pad can be seen. The cut-off frequency f_{T} seems to be almost independent of the emitter length. The reason is, that the delay time τ_T and the product of R_{je} and C_{je} are approximately constant for any emitter length (s. Fig. 3.19) and dominate the result of Equation 2.5 for long emitter length, where C_{je} is much larger than the other capacitance parameters (s. Fig. 3.17).

An explanation for the scaling behavior of $f_{\max}$ at long emitter length is offered by equation 2.6 and the extracted small-signal model parameters. Any variation of the ratio between f_{T} and $f_{\max}$ depends on the product of R_{bb} and C_{jc}, whose value is plotted in Fig. 3.19 as a function of emitter length. For emitter lengths larger than 6 μm the values gradually increase explaining the decrease of $f_{\max}$. The increase of this product term is probably due to parasitic effects, which are absorbed by the model parameters R_{bb} and C_{jc}. Especially C_{jc} does not scale linearly, which would be unexpected, if the parameter did only depend on the intrinsic collector capacitance. Possible reasons for this observation are resistive and inductive effects in the narrow base contacts along the emitter (s. Fig. 2.25). This is conceivable, if the current flow is considered. The emitter and collector wiring is arranged in such a way that the current densities for all emitter lengths are comparable. However, the base current flows along the emitter through the narrow base contact, which does not scale in width with emitter length. As a consequence, the finite resistance of the contact could become significant. In addition, the skin effect in the narrow contact could play an important role, which was neglected in the derivation of the small-signal model topology (s. Figs. 2.3 and 2.4).

3.4 Summary

In this chapter the large-signal UDSD model as well as a small-signal model were described and their contributions to the design of high bit rate circuits and the investigation of transistors were explained. Starting from a preliminary status the extraction procedure for the large-signal model could be refined and adjusted to the demands of designing digital and analog circuits targeting 80 Gbit/s. In this process a suitable compromise between complexity and accuracy was made. These models successfully contributed to the realization of state-of-the-art analog and digital circuits [94], [74], [72].

The small-signal model is a very useful tool in addition to the large-signal model. Even though it is only valid in small-signal operation, its accuracy is higher and the extraction is faster than in the large-signal case. It can be used for fine tuning of the small-signal responses of amplifiers. In addition, it can be used to investigate transistors. Extracted model parameters were used to explain the complicated dependence of the high-frequency parameters f_{T} and f_{max} on emitter length.

Chapter 4

Methods of Broadband Amplifier Characterization

4.1 Figures of Merit

Most optical communication systems employ binary amplitude modulation using the "non-return to zero" (NRZ) data format [86]. The transmitted data signals experience various imperfections in the optical and electrical domain, which can increase the bit error rate (BER) or even make a transmission of data impossible. To evaluate the performance of system components, different figures of merit (FOM) can be defined, in addition to the BER. For broad-band amplifiers suitable as modulator drivers bandwidth, group delay and output power are important. These FOMs support the interpretation of the experimental results in Chapters 5 and 6.

4.1.1 Bandwidth

In digital communications systems, the available bandwidth is limited, because many system components, e.g. amplifiers, have limited bandwidth. For this reason, it is important to consider the spectral content of the digitally modulated signals in order to design suitable amplifiers. Calculations show, that the spectral shape of a random binary sequence in NRZ data format is determined by the spectral characteristic of a single signal pulse [83].

A comparison of two different signal pulses and their power density spectra is shown in Fig. 4.1. In the time domain, the rectangular pulse $u_{\sqcap}(t)$ (dashed line) of length T has the ideal shape for data transmission at a bit rate of 1/T with a rise and fall time of zero. The raised cosine pulse $u_{\cos}(t)$ (solid line) has twice the length in the time domain and a rise an fall time of length T.

In the frequency domain, the spectral power density of the raised cosine pulse $|U_{\cos}(f)|^2$ (solid line) is almost entirely confined to frequencies below 1/T. In contrast, the rectangular pulse has noticeable frequency components ($|U_{\sqcap}(f)|^2$, dashed line) at frequencies greater than 1/T. An explanation for this behavior is given by the equations for both pulses in the

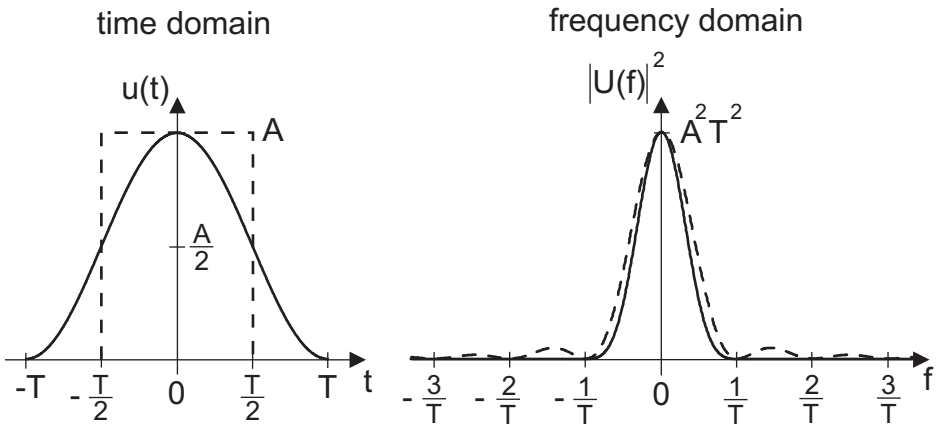


Figure 4.1: Time and frequency representation of raised cosine pulse (solid lines) in comparison to the rectangular pulse (dashed lines). Analytic formulas are given in Eqs. 4.1 to 4.4.

time domain and their Fourier transforms (s. Eqs. 4.1 to 4.4). In contrast to the rectangular pulse, the Fourier transform of the raised cosine pulse decays inversely as f^3.

$$u_{\sqcap}(t) = A\,;\ |t| < \frac{T}{2} \tag{4.1}$$

$$U_{\sqcap}(f) = AT\frac{\sin \pi fT}{\pi fT} \tag{4.2}$$

$$u_{\cos}(t) = \frac{A}{2}\left[1 + \cos\frac{\pi t}{T}\right]\ ;\ |t| < T \tag{4.3}$$

$$U_{\cos}(f) = AT\frac{\sin 2\pi fT}{2\pi fT\,(1 - 4f^2T^2)} \tag{4.4}$$

Electrical digital circuits always generate signal pulses with finite rise and fall times. These delay times depend not only on the high-speed performance of the transistor, but also on the circuit concept. E.g., the current switch, a fundamental building block of digital circuits, utilizes not only transistors, but also resistors. These resistors, in combination with the intrinsic and parasitic capacitances of the transistor, give rise to delay time constants, which affect the shape of the signal pulses. As a result, high-speed NRZ data signals typically have a shape which is close to the raised cosine pulse, which can be seen in the eye diagram in Fig. 4.4 (top). The eye diagram shows the superposition of a pseudo random bit sequence, which is generated by a 80 Gbit/s electrical multiplexer (s. Sec. 4.3.2).

Although the shape of the signal pulses is not exactly a raised cosine pulse, the similarity can be used to interpret the eye diagrams in Fig. 4.4 in the following way. First, the fact that the pulses have a raised cosine like shape (top figure) implies, that the bandwidth of the data signal is probably not greater than 80 GHz (s. Fig. 4.1) and that this bandwidth is sufficient to transmit digital signals in NRZ data format at a data rate of 80 Gbit/s. Second, from the middle and bottom figure we conclude, that a further reduction in bandwidth

results in even broader and not as clearly modulated signal pulses, which increases the BER and eventually makes data transmission impossible. This is discussed in more detail in Sec. 4.3.2.

The degradation of the ideal rectangular pulse of length T to a pulse with finite rise and fall times with a pulse length greater than T is also called intersymbol interference (ISI) [86], as each pulse interferes with the preceding and succeeding pulses. The discussion above shows that ISI can be cause by a limitation in bandwidth, but that a bandwidth of 1/T is still sufficient for data transmission. However, ISI is also influenced by characteristics other than bandwidth, i.e., group delay described in the following section.

4.1.2 Group Delay

Dispersion is a familiar phenomenon in electronics. If different frequencies comprising a pulse propagate at different velocities in a medium or system component, then phase distortion occurs and the pulse shape can change. Group delay is a useful measure of phase distortion in components such as amplifiers and filters. It measures the transit time, with respect to frequency, of a signal through the device under test. The group delay T_g is calculated as the derivative of the phase response with respect to frequency [3]:

$$T_g = \frac{d\Phi}{d\omega} \tag{4.5}$$

The pulse shape does not change, if the group delay is constant over the full width of its spectrum. In broadband amplifier design, this is not always the case, because special design techniques like peaking mechanisms (s. Sec. 6.1.2) can cause resonance effects within the amplifier, which can seriously affect the group delay flatness. Therefore, in amplifier design group delay is an important design criterion. Evaluating the group delay has been used before to assess the quality of broadband amplifiers [95] and can be used to estimate the possible strength of ISI. However, eye diagrams are the preferred method of characterization (s. Sec. 4.3.2), because they show the overall effect of ISI on the quality of the data signal.

4.1.3 Output Power

The output power is an important FOM for amplifiers used as modulator drivers. The modulator at the transmitter side of an optical communication system (s. Fig. 1.2) can be realized using different materials and concepts [24]. The most commonly used materials for external modulators are $LiNbO_3$, III-V compound semiconductors, and electric polymers. They are used to realize Mach-Zehnder (MZ) and electroabsorption (EA) modulators. Common for all modulator concepts is the fact, that an acceptable extinction ratio of the optical modulated signal can only be achieved, if the voltage swing of the electrical signal driving the modulator matches or exceeds a certain value, which depends on the materials and concepts. Recently, a traveling-wave EA modulator suitable for 80 Gbit/s was reported, which required a voltage swing of 3 V peak-to-peak delivered to a 50 Ω

Figure 4.2: Photograph of S-parameter measurement system HP 8510XF.

load [111]. Therefore, modulator drivers should deliver a voltage swing of at least 3 V in a 50 Ω circuit environment, which corresponds to an output power of 13.5 dBm.

4.2 Small-Signal-Characterization (S-Parameter)

In order to characterize the behavior of a two-port network, measured data of both its transfer and impedance functions must be obtained. At low frequencies, the impedance (Z), admittance (Y), or hybrid (H) parameters are examples of network functions used in the description of two-port networks. However, these parameters cannot be measured accurately at frequencies in the GHz-range because the required short- and open-circuit tests are difficult to achieve over a broadband range of microwave frequencies [36].

A set of parameters that is very useful in the microwave range are the scattering parameters (S-parameters). These parameters are defined in terms of traveling waves and present the reflection and transmission coefficients between the incident and reflection waves. They describe completely the behavior of a device under linear conditions at microwave frequencies. Each parameter is typically characterized by magnitude and phase. In the case of two-port networks embedded in an environment with a characteristic impedance of $Z_0 = 50\ \Omega$ the following parameters are defined:

S_{11}: input reflection coefficient of 50 Ω terminated output,

S_{21}: forward transmission coefficient of 50 Ω terminated output,

S_{12}: reverse transmission coefficient of 50 Ω terminated input,

S_{22}: output reflection coefficient of 50 Ω terminated input.

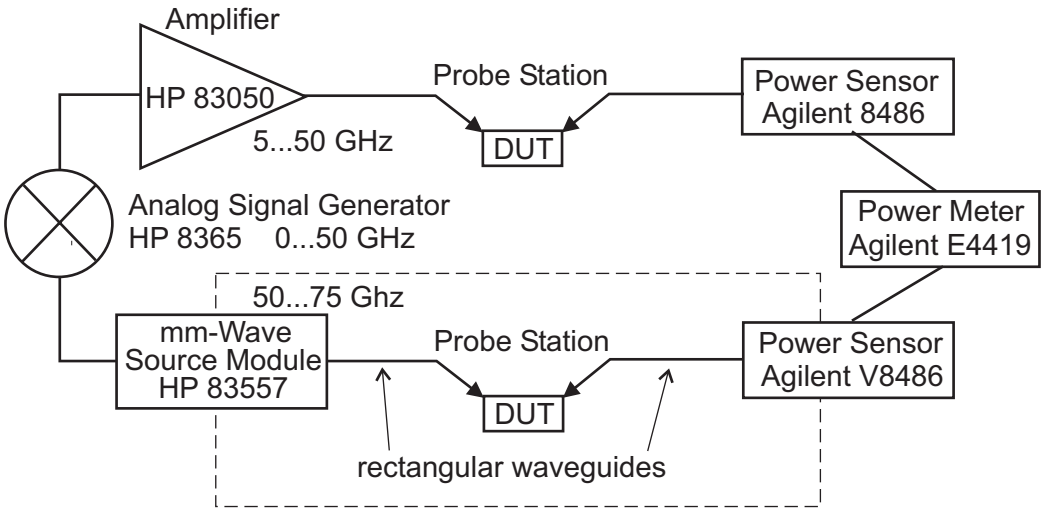


Figure 4.3: Schematic of power measurement setup. The device under test (DUT) was measured using coaxial cables up to a frequency of 50 GHz (top). For the frequency range from 50 to 75 GHz (bottom) rectangular rigid waveguides were used (inside dashed box).

The advantage of S-parameters does not only lie in the complete description of the device performance at microwave frequencies but also in the ability to convert to other parameters such as Z, Y and H-parameters.

In this work the measurement system HP 8510XF by Agilent was used. It allows to measure S-parameters in the frequency range from 250 MHz to 120 GHz. A photograph of the system is shown in Fig. 4.2.

4.3 Large-Signal-Characterization

4.3.1 Output Power

A schematic of the power measurement setup is shown in Fig. 4.3. The signal generator provides sinusoidal electrical signals at frequencies up to 50 GHz. For measurements up to this frequency the signal is amplified by a broadband amplifier and the power is delivered to the device under test (DUT) at the probe station. The output of the DUT is detected by a power sensor, whose signal is processed and displayed by a power meter. In order to measure frequencies above 50 GHz a mm-wave source module is used to generate frequencies in the range from 50 to 75 GHz. The source module contains a frequency multiplier and an amplifier. The coaxial cables used at lower frequencies exhibit large losses above 50 GHz. Therefore, rectangular waveguides specified for frequencies from 50 to 75 GHz were used in parts of the system, which operate above 50 GHz marked by a dashed box.

The amplifiers were measured according to the following procedure. First, the measurement system was calibrated. Second, the losses between the probe station and the power sensor were measured. Based on this data the power levels at the output of the DUT could

be calculated from the measured power levels. During each measurement the power level of the signal generator was swept over a wide range and the output power of the DUT was measured.

Such measurements were conducted to evaluate the output power capabilities of the amplifiers presented in Chapter 6. A typical measurement result showing the output power P_{OUT} as a function of input available power P_{IN} is shown in Fig. 6.8. These measurements were conducted at several frequencies between 5 and 75 GHz. An example of a graph summarizing the results of such frequency dependent power measurements is given in Fig. 6.7. The output power at 1-dB compression as well at the gain are plotted as a function of frequency. The discontinuity at 50 GHz results from the fact that this frequency was measured with both measurement setups shown in Fig. 4.3. The accuracy of the power sensors is 0.2 dB, which results in slightly different results between two measurements.

4.3.2 Signal Distortion (Eye-Diagram)

Although the characterization methods described above can give valuable information about the characteristics of amplifiers (bandwidth and output power), they cannot give conclusive results with respect to digital signal transmission. To investigate the quality of transmission of arbitrary bit patterns through a system component, long sequences of random bit patterns must be examined. A common tool for visualizing the nonidealities in random data is the eye diagram. Such a diagram is generated by dividing a long pseudorandom bit sequence [71] into short sequences of equal length and superimposing the resulting short signal traces in a single diagram. Then the eye diagram displays the accumulation of distorted edges and levels, which have to be expected during the transmission of random data. Eye diagrams revealing different signal qualities are shown in Fig. 4.4. The larger and clearer the opening of each eye is, the lower the bit error rate can be expected to be.

Electrical eye diagram measurements were performed using a $2^{31}-1$ pseudorandom bit sequence (PRBS) pattern generated by a four-channel 10-Gbit/s signal pattern generator, and three levels of 2:1 multiplexers (MUX) manufactured in-house. The 80 Gbit/s pattern is generated by a full 2:1 multiplexer including data-flip-flops realized in an in-house metamorphic HEMT technology [93]. The multiplexer has single-ended in- and output and implements source coupled FET logic (SCFL). The output eye-diagram was measured using an Agilent 86100B oscilloscope with remote sampling modules 86118A specified up to 70 GHz. The eye diagram of an on-wafer measurement of the multiplexer at 80 Gbit/s is shown in Fig. 4.4. After packaging and extending the module by a 3-dB attenuator (HP 8490D), the eye diagram was degraded (Fig. 4.4). For on-wafer measurements of the amplifiers presented in Chapters 5 and 6 the measurement setup was expanded by two probes (Cascade ACP 65) and two bias-Ts (SHF BT 65B) for biasing. A reference on-wafer measurement of a 200 μm long 50 Ω CPW resulted in a further degenerated eye diagram (Fig. 4.4).

The strong degeneration of the eye is attributed to the packaging, and to the attenuator, bias-Ts, and probes, which are specified only up to a frequency of 50 to 65 GHz. A major factor is the reduction of available bandwidth below 80 GHz, which increases the

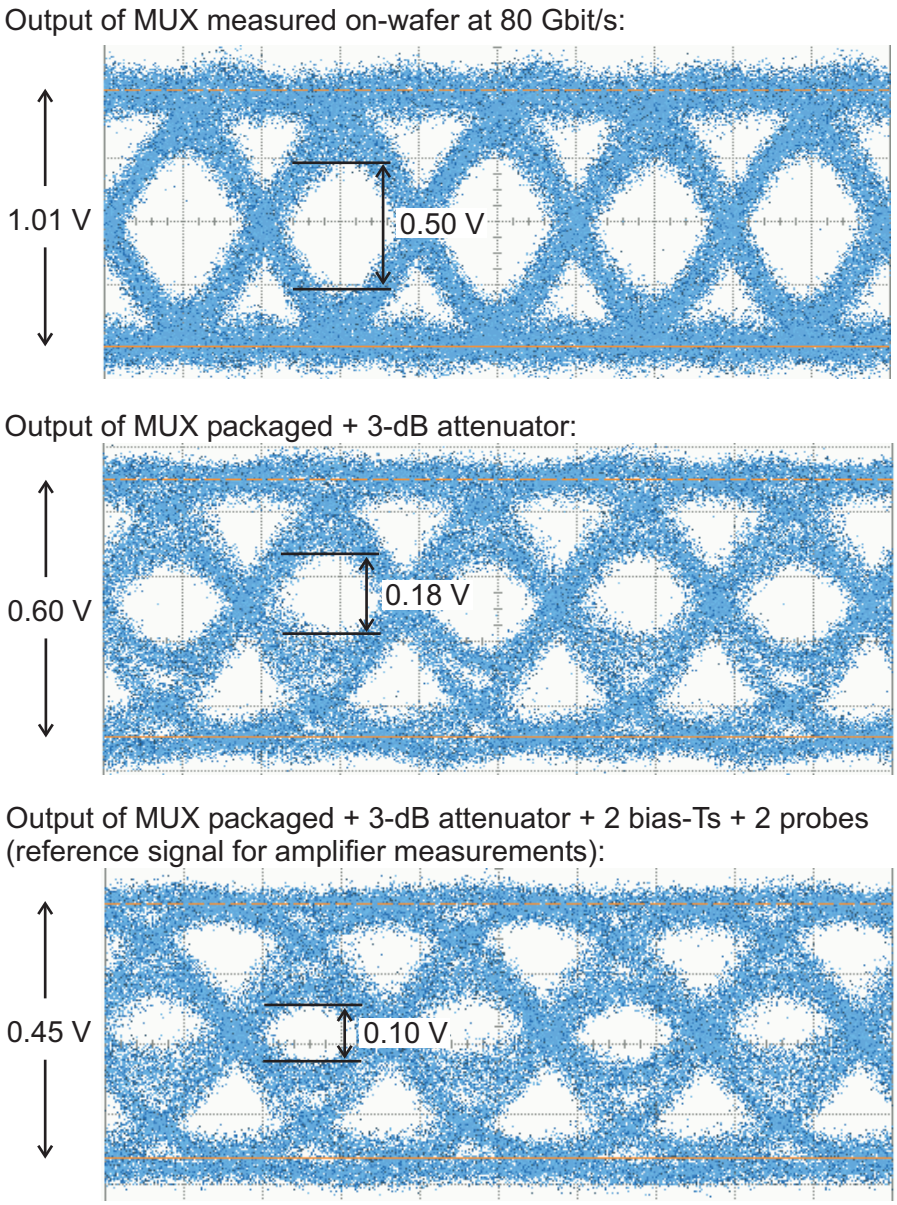


Figure 4.4: Eye diagrams of MUX output signal at 80 Gbit/s. The on-wafer measurement showed a clear eye (top). After packaging and adding an attenuator the eye was degraded (middle). After extending the measurement setup with two bias-Ts and two probes, an on-wafer reference measurement of a short through line resulted in a further degraded eye (bottom). The relative eye opening fell from 50 % (top) to 22 % (bottom).

intersymbol interference and results in a degradation of signal quality (s. Sec. 4.1.1). Still this eye diagram was used as a reference for eye diagram measurements at 80 Gbit/s. The reference eye has a voltage swing of 0.45 V_{pp} (peak-to-peak) and an eye opening of 0.10 V_{pp}. A measure for the quality of an eye diagram is the relative eye opening, i.e. the ratio of eye opening to amplitude, which is 22 % in the case of the reference signal.

Chapter 5

Realization of Compact Lumped Amplifiers

For amplifiers two major design principles exist: the lumped amplifier approach and the distributed amplifier approach. Both designs have advantages and limitations. Among the advantages of lumped amplifiers are features like compact design and low power consumption. Disadvantages are the comparatively low output power and smaller bandwidth.

This chapter investigates the performance that can be achieved with lumped amplifiers in InP HBT technology. The first section, describes major concepts that can be used in lumped amplifier design. Following this an experimental comparison of several amplifier concepts is presented. The best concept is further investigated with regards to digital signal transmission at 40 and 80 Gbit/s. Finally, the output power capability of these lumped amplifiers is investigated.

5.1 Design Considerations

Single-Stage Amplifiers

Important characteristics of broadband amplifiers are gain and bandwidth, which depend on the transistors' gain-bandwidth product. Feedback is a common technique used in wideband amplification to control the gain and bandwidth of an amplifier [44]. A schematic of a simple amplifier using resistive feedback is given in Fig. 5.1(a). Local series feedback in the form of emitter degeneration is commonly used in wideband amplifiers that incorporate bipolar transistors. Series feedback does not change the current gain of the device, but it increases the input impedance. The use of local shunt feedback is also a common technique in the design of wideband amplifiers. Even when local shunt feedback is not intentionally introduced, the parasitic collector-base junction capacitance in a bipolar transistor gives unwanted shunt feedback. Ideal resistive feedback does not change the gain-bandwidth product of a single-pole amplifier, it rather trades gain for bandwidth. However, device parasitics reduce the gain-bandwidth performance.

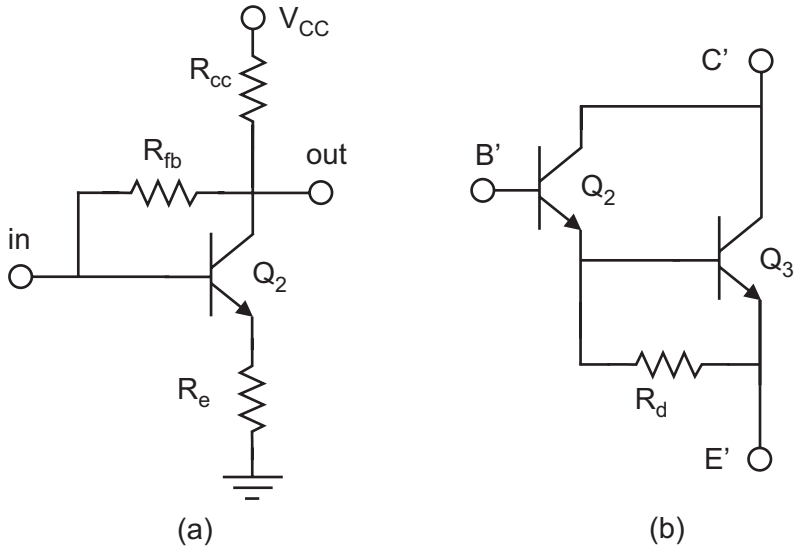


Figure 5.1: (a): Schematic of single-stage amplifier employing shunt feedback (R_{fb}) and series feedback in the form of emitter degeneration (R_e). (b): Schematic of Darlington compound transistor cell with effective base (B'), collector (C'), and emitter (E') nodes.

In order to prevent standing waves in transmission lines connecting amplifiers that are separated by many wavelengths, it is necessary to design amplifiers with terminal impedances matched to the specified system impedance (usually 50 Ω). Even within circuits there are advantages in keeping impedance levels matched, such as reduced sensitivity to parasitics. Using impedances that are too high makes a circuit sensitive to capacitive parasitics, and using impedances that are too low, makes a circuit sensitive to series resistive and inductive parasitics. A good approach to impedance matching is the use of shunt feedback and series feedback described above. For a matched-impedance amplifier voltage gain and current gain are identical. In this work all amplifiers are matched to 50 Ω.

Dual-Stage Amplifiers

While single-stage matched-impedance amplifiers are very simple, they often do not provide adequate gain. Two-stage amplifiers offer superior gain for a given bandwidth with flatter band-pass characteristic. Because a two-pole amplifier is generally made up of two gain stages, a reduction in bandwidth by a given factor increases the gain by the square of that factor. The reason for this is that each of the two stages increases in gain when the bandwidth is reduced. Therefore, the $\sqrt{\text{gain}}\times$bandwidth product is the proper performance measure for a two-stage amplifier [44].

Improved performance can be obtained by applying feedback around more than only one stage. Feedback around a two-pole amplifier causes substantially different effects from feedback around a single-pole amplifier. Typically, two-stage amplifiers have two significant poles, which can lead to excessive gain peaking in the amplifier response. If an appropriate feedback mechanism around both stages is used, the excessive peaking vanishes.

Compound Transistor Stages

Often compound transistor configurations in place of single bipolar transistors are used in wide-band amplifier design. The primary motive is to achieve a higher current gain at a given frequency. Particularly common is the combination of two transitors in a common-collector and a common-emitter configuration, which is shown in Fig. 5.1(b). This configuration is referred to as a Darlington configuration. From the schematic of the Darlington cell, it can be shown that the effective small-signal current gain of the Darlington cell can be two times larger than in the case of a single device [61]. This occurs when the resistance is adjusted such that the emitter current in both devices is equal. Consequently, the Darlington cell significantly improves the gain-bandwidth product of the common emitter amplifier topology. The cost of this improvement is added circuit complexity and increased phase shift in current gain.

Amplifier Configurations

Two common configurations in the design of two-stage amplifiers are, first, the so called Meyer configuration [78] and, second, the Kukielka configuration [54]. In literature the advantages and disadvantages of using single transistor stages, compound transistor stages, the Meyer configuration, and the Kukielka configuration are discussed [44], [21], [43]. However, it has not been shown, that one of the circuit concepts always yields superior performance in any technology. For each technology the amount of parasitics, the high-frequency parameters, and their interrelation differ. These factors influence the performance of each amplifier concept. To clarify which lumped amplifier approach achieves best performance in the InP-DHBT technology of this work, four different amplifiers were designed and experimentally tested. The results of this comparison are presented in Section 5.2.1.

5.2 Realized Amplifiers

5.2.1 Comparison of Lumped Amplifier Approaches

Four amplifier concepts were investigated with regards to their gain-bandwidth product: two Meyer configurations and two Kukielka configurations. All amplifiers use devices of a nominal emitter size of $4 \times 1\,\mu\text{m}^2$. In the design the small-signal model (Sect. 3.3) was used and all amplifiers were designed for equal bandwidth.

Meyer Configurations

A Meyer configuration with a Darlington compound transistor is shown in Fig. 5.2. The circuit can be seen as a hybrid between the voltage-feedback pair (via R_{fb1}) and a current-feedback pair (via R_{fb2}). The configuration can also be considered a voltage-gain stage with shunt feedback around it (via R_{fb1}) for matching. A detailed calculation of the circuit

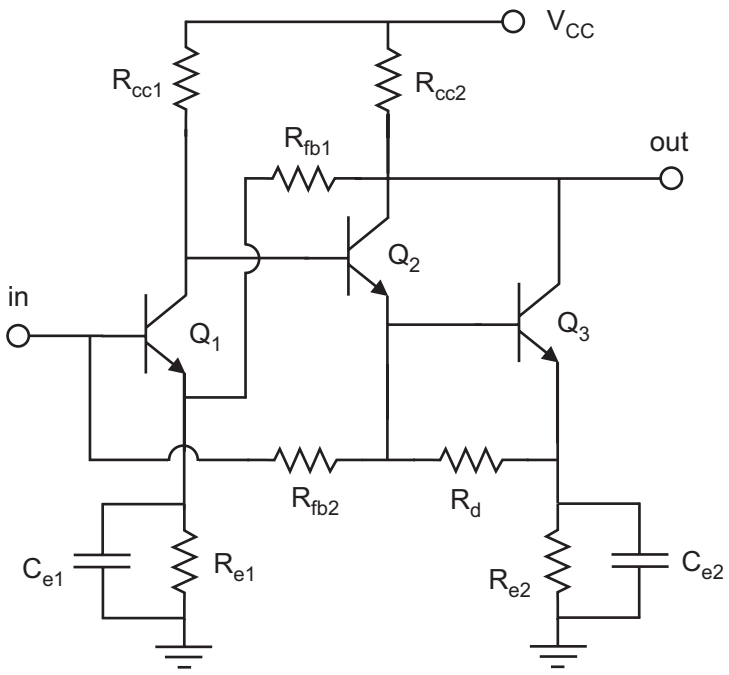


Figure 5.2: Schematic of lumped amplifier in Meyer configuration with a Darlington compound transistor stage. The original configuration used a single device instead of the Darlington configuration (Q_2, Q_3, R_d).

concept has been given before [44]. Capacitors were added in parallel to R_{e1} and R_{e2} to introduce peaking and enhance the bandwidth characteristic [22].

Two variations of this circuit were realized. First, a Meyer configuration with Darlington described above and, second, the original Meyer configuration [78], that uses a single transistor as a second stage.

There is usually little advantage in using a compound first stage, as the impedance at the input node at high frequencies is fairly low. Additionally, making the first stage compound reduces the available output swing. Finally, compound stages are usually noisier than single-transistor stages. For these reasons, compound stages are usually not used in the input stage of a multi-stage amplifier.

Kukielka Configurations

The Kukielka configuration is shown in Fig. 5.3 and can be considered a voltage gain stage with shunt feedback around it and is a variant of the current feedback pair. Series-series feedback in the form of emitter degeneration at Q_1 gives a high input impedance (in the absence of R_{fb1}). Local shunt feedback around Q_2 gives low impedance node at the base of Q_2. Then shunt feedback is applied around this voltage amplifier via R_{fb1}. Capacitors were added in parallel to the emitter degeneration resistance like in the Meyer configuration. A detailed theoretical analysis of this amplifier type is given elsewhere [21].

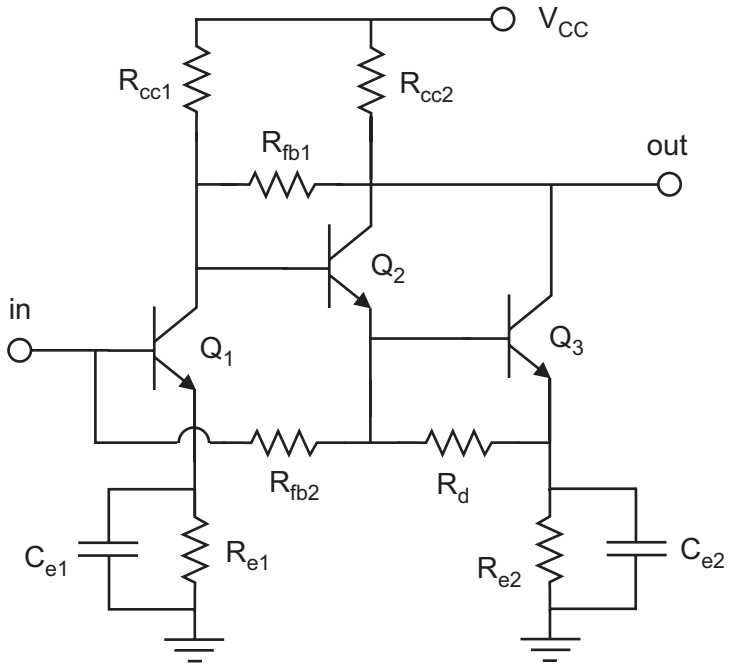


Figure 5.3: Schematic of lumped amplifier in Kukielka configuration and with a Darlington compound transistor stage. The original configuration used a single device instead of the Darlington configuration (Q_2, Q_3, R_d).

As for the Meyer configuration this circuit was investigated with and without a Darlington compound transistor.

Results

The measured small-signal gains for all four amplifier versions are shown in Fig. 5.4. As both the gain as well as the 3-dB bandwidth (BW) varies between amplifier versions, the characteristics of the amplifiers are summarized in Table 5.1. As the amplifiers are dual-stage amplifiers the critical figure of merit is the $\sqrt{\text{gain}}\times$bandwidth product (s. Sec. 5.1). The comparison reveals, that the Kukielka configurations yield a better high-frequency performance, while the Kukielka configuration with Darlington compound stage shows the highest value of this figure of merit.

amplifier configuration	3-dB BW [GHz]	gain [dB]	$\sqrt{\text{gain}} \times$ BW [GHz]
Meyer	63	8	100
Kukielka	84	7	125
Meyer with Darlington	71	9	119
Kukielka with Darlington	68	11	128

Table 5.1: Table comparing the characteristics of four different lumped amplifiers. The critical figure of merit is the $\sqrt{\text{gain}}\times$bandwidth product.

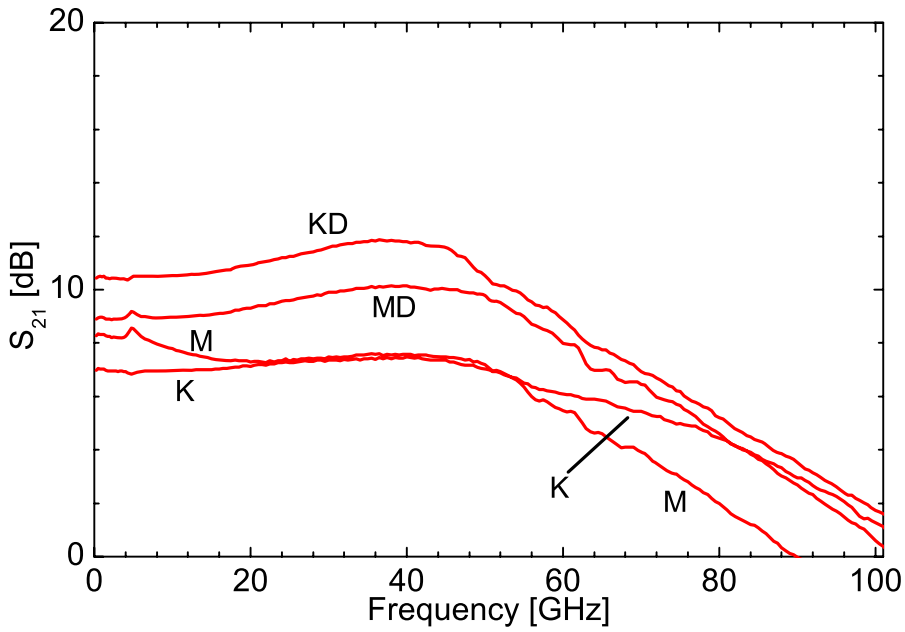


Figure 5.4: Comparison of the small-signal gain of the four different amplifiers using Meyer configuration (M), Kukielka configuration (K), Meyer configuration with Darlington (MD), and Kukielka configuration with Darlington (KD).

Discussion

The comparison of the amplifier results reveals that the Kukielka configuration with Darlington compound transistor seems to have superior gain-bandwidth performance in the present InP technology. This confirms the usefulness of compound transistor stages. The superior performance of the Kukielka configuration in comparison to the Meyer configuration might be due to the fact, that the latter concept is more sensitive to parasitic effects [44]. The amplifiers with Darlington stage show the desired small-signal gain with a moderate peaking of 1-dB before the roll-off. The curve progressions of the small-signal gains of the simple Kukielka and Meyer configurations have a different shape. The differences cannot only be attributed to a change in amplifier concept. Deficiencies in circuit layout and design are likely to have contributed to this result. Nevertheless, the Kukielka configuration with Darlington compound transistor was further investigated and the results are given in the following sections.

5.2.2 High-Gain Amplifier for 40 Gbit/s

Design

On the basis of the result of the previous section a high-gain amplifier according to Fig. 5.3 was designed. The amplifier employs transistors with a nominal emitter size of $4 \times 1\,\mu\text{m}^2$. The simulations were conducted using a small-signal model (Sect. 3.3). The amplifier was

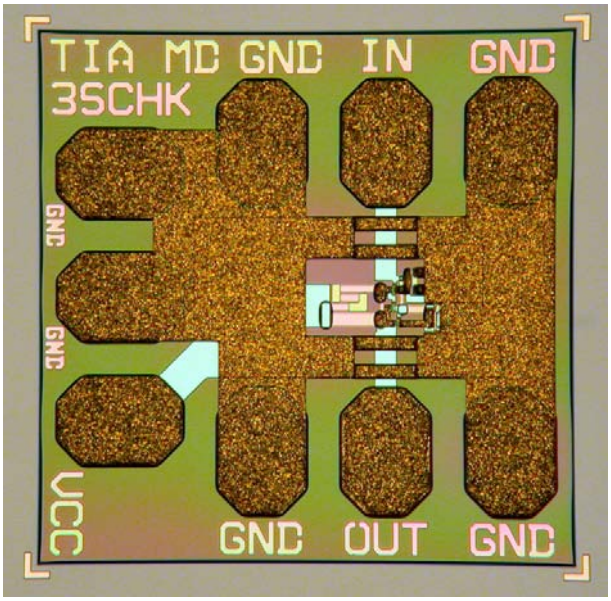


Figure 5.5: Chip photograph of high-gain lumped amplifier resembling the layout of all lumped amplifiers. The design differences between the amplifiers of this chapter are not visible in a photograph. The chip size is $500 \times 500\,\mu\text{m}^2$.

designed for operation at 40 Gbit/s data rate. As the bandwidth requirement is reduced in comparison to the amplifier results above much higher gain can be achieved. A chip photograph of the amplifier is given in Fig. 5.5.

Experimental Results

The amplifier was operated at a supply voltage of 6 V (V_{cc}), resulting in an input current of 29 mA. This corresponds to a current density of 2.4×10^5 A/cm^2 in each transistor. The power consumption was 170 mW.

The results of the S-parameter measurement are shown in Fig. 5.6 in solid lines. The gain of the amplifier is 19 dB. It exhibits a peaking of 1 dB and a 3-dB bandwidth of 43 GHz. The roll-off is very gradual. The simulated data are shown in dashed lines.

The eye diagram was measured with a pseudo random bit pattern at 40 Gbit/s (s. Sect. 4.3.2). The result is shown in Fig. 5.7. The signal gain is 19 dB. The noisy pattern of the reference signal is caused by the oscilloscope. In order to save the detector from damages a 20-dB attenuator was inserted at the input of the oscilloscope. The attenuator causes noise, which leads to the blurry reference eye pattern. As this noise is due to the detection unit, it is not amplified and a clear large output eye is measured. The larger voltage swing of the output eye pattern improves the signal-to-noise ratio at the detector and makes the signal look less noisy. The peak-to-peak voltage swing of 590 mV corresponds to about -1 dBm or 0.8 mW output power.

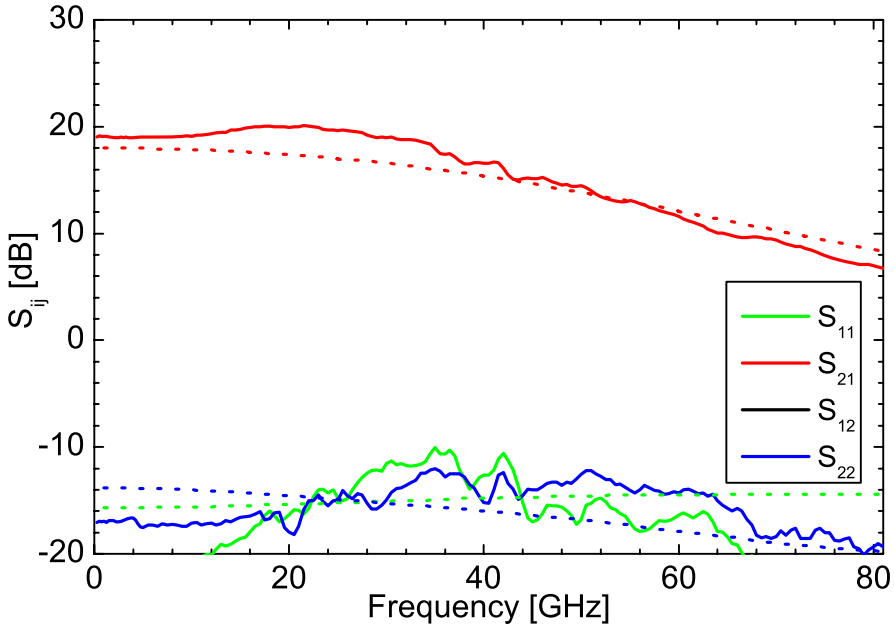


Figure 5.6: Measured (solid lines) and simulated (dashed lines) scattering parameters of lumped amplifier for 40 Gbit/s data rate. The gain is 19 dB and the 3-dB bandwidth is 43 GHz.

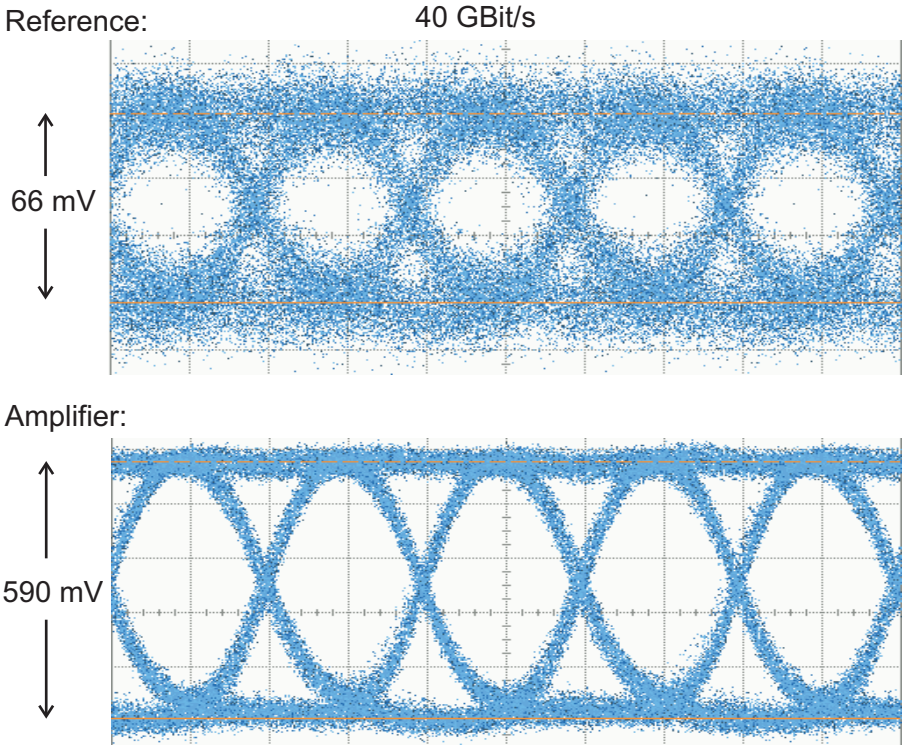


Figure 5.7: Eye diagram of high-gain lumped amplfifier at 40 GBit/s. The signal gain is 19 dB.

Discussion

The amplifier exhibits excellent performance with regard to the transmission of 40 Gbit/s digital data signals. The bandwidth exceeds 40 GHz, the roll-off is very gradual and the matching is better than −10 dB. The eye diagram shows an almost ideal output signal.

This result confirms, that this amplifier concept and design is very well suited for digital applications. Such amplifers can be used as transimpedance amplifiers (TIA) in combination with a photodiode in an optical receiver. However, as will be shown later in this chapter, theses amplifiers are not suitable as modulator drivers.

The design goal of the circuits in this work is the application in 80 Gbit/s data transmission systems. Therefore, the presented amplifier has not sufficient bandwidth. In the following section another amplifier is presented, which has much wider bandwidth and reduced gain.

5.2.3 Low-Power Amplifier for 80 Gbit/s

Design and Layout

While the previous amplifier was designed for high gain, this amplifier is design for maximum bandwidth. The concepts of the amplifiers are identical and devices with an emitter size of $4 \times 1\,\mu\text{m}^2$ were used. By adjusting the design values of the circuit components (s. Fig. 5.3), the gain and bandwidth of the amplifier can be adjusted. The emitter degeneration resistances (R_{e1}, R_{e2}) were increased to reduce the gain and increase the bandwidth. To move the frequency at which capacitive peaking occurs to higher values, the value of the capacitors C_{e1} and C_{e2} were reduced. The feedback resistors were adjusted to account for the change in bias conditions within the circuit. According to simulations these measures yield an amplifier with a small-signal gain of 10.5 dB and 80 GHz 3-dB bandwidth (Fig. 5.8). In comparison to the first amplifier the design changes are not visible in a chip photograph and this amplifier's layout looks identical to the first one (Fig. 5.5).

Experimental Results

The amplifier was measured with a power supply voltage $V_{cc} = 5$ V. The circuit drew a current of 24 mA corresponding to power consumption of 120 mW. The current density in each transistor was 2.0×10^5 A/cm^2.

A comparison of the measured and simulated scattering parameters is given in Fig. 5.8. The solid lines mark the measurement data. The small-signal gain of the amplifier is 10.5 dB with a peaking of 1.5 dB and a 3-dB bandwidth of 70 GHz. The matching of input and output port is better than −10 dB over the full bandwidth. The simulated data are displayed in dashed lines. The bandwidth was predicted to be larger than measured.

As the bandwidth of this amplifier is 70 GHz the eye diagram measurement was conducted with an 80 Gbit/s signal pattern. Because of the lower gain of the amplifier the signal was not attenuated as much as in the case of the first amplifier. The result of the measurement is shown in Fig. 5.9.

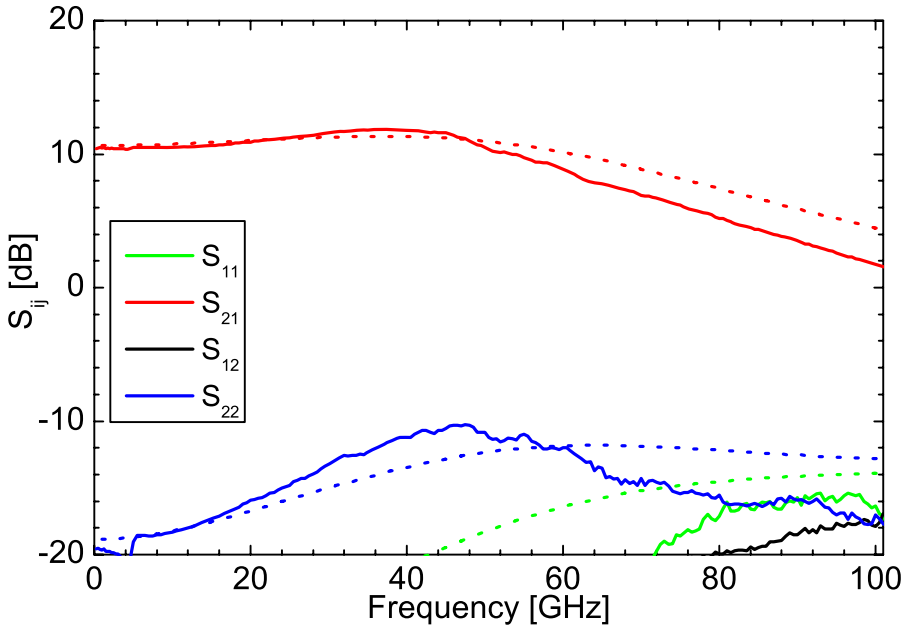


Figure 5.8: Measured scattering parameters of broad bandwidth lumped amplifier. The amplifier has a 3-dB bandwidth of 70 GHz and a gain of 10.5 dB. The simulated response is added in dashed lines.

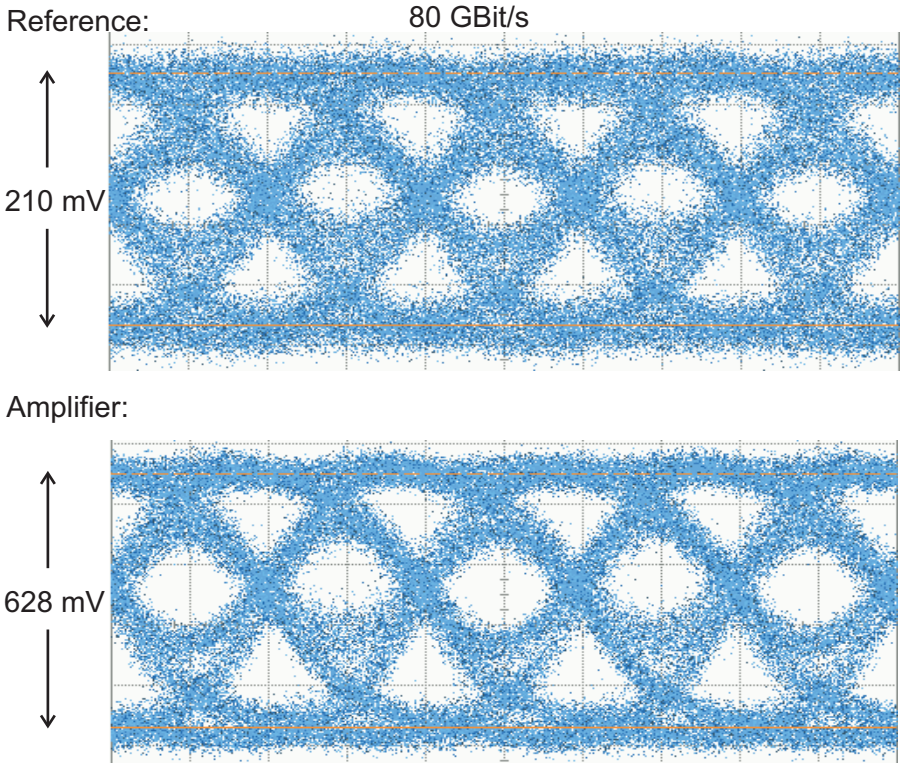


Figure 5.9: Eye diagram of broadband lumped amplifier. The gain of the digital signal is 19 dB.

The origin of the degraded input eye diagram in comparison to the 40 Gbit/s eye diagram shown above is the limited bandwidth of the measurement setup (s. Sec. 4.3.2). The output and input eye pattern have comparable shape and the quality of the output pattern is rather improved in comparison to the input eye diagram. The signal gain is 9.5 dB and the output voltage swing of 628 mV corresponds to an output power of 0 dBm or 1 mW.

Discussion

The results show, that the small-signal models (s. Sec. 3.3) can predict the behavior of the lumped amplifiers quite accurately. A likely reason for differences are parasitic effects which result from the wiring within the circuit. Accurate modeling of the short interconnects is difficult because of coupling between adjacent structures.

The bandwidth of the amplifier is 70 GHz. Still, the transmission of 80 Gbit/s signals is possible. The relative eye opening of the output eye diagram is slightly improved by the amplifier. Possible explanations for this effect are the gain peaking, which compensates part of the losses at high frequencies, and the higher output voltage swing, which increases the relative accuracy of the measurement system.

The use of only three transistors results in a small chip size, low power consumption, and high yield. All three points are advantages of the realized amplifiers. Consequently, the lumped approach is advisable for the realization of amplifiers for 80 Gbit/s signals at low output power levels. As will be shown in the next section, the lumped approach is not suitable for realizing high output power amplifiers. However, an application as a pre-amplifier for modulator drivers is possible.

5.3 Output Power Limit

This chapter has presented two lumped amplifiers, which can amplify 40 and 80 Gbit/s signals. Their output voltage swings were about 600 mV. The maximum available output power from these amplifiers is investigated in this section.

The limitation in output power can be observed in an eye diagram measurement. If the input signal power level is large enough, the output signal is limited by the maximum available output power. The gain drops and the eye diagram shows the limiting behavior by distorting the signal. This observation is shown in Fig. 5.10. The input signal has a voltage swing of 468 mV. The eye diagram measurement of the same amplifier from above (Sec. 5.2.3) predicts a gain of 9.5 dB, i.e. an output voltage swing of 1.4 V in this case. However, the swing of the output signal is limited to 806 mV, which corresponds to a gain of only 5 dB. The signal is also distorted. This behavior was found for both amplifiers at various input power levels and at data rates of 40 and 80 Gbit/s. The output voltage swing was always limited to about 800 mV ($\widehat{=}$ 2 dBm).

The power limit is due to the small transistors of the amplifier. As explained in Section 5.1, the amplifier is matched to a 50 Ω environment. A power of 2 dBm corresponds to a current swing of 16 mA. This is twice the bias current of the output transistor Q_3

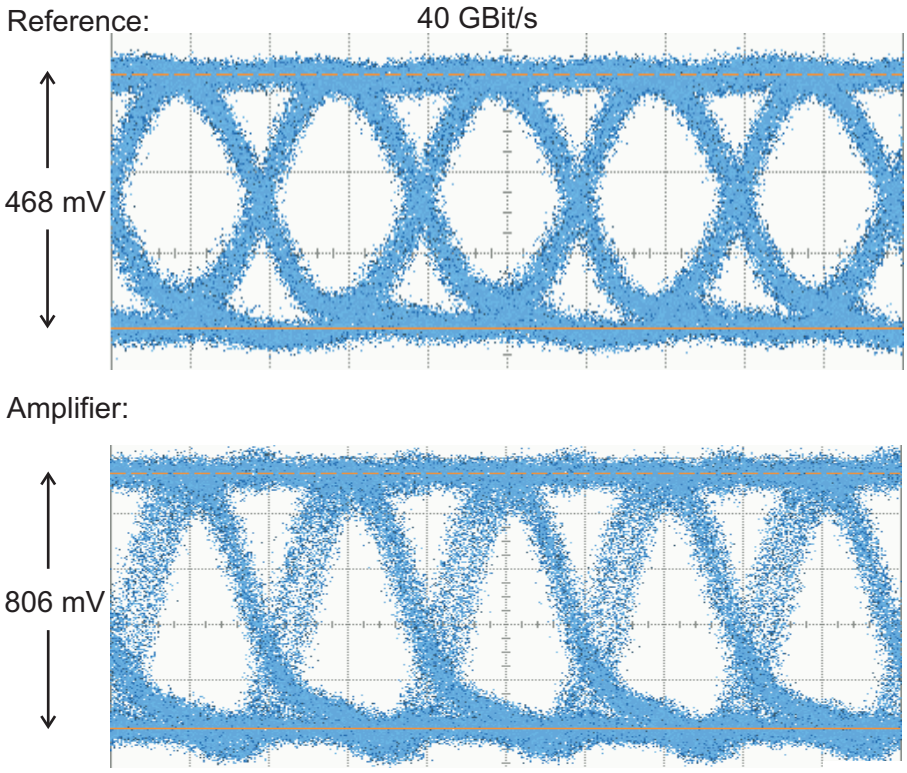


Figure 5.10: Distorted eye diagram of broad bandwidth lumped amplifier due to limited output power at 40 GBit/s.

in Fig. 5.3 and indicates, that the output signal is limited by the available current of the output transistor.

The output voltage swing is 804 mV, which is much smaller than the breakdown voltage of the transistors, which is higher than 4 V. Therefore, if the amplifier would be matched to a higher output impedance, a current swing of 16 mA would cause a higher voltage swing. The output power would increase without changing the transistor size or circuit concept. For this reason, narrow band power amplifiers use impedance matching networks at the output. A schematic of this concept is shown in Fig. 5.11.

Impedance matching networks cannot be used in broadband amplifiers. Although, these matching networks can increase the output power considerably, they have limited bandwidth. In order to keep the losses of the matching network to a minimum, reactive components have to be used. However, the efficiency of transmission and the bandwidth are exchangeable quantities in the impedance matching of any load having a reactive component [76], [32]. Therefore, the use of an impedance matching network to increase the output power of the amplifier, would reduce the bandwidth performance, which is not acceptable for amplifiers in high-bit-rate data transmission systems. Such amplifiers have to operate from frequencies below 1 MHz up to frequencies in the order of the bit rate [87].

As impedance matching networks cannot be used, the output power of the amplifier can only be increased by increasing the current that is switched by the output stage. There

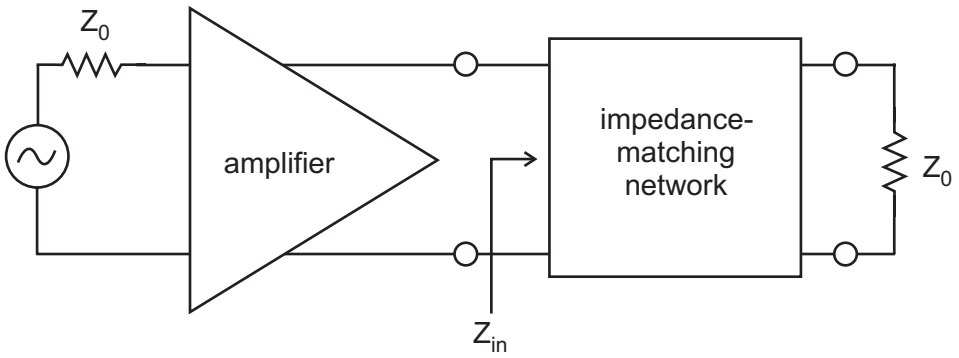


Figure 5.11: Schematic of an amplifier and an impedance matching network in a circuit environment with an impedance level Z_0. The matching network transforms the impedance of the environment to an impedance Z_{in}, which is optimized for maximum power transmission from the amplifier to the environment. Unless the impedance matching network is purely resistive (i.e. very lossy) the network has a limited bandwidth [32].

are two main concepts to achieve this. First, larger transistors can be used. Second, a distributed amplifier can be employed in which several amplifiers operate in parallel.

Using larger transistors has disadvantages. Good matching of the amplifier is more difficult to achieve, because the input capacitance of the devices and parasitic capacitances are increased [9]. This has negative consequences for the bandwidth performance. As the lumped broadband amplifier from above has a bandwidth of 70 GHz, a further reduction in bandwidth is not acceptable. Furthermore, the reliability of the circuit suffers, if high power is dissipated in the very small area of the output device [107]. Therefore, a fully lumped approach for high power driver circuits is not advisable.

To overcome these limitations, distributed amplifiers can be used. Distributed concepts do not only yield higher output power, they also allow to increase the bandwidth considerably in comparison to lumped amplifiers [8]. The following chapter deals with the development of distributed amplifiers and demonstrates their advantages with respect to bandwidth and power performance.

Chapter 6

Distributed Amplifiers for 80 Gbit/s

Distributed amplifiers (DA) achieve larger bandwidth values and more output power than lumped amplifiers. The lumped approach is mainly limited by the input capacitance and the parasitic capacitances of the transistors as stated in Section 5.3. Distributed amplifiers, on the other hand, turn these capacitances into circuit components neutralizing their limiting effect. These amplifiers have input and output artificial lines made up of a series of transmission line sections and the capacitances of the transistors. Because these artificial lines have a very high cut-off frequency, the amplifiers inherently have wideband characteristics. Therefore, the distributed amplifying technique is an excellent way to build broadband amplifiers. A schematic of a distributed amplifier is depicted in Fig. 6.1.

In this topology the gain cells couple the signal from the input transmission line to the output transmission line through each section. Each gain cell provides amplification of the signal. Ideally, the traveling wave gets additively superimposed on the output transmission line at each section as it travels toward the output. The traveling wave concept and its advantages are well documented in literature [5],[6],[14],[108].

The fact that several amplifier cells are arranged in parallel, enables distributed amplifiers to drive high currents. Therefore, higher output power levels can be expected in comparison to a lumped amplifier approach. This second advantage makes the distributed amplifier concept a very attractive candidate for the development of high-bit-rate modulator drivers, which depending on the modulator can require voltage swings of more than 5 V.

The following sections describe the development of modulator driver concepts for 80 Gbit/s applications using InP-HBTs. Section 6.1 gives details about distributed amplifier design issues. Section 6.2 describes design concepts and presents circuit results for three generations of distributed amplifiers. The final amplifier version features excellent power and bandwidth performance. The last section summarizes the key points of this chapter.

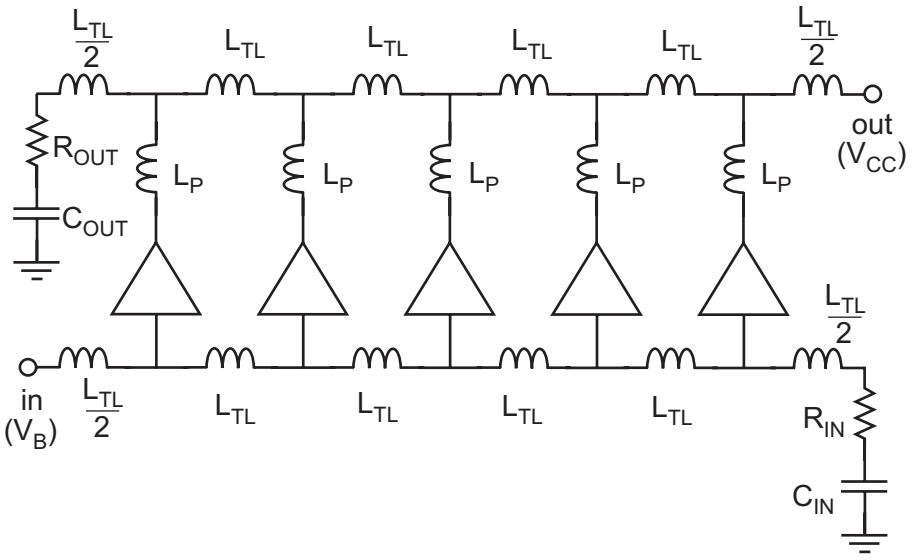


Figure 6.1: Schematic of a 5-cell distributed amplifier.

6.1 Design Considerations

The basis of the distributed amplifier concept are artificial transmission lines. These lines have cut-off frequencies, electrical delays and impedances, which are characterized by their inductance L and their capacitance C. The latter consists mostly of the input and output capacitances of the gain cells (Fig. 6.1). The cut-off frequency f_c, the electical delay T, and the impedance Z are calculated using [14]:

$$f_c = \frac{1}{\pi\sqrt{LC}} \tag{6.1}$$

$$T = \sqrt{LC} \tag{6.2}$$

$$Z = \sqrt{\frac{L}{C}} \tag{6.3}$$

.

The inductance L is usually adjusted to match the impedance Z to the circuit environment, which has an impedance level of $50\,\Omega$. From these boundary conditions it can be calculated that the cut-off frequency f_c will exceed 80 GHz, if the capacitance value C is smaller than 80 fF. In first order approximation the largest contributing to C is the input capacitance of a transistor, which is approximately equal to the capacitance C_{je} (s. Fig. 3.16). This capacitance was extracted for several transistors and its value is smaller than 80 fF, if the devices have an emitter size of less than $1 \times 10\ \mu\text{m}^2$ (s. Fig.3.17). Therefore, the cut-off frequency of the artificial transmission lines are not an issue, if small transistors are used in amplifier design. However, the electrical delay T is of interest as it can be different

in input and output transmission line due to differences in input and output capacitance of the gain cells.

Many accounts of the general theory about dimensioning and calculating distributed amplifiers in HEMT as well as HBT technology are given in literature [2],[1],[108]. In the following sections design issues are discussed, that were important for the development of distributed amplifiers presented later in this chapter.

6.1.1 Gain Cell Design Concepts

The simplest form of a gain cell consists of a single transistor in common-emitter configuration. The high-frequency performance of a such an amplifier cell is limited by the Miller capacitance [61]. The value of this capacitance can be estimated to be the collector-base capacitance multiplied by the voltage gain of the device, which is typically much larger than one. It has been shown that the Miller capacitance adds directly to the input capacitance of the transistor stage degrading its frequency response [37].

To overcome these limitations in the design of distributed amplifiers, two concepts can be used. First, the devices are scaled down to reduce all capacitance values, in general [61]. Second, the common-emitter single stage is replaced by two transistors in a common-emitter common-base configuration called cascode configuration [60], [7]. The former concept is followed, naturally, because the devices used in this work are at the lower scaling limit of the present technology (s. Sect. 2.4). The latter concept has been demonstrated to be an excellent method to achieve high performance distributed amplifiers [9], [8]. Fig. 6.2 shows a schematic of a gain cell using the cascode configuration. The cascode derives its advantage at high frequencies from the fact that the collector load for the common-emitter stage is the very low input impedance of the common-base stage. This reduces the voltage gain to nearly unity diminishing the Miller effect [37]. Other advantages of cascode cells are their very high output impedance, which reduces the losses in the output transmission line of the distributed amplifier (s. Sect. 6.1.2), and high degree of isolation from the output to the input.

More sophisticated gain cells have been proposed [4], [62]. These concepts result in state-of-the-art distributed amplifiers. However, they do not show a significant improvement in performance over concepts based on simple cascode designs. Disadvantages of such very complex amplifiers include higher power consumption, more dissipated power and possibly lower yield due to the large number of devices.

In this work the inner core of all gain cells is a cascode configuration. Modified gain cells with large devices are used to improve the power performance (s. Sect. 6.2.3). Attenuation compensation techniques are used to enhance their high-frequency performance.

6.1.2 Attenuation Compensation

Without loss compensation techniques distributed amplifiers cannot achieve their maximum wideband characteristic. The artificial transmission lines of a distributed amplifier have losses which are mainly attributed to the limited input and output impedance of

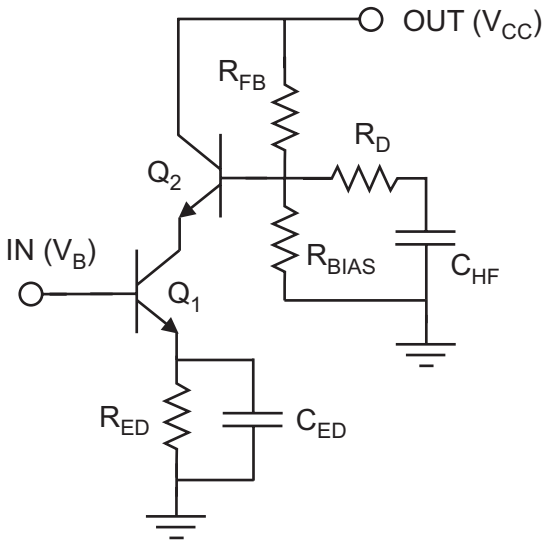


Figure 6.2: Gain cell of first distributed amplifier version employing a cascode configuration (Q_1, Q_2), self-biasing (R_{FB}, R_{BIAS}), and emitter degeneration (R_{ED}, C_{ED}).

the gain cells. These losses are particularly high at high frequencies. Attenuation compensation techniques can be used to increase the gain-bandwidth product of a distributed amplifier [26]. For the development of broad-bandwidth and high-gain amplifiers the implementation of such techniques is mandatory. This section gives an overview of most common techniques.

Cascode Configuration

Cascodes feature a higher output impedance than single-transistor common-emitter configurations [37]. Hence, the implementation of cascode configurations reduces the losses in the output transmission line [58]. A distributed amplifier using cascodes can employ more cells, which results in higher gain. In this case, the losses of the input transmission line dominate.

Emitter Degeneration

Attenuation compensation can be achieved using a resistive emitter degeneration at the common-emitter transistor of the cascode configuration. The input impedance is increased and the resistive loading of the input line is reduced [87]. The emitter resistance is also used to control the gain of the amplifier.

The emitter degeneration can be extended by a parallel capacitance (C_{ED} in Fig. 6.2). This capacitance causes peaking at high frequencies and can compensate the attenuation of the transmission lines. This technique is particularly useful for extending bandwidth [7].

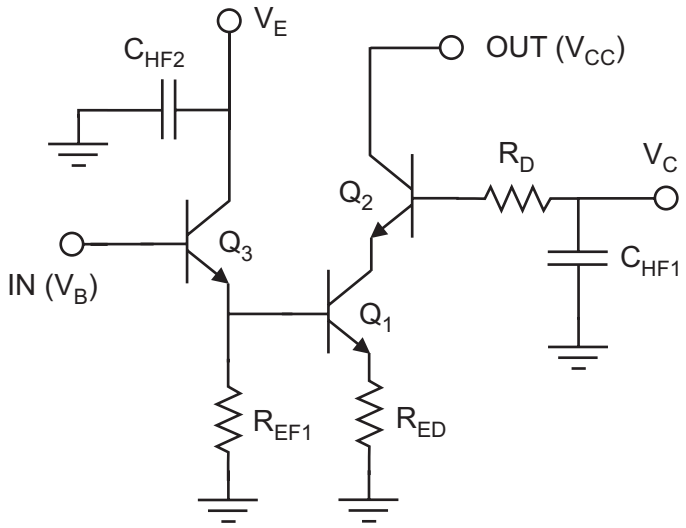


Figure 6.3: Gain cell of second amplifier version employing a cascode configuration (Q_1, Q_2), emitter degeneration (R_{ED}), and an emitter follower (Q_3).

Emitter Followers

Emitter followers at the input of a gain cell alleviate two limiting factors. The first limit in HBT distributed amplifiers is the input capacitance of the common emitter device, which creates the dominant cut-off pole for high-frequency response [61]. This high input capacitance in comparison to HEMTs is due to the close proximity of the base to the active conduction region when viewed as a charge controlled device [50]. Second, the input impedance of HBTs is low even if emitter degeneration is used and causes high losses in the input transmission line.

A gain cell using emitter follower buffering is shown in Fig. 6.3. The input impedance is increased as the emitter resistance of Q_3 is larger than emitter resistance of Q_1. At high frequencies the emitter follower device actively transforms the capacitive impedance at the emitter and generates a negative resistance at its base [58]. The combination of both effects results in lower losses in the input transmission line and increases bandwidth.

Capacitive Division

Capacitive division is a technique, which introduces a series capacitance at the input of each gain cell. It increases the input impedance of the gain cell and reduces its input capacitance [84], [5]. The former effect reduces the losses of the input transmission line and the latter effect increases the bandwidth performance of the amplifier. Very broad bandwidth HBT distributed amplifiers have been demonstrated [1]. However, capacitive division has significant disadvantages.

First, capacitive division reduces gain, as the input signal is divided between the series capacitor and the base node of the input transistor [53]. Second, due to practical geometrical considerations the series capacitor value is limited to a few pF. This results in

a deteriorated low-frequency behavior of the amplifier [1]. However, for the transmission of digital signals, constant gain performance down to frequencies as low as a 800 kHz for 80 GBit/s are desirable [2]. Simulations show, that capacitance values of 50 pF and larger would be necessary to achieve acceptable low-frequency performance. Therefore, capacitive division was not used in the development of driver amplifiers.

Peaking Lines

This attenuation compensation technique employs transmission lines between the elements of the gain cell or at its output. They transform the impedance between the elements connected. A detailed analysis of peaking lines reveals that the performance of distributed amplifiers using cascode gain cells can be improved significantly [53], [52]. In particular, a transmission line extending the output of the gain cell (L_P in Fig. 6.1) can contribute to excellent broad-bandwidth performance [7]. If used correctly, the output capacitance of the gain cells can be matched to the input capacitance of the gain cells, which is important for matching the electrical delays in input and output transmission line (s. Equ. 6.2). Also, the losses of the output transmission line can be compensated by high-frequency gain peaking. Finally, this attenuation compensation technique can also control the roll-off characteristic of the amplifier.

6.1.3 Coplanar and Microstrip Transmission Lines

Three different transmission line technologies for the realization of distributed amplifiers can be found in literature. First, microstrip lines (MSL), which use a metalization at the backside of the substrate as the ground plane. Second, inverted microstrip lines (IMSL), which use a metal layer on top of the wafer separated from the circuit by a dielectric as ground. Third, coplanar waveguides (CPW), with a ground-to-ground spacing of usually 50 μm.

IMSLs show little frequency dispersion and will avoid proximity effects, if chips are flip-chip bonded. High performance DAs have been demonstrated using this approach [75]. MSLs cannot confine the electric field in a small region. Substrates typically have a thickness of several hundred micrometer and the electric field spreads out laterally approximately the same distance. This can cause coupling between adjacent structures. Moreover, MSLs show a significant amount of frequency dispersion. Therefore, MSLs are not a good choice for the realization of distributed amplifiers.

CPWs show lower dispersion, lower coupling and require less chip area than MSLs. A drawback are higher conductor losses. However, in HBT DAs losses due to the transistor devices dominate the losses of the distributed transmission line network [59]. These losses can be compensated by using attenuation compensation techniques (s. Sect. 6.1.2). Hence, CPWs are well suited for DA design and were used in this work. A distributed amplifier layout using CPWs is shown in Fig. 6.4.

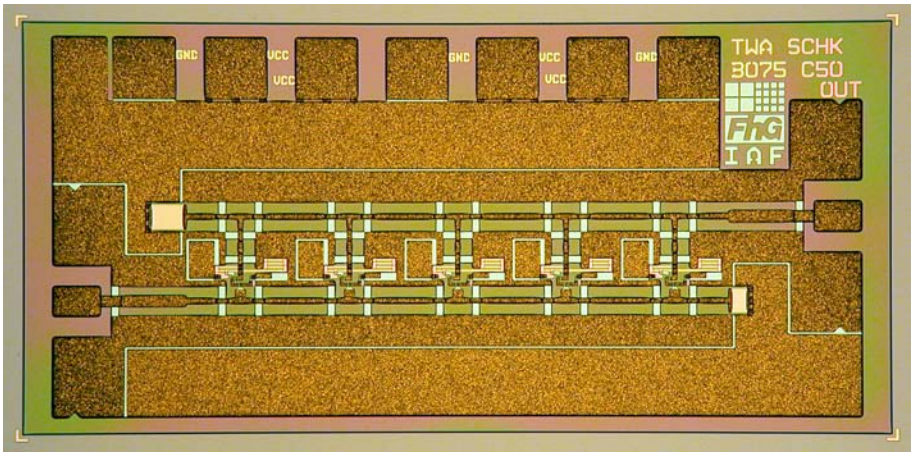


Figure 6.4: Photograph of a five-cell distributed amplifier. The chip size is $640 \times 1340\,\mu\text{m}^2$.

6.1.4 Transmission Line Termination

The artificial transmission lines of distributed amplifiers have to be terminated in order to prevent reflections. Reflections would lead to pronounced gain variations over the frequency band and degrade matching of input and output port. A straightforward method of termination is a series combination of a resistance matching the effective impedance of the artificial transmission line and a large capacitance (s. Fig. 6.1). The capacitor prevents unwanted power dissipation in R_{IN} and R_{OUT}, when the circuit is biased. The capacitance value has to be large enough to secure acceptable low-frequency performance (s. Sect. 6.1.2). In cases where this is not possible due to limited available chip area, three alternative concepts can be used to extend the low-frequency performance of distributed amplifiers: active loads [60], off-chip capacitors, or direct DC-matching terminations.

Active loads transform the on-chip capacitance into a larger value using a small network of a transistor and resistors. It has been shown, that this concept can result in distributed amplifiers with excellent low-frequency performance [57]. Simulations confirmed that the active load concept can improve the low-frequency performance of a distributed amplifier for a given on-chip capacitance value. However, the investigation also revealed that active loads reduce the bandwidth performance. The impedance transformation is frequency dependent and degrades matching of the terminations at very high frequencies and, hence, limits the bandwidth.

Instead of using active loads, large value off-chip discrete capacitors can be integrated to extend the low-frequency response. However, the integration of off-chip capacitors can cause problems. Often these discrete capacitors have self-resonances which fall within the passband of the amplifier [60]. The problem may be compounded by the effects of wirebond parasitics in combination with the discrete capacitor component [7]. These effects can sometimes result in in-band resonances and gain ripples due to standing waves in either of the input or output distributed transmission lines and are cumbersome to manage.

DC-matching termination, leaving out the capacitor, can be used to obtain flat gain

from frequencies as low as 0 Hz [53]. The drawback of this simple design is the increased power consumption caused by the bias voltage in the transmission line. Also, the circuit cannot be biased through the termination requiring the implementation of bias-Ts at the input and output of the circuit.

6.1.5 Linear and Matrix Design

Distributed amplifiers can follow linear or matrix concepts. A linear concept features two artificial transmission lines for the input and the ouput connected by a row of gain cells (s. Fig. 6.1). This is the most widely used concept. Matrix amplifiers consist of two or more rows of gain cells. The output transmission line of a lower row serves as input transmission line for the upper row. This results in a layout where gain cells are arranged like elements in a matrix.

Matrix amplifiers offer more flexibility in using device scaling and capacitive coupling techniques [61]. However, there is no indication that a matrix amplifier has a better performance than a cascade of several linear amplifiers with the same total number of gain cells. An experimental comparison between a two-row matrix distributed amplifier and a dual-stage linear distributed amplifier did not show any advantage of either concept over the other [25]. Recently, half-matrix half-dual-stage distributed amplifier concepts have been reported without demonstrating significant advantages [56]. If complex concepts do not lead to improved circuit performance, they should not be implemented. It would lead to higher power consumption and lower yield, i.e. higher costs. Consequently, this work deals with the development of linear distributed amplifiers.

6.1.6 Single-ended and Differential Design

Distributed modulator driver amplifiers can be designed for differential or single-ended input and output. An advantage of differential designs is, that the voltage output swing can be considerably larger than in a single-ended design. Also, multiplexers usually have a differential output, because differential multiplexer designs reach higher speed than single-ended concepts. For 40 GBit/s excellent modulator driver performance using differential designs has been demonstrated [110], [9].

Differential amplifiers usually consist of two single-ended amplifiers, which are coupled at each gain cell. In first order approximation the performance of the differential amplifier is determined by the single-ended design concept. Therefore, an evaluation of modulator driver concepts can be conducted using the single-ended approach. The lower complexity results in higher yield and shorter design cycles during the development. This work concentrates on the optimization of single-ended amplifiers.

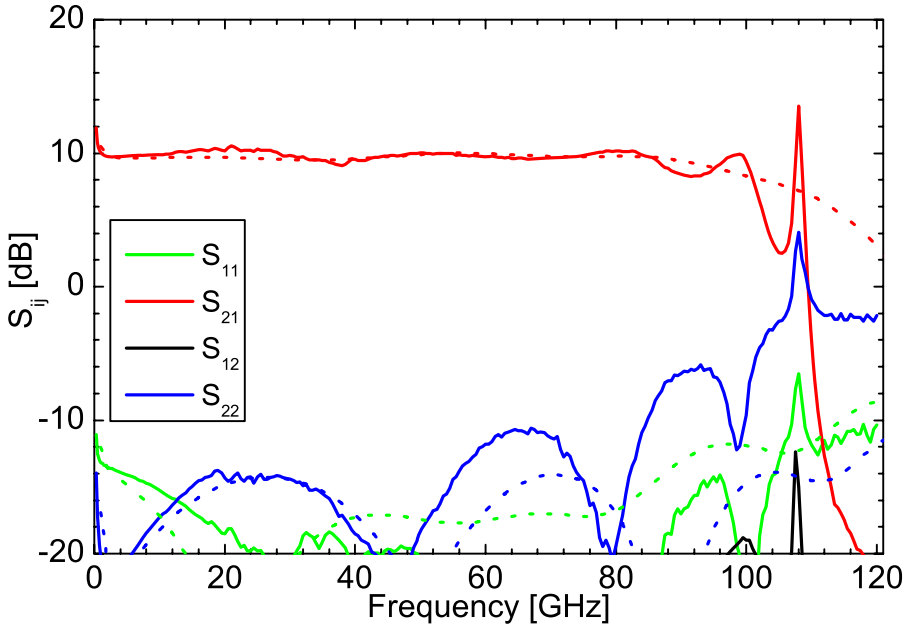


Figure 6.5: Measured scattering parameters (solid lines) of distributed amplifier using resistive feedback. The gain is 10 dB and the 3-dB bandwidth is 102 GHz. The simulated response is added in dashed lines.

6.2 Realized Amplifiers

In this section three distributed amplifiers are presented, corresponding to three major development steps toward the realization of high-power modulator drivers for 80 GBit/s. First, a low-power very broad bandwidth amplifier is shown. Second, a more sophisticated amplifier removes some deficiencies of the previous version and introduces a loss compensation technique. Third, an amplifier optimized for high output power operation is presented. All amplifiers are characterized using scattering parameter, eye diagram, and power measurements.

6.2.1 Low-Power 100 GHz Bandwidth Amplifier

The design goal of the first version was simplicity and demonstration of the technology's potential, i.e. bandwidth. A low number of devices results in high yield. A direct comparison to the lumped amplifiers in Chapter 5 can be made.

Design and Layout

The gain cell of this amplifier is shown in Fig. 6.2 and uses transistors with an emitter size of $1 \times 4\ \mu m^2$. It features a cascode configuration, whose advantages were described in Sects. 6.1.1 and 6.1.2. Although the cascode configuration improves the output resistance and bandwidth performance, it tends to be unstable for HBTs [58]. To ensure stability the

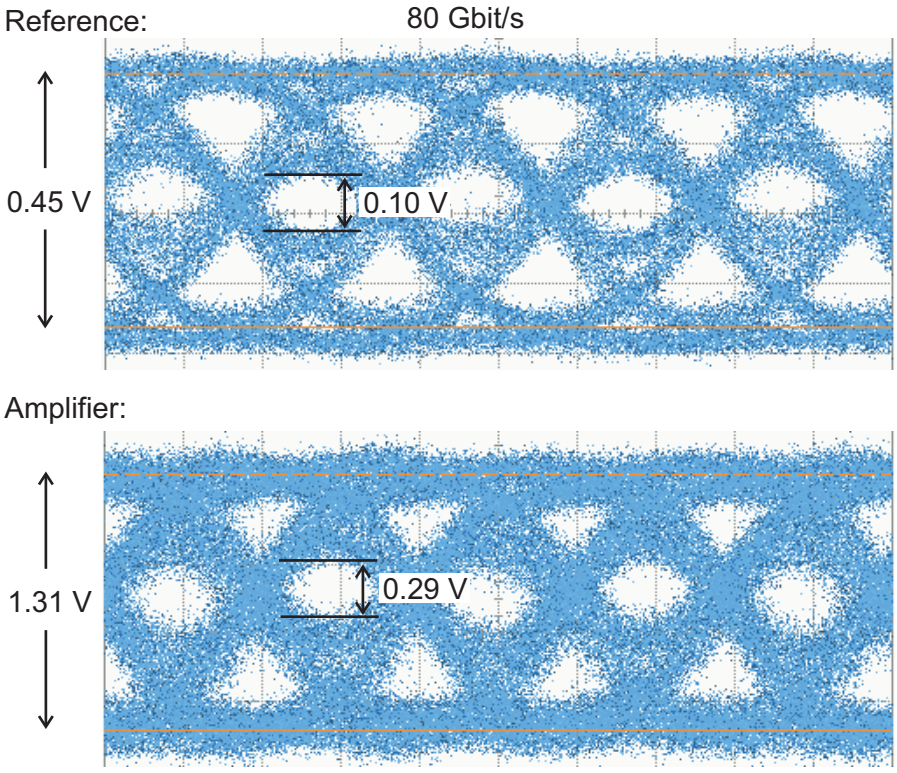


Figure 6.6: Eye diagram of first distributed amplifier version at 80 GBit/s. The signal gain is 9 dB. The relative eye openings of reference and amplifier eye are 22 %.

gain cell is biased through the output transmission line using a voltage divider provided by two resistors. One of them provides negative feedback (R_{FB}), which improves amplifier stability [59]. The stability behavior of the cascode stage is further enhanced by adding a small resistance (R_D) for negative feedback in series with the shorting capacitance (C_{HF}).

Any inductance between the input line and the input of the active cell can seriously degrade the bandwidth of the amplifier [7]. Hence, the active cell is placed as close as possible to the center conductor of the coplanar transmission line.

For attenuation compensation emitter degeneration (R_{ED}) and capacitive peaking (C_{ED}) are used. At the output the gain cell is extended by a transmission line (L_P) to achieve line peaking (s. Fig. 6.1). Careful tuning of these three parameters allows adjusting the input and output impedance of the gain cell. However, the input and output impedance of the gain cells cannot be made equal as the input capacitance of the gain cell is large (s. Sec.6.1.2). Consequently, the input and output transmission lines have different impedance levels. To optimize matching of the amplifier, the input line impedance was adjusted to be below 50 Ω and the output impedance was adjusted to be above 50 Ω. Together with termination resistances of 35 Ω (R_{IN}) and 70 Ω (R_{OUT}) good matching can be achieved. The terminations are extended by large value series capacitors ($C_{IN} \approx C_{OUT} \approx 35$ pF) to reduce the power dissipation while enabling low-frequency operation.

The number of cells in this amplifier is limited by two factors. First, the difference

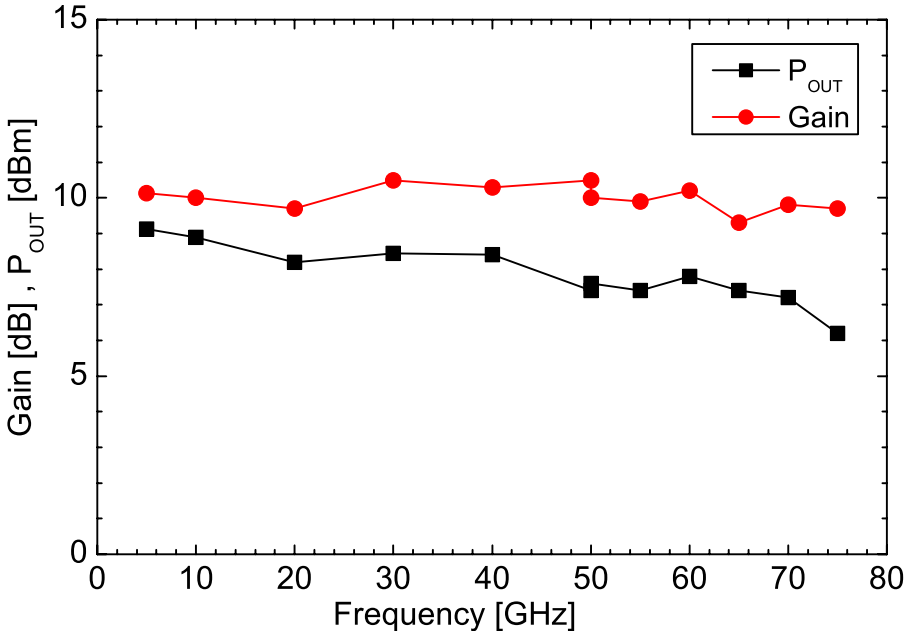


Figure 6.7: Output power at 1-dB compression point and gain of distributed amplifier employing resistive feedback. The discontinuity at 50 GHz is due to the measurement system (s. Sec. 4.3.1).

in impedance level results in different electrical delays in input and output transmission lines (Eq. 6.2). Hence, the electrical signals of each gain cell, which are superimposed in the output transmission line, are not exactly in phase. If the lines are too long, the effect degrades the bandwidth. Second, the low input impedance of the gain cells attenuates the input signal, significantly. If the number of cells is too high, the signal at the last cell will be too low for the gain cell to contribute a noticeable amount of gain. Still, this gain cell would cause finite losses in the output line reducing the gain of the amplifier. If the number of cells is too low, the gain will be lower than necessary. According to simulations a five-cell amplifier as shown in Fig. 6.1 is a good compromise. A chip photograph of the amplifier is shown in Fig. 6.4.

Results

The amplifier was measured with a power supply of 3.5 V (V_{CC}) and 1.0 V (V_B). The bias current was 43 mA resulting in a power consumption of approximately 150 mW. This corresponds to a current density of 1.9×10^5 A/cm^2 in each transistor.

The bandwidth performance is shown in Fig. 6.5. The small-signal gain is 10 dB and the 3-dB bandwidth is 102 GHz. The input matching is better than -13 dB over the full bandwidth and the output matching is better than -10 dB up to 85 GHz. The simulated data are also shown in dashed lines in Fig. 6.5. The oscillations in the measured amplifier response beyond 80 GHz are attributed to excessive peaking. Capacitive peaking enhances the bandwidth, but it is very sensitive. Because a slight reduction in bandwidth

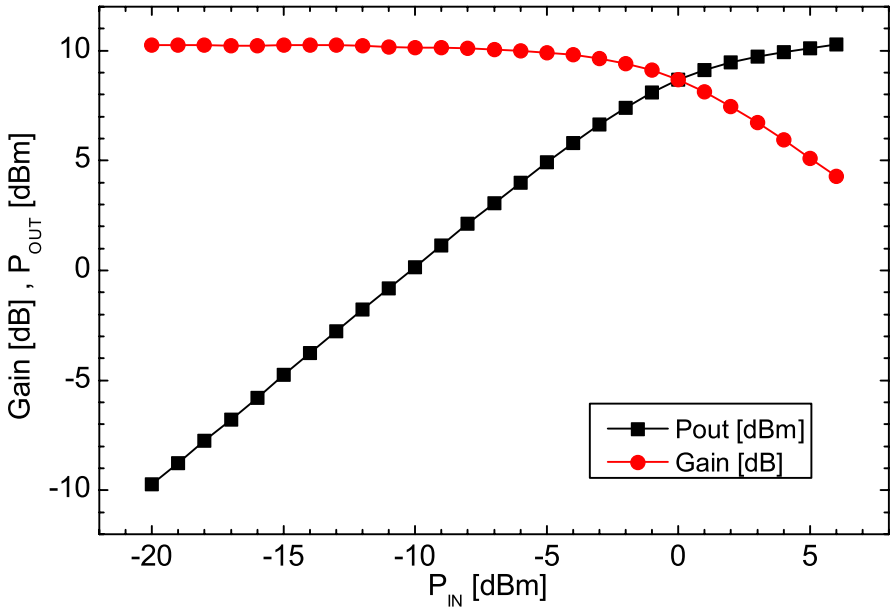


Figure 6.8: Output power vs. input available power and gain of distributed amplifier employing resistive feedback measured at a frequency of 60 GHz. The 1-dB compression point is at 8 dBm output power.

is acceptable for the transmission of 80 Gbit/s data signals, this peaking technique is not used in the amplifiers presented in the following Sections.

An eye diagram measurement is presented in Fig. 6.6. The input eye is already degraded due to bandwidth limitations in the measurement setup (s. Sec. 4.3.2). The output eye is slightly degraded in comparison to the input eye. Still, the relative eye openings of both the input and output eyes are 22 %. The signal gain is 9 dB. The output eye does not show any limiting effect, which indicates that the amplifier can support larger signals. In order to estimate this maximum output swing, power measurements were conducted.

A power and large-signal gain measurement is depicted in Fig. 6.7. The graph shows the output power at 1-dB compression point. The slight discontinuity at 50 GHz is due to different measurement setups for the frequency bands above and below this frequency (s. Sec. 4.3.1). The output power starts at 9 dBm for low frequencies and gradually decreases to 6 dBm at 75 GHz. The large-signal gain is 10 dB and matches the small-signal gain in Fig. 6.5. Fig. 6.8 shows a power sweep of the amplifier at 60 GHz. The 1-dB compression point is reached at an output power of almost 8 dBm, which corresponds to a peak-to-peak voltage swing of 1.6 V. The results of the power measurement indicate, that the maximum voltage swing of the amplifier for an 80 GBit/s signal is between 1.3 V ($\widehat{=}$ 6 dBm) and 1.8 V ($\widehat{=}$ 9 dBm).

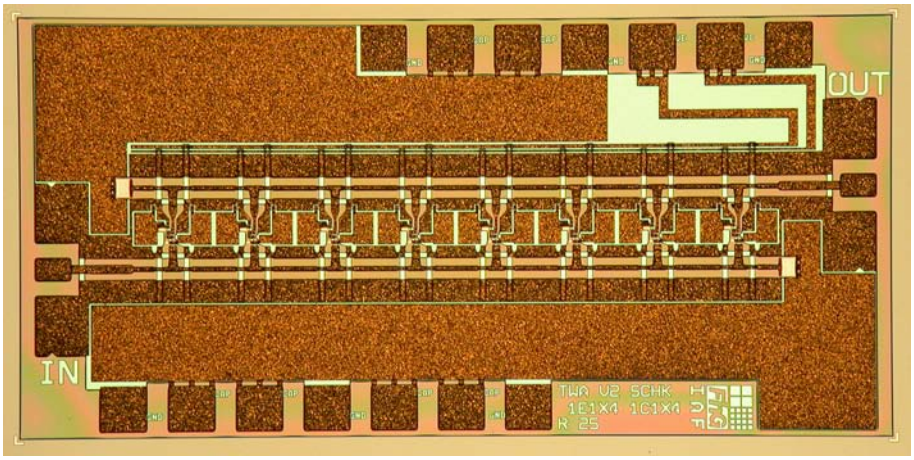


Figure 6.9: Photograph of second amplifier version using emitter followers and external bias ports. The chip size is $900 \times 1860\,\mu\text{m}^2$.

Discussion

This first distributed amplifier version clearly demonstrates the advantages of the distributed concept. Due to the lower operating voltage the power consumption of this amplifier is comparable to the lumped amplifiers presented in Chapter 5. However, the bandwidth is 45 % larger than in the case of the lumped amplifier while reaching the same gain. The maximum output swing is 60 % larger. The following versions of distributed amplifiers modify the concept to enhance gain and power performance while maintaining a bandwidth that supports the amplification of 80 GBit/s signals.

6.2.2 Tunable Amplifier for Loss Compensation

While the first distributed amplifier version achieved remarkable results using basic design principles, it lacks features that are necessary for modulator drivers. The number of cells is limited, resulting in a limitation of gain and output power. Both figures of merit are improved in this second version by using attenuation compensation techniques in combination with a larger number of cells.

Design and Layout

The gain cell of this amplifier uses transistors with an emitter size of $1 \times 4\ \mu\text{m}^2$ and employs an emitter follower at its input (Fig. 6.3). This results in positive consequences. The resistive loading as well as the capacitive loading of the input transmission line are reduced (s. Sec. 6.1.2). Lower resistive loading lessens the attenuation in the input transmission line. The input signal after the fifth cell is still strong enough for utilizing additional gain cells. This adds to the amplifier gain. According to simulations a total number of eight cells can be used. For simulations a small-signal model was used.

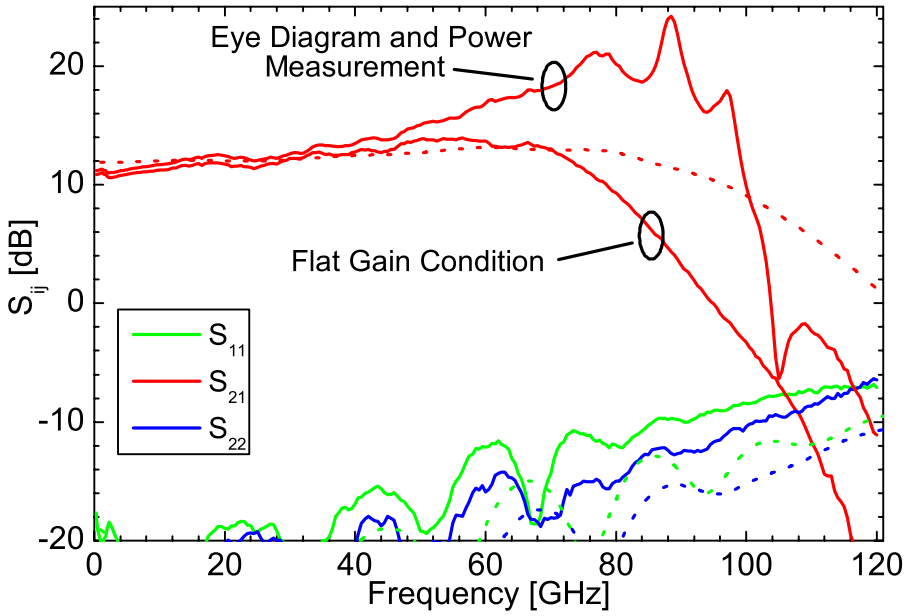


Figure 6.10: Scattering parameters of tunable distributed amplifier. Solid lines: response at operating point for flat gain and small-signal gain at operating point for eye diagram and power measurements. Dashed line: Simulated response for flat gain.

The lower input capacitance makes it possible to match the input and output capacitance of the gain cells. In contrast to the previous amplifier version the impedance of the artificial input and output transmission lines can be made equal and matched to the 50 Ω circuit environment. This leads to significant improvement in the matching of amplifier input and output. Furthermore, the electrical delays in input and output transmission line are equal (Eq. 6.2), which ensures that the gain cell output signals are in phase.

Line peaking and emitter degeneration as additional attenuation compensation techniques are used. These techniques were employed in first amplifier version, too. However, capacitive peaking is not used for two reasons. First, capacitive peaking often leads to a steep roll off, which can cause signal distortions at very high bit rates. For data transmission systems it is advisable to trade a certain amount of bandwidth for a gradual roll off [8]. Second, this amplifier does not use negative feedback at the cascode (R_{FB} in Fig. 6.2) which makes the amplifier more sensitive to oscillations. Reducing the number of peaking mechanisms alleviates this problem.

Negative feedback at the cascode is not used, because the amplifier has external bias ports for the base of transistor Q_2 (Fig. 6.3). The additional external bias port can be used to compensate minor deficiencies in design or technology to a certain degree. The bias conditions of Q_1 and Q_2 can be set independently allowing to alter the frequency characteristic of the amplifier. A chip photograph of the amplifier is depicted in Fig. 6.9.

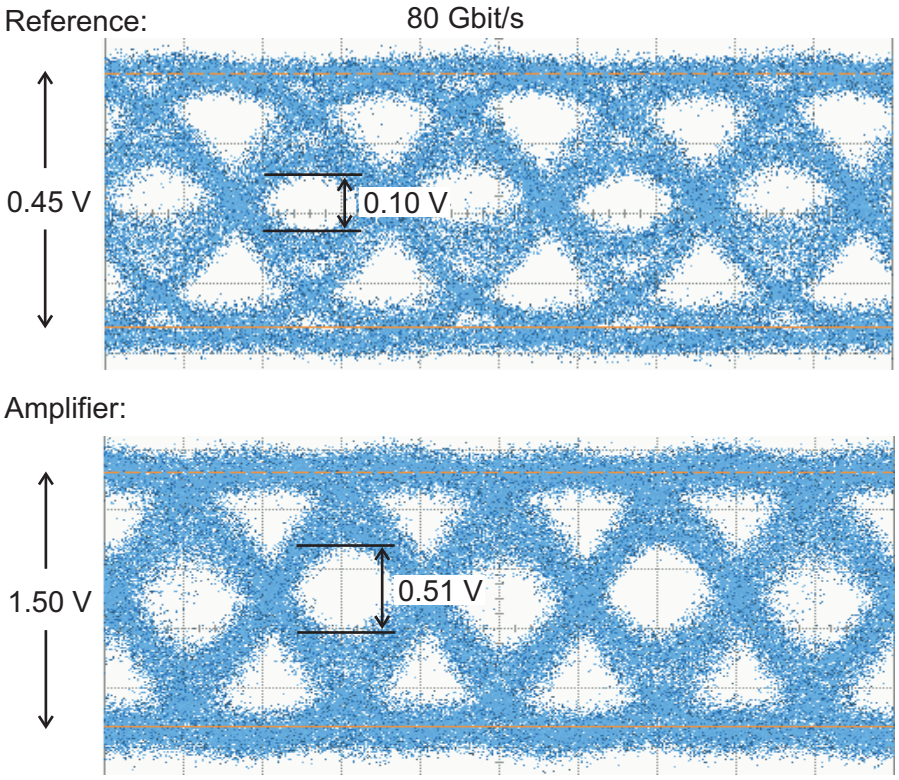


Figure 6.11: Eye diagram of second distributed amplifier version at 80 GBit/s. The gain of the signal is 10.5 dB. The relative eye openings of input and output eye are 22 % and 34 %, respectively.

Results

The amplifier was measured in two different operating conditions. First, it was biased for flat small-signal gain. Second, a bias point was chosen, which optimizes the shape and height of the eye diagram and maximizes the output power. The bias conditions and corresponding power consumptions are given in Table 6.1. The values correspond to a current density of $1.9...2.4 \times 10^5$ A/cm^2 in each transistor.

Measured scattering parameters of this amplifier are shown in Fig. 6.10. If operated in

condition	V_{CC} (V)	I_{CC} (mA)	V_C (V)	I_C (mA)
flat gain	3.8	59	2.0	≈ 1
max. power	4.2	78	2.6	≈ 1

condition	V_E (V)	I_E (mA)	V_B (V)	I_B (mA)	P (mW)
flat gain	2.4	67	2.1	≈ 1	420
max. power	3.0	71	2.2	≈ 1	540

Table 6.1: Bias conditions and power consumption for flat gain and maximum output power condition of distributed amplifier with one emitter follower in each gain cell.

flat gain condition, the small-signal gain is 11 dB and the 3-dB bandwidth is 82 GHz. The roll off is very gradual. The matching of the input and output is better than −10 dB over the full bandwidth. The small-signal gain for eye and power measurement is shown also. In this operating condition the gain increases to higher frequencies and starts to oscillate above 80 GHz. The simulated response for flat gain is shown in dashed lines. The reduced bandwidth in the measurement is attributed to a fact, that the peaking mechanisms are too weak. In comparison to the previous amplifier, which exhibited too much peaking, this amplifier rather uses too little.

The result of an 80 GBit/s eye diagram measurement is depicted in Fig. 6.11. The output swing is 1.5 V and is limited by the available input swing of the measurement system. The gain of the digital signal is 10.5 dB. The output eye is clearer and the eye opening is larger than in the case of the input eye. The relative eye opening has improved from 22 % for the reference signal to 34 % for the amplifier output signal.

Power measurements were conducted to investigate the maximum output power of this amplifier. Fig. 6.12 shows the output power against input power at a frequency of 60 GHz. For low and medium output power levels the gain is constant at 17 dB and the amplifier behavior is linear. The border of the linear region is the 1-dB compression point, where the gain has decreased by 1 dB. The output power at this point is 13 dBm. Fig. 6.13 summarizes the result of this measurement over a frequency range from 5 to 75 GHz. The output power at 1-dB compression is almost constant at 14.5 dBm up to 40 GHz. At 75 GHz the output power has moderately decreased to 12 dBm. The large-signal gain of the power measurement corresponds to the small-signal gain of the scattering parameter measurement (Fig. 6.10).

Discussion

The scattering parameters demonstrate improvements in comparison to the first DA version. The matching of the amplifier is excellent. This is due to the adjustment of the impedance in the artificial input and output transmission lines, which was enabled by the use of an emitter follower in each gain cell. The gain could also be increased by 1 dB at low frequencies and it gradually rises up to 70 GHz by 2 dB. This behavior is called pre-emphasis and is useful for compensating losses, that get introduced, when the chip is packaged. These losses rise with frequency. The bandwidth performance is reduced. However, the bandwidth still exceeds 80 GHz and the roll off is very gradual, which is of advantage for digital signal transmission.

The output eye diagram opening is much larger and clearer, and the amplitude is slightly higher than in the case of the previous amplifier. The increase in amplitude corresponds to the increase in small-signal gain. The improved eye opening can be attributed to the gain peaking in the operating point of the eye measurement (s. Fig 6.10). As explained in section 4.3.2, the setup for the eye diagram measurement is bandwidth limited with a nominal bandwidth of 60 GHz. Higher frequencies get attenuated significantly. Apparently, gain peaking can compensate these losses producing a clearer eye opening in comparison to the reference signal.

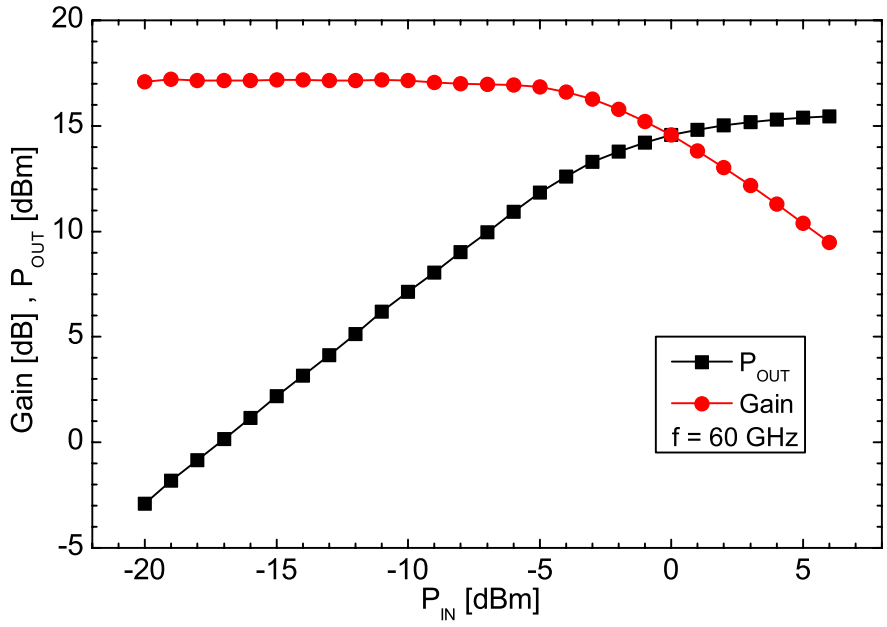


Figure 6.12: Output power against available power at 60 GHz of second DA version. The output power at 1-dB compression point is 13.5 dBm.

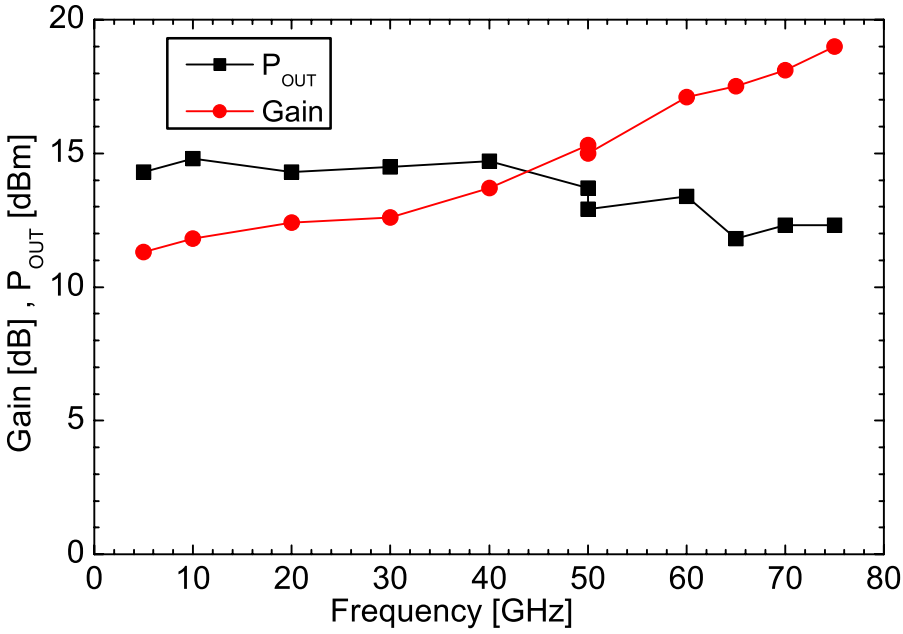


Figure 6.13: Output power at 1-dB compression point and gain of second DA version.

The output power capability of this amplifier has been improved by 5 dB over the full bandwidth with a very gradual decrease of only 2.5 dB from 5 to 75 GHz. This corresponds to an increase in voltage output swing of 80 %. From the results of the output power measurement it can be concluded that the maximum peak-to-peak voltage swing of an 80 GBit/s signal would have a value between 2.6 V ($\widehat{=}$ 12 dBm) and 3.4 V ($\widehat{=}$ 14.5 dBm).

Besides the performance improvements a striking feature of this amplifier concept is the tuning capability. As demonstrated it can be used to adjust the high-frequency performance to compensate system limitations. An unavoidable drawback of this design is higher power consumption due to the larger number of devices used and a larger chip size. Both can lead to a lower yield. The last amplifier version described in the following section takes another step toward higher output power while maintaining bandwidth.

6.2.3 High-Power Amplifier for 80 Gbit/s

The previous amplifier has shown an excellent performance regarding the amplification of 80 Gbit/s digital signals. For modulator driver applications higher output power is still desirable. Therefore, the last development step focuses on achieving more output power without sacrificing advances of previous development steps.

Design and Layout

The breakdown voltage of the InP-HBTs is high and the expectations regarding power performance are not met by the power performance of amplifiers presented in the last two sections. An advancement in circuit design is needed to exploit the potential of the devices. In narrow bandwidth amplifiers a matching network can be used to transform the 50 Ω circuit environment into a load which extracts the maximum power out of a device for a given operating point. However, these matching networks have a very limited bandwidth and cannot be used in broadband amplifiers necessary for high-bit-rate application (s. Sect.5.3).

Peaking lines at the output of the gain cells of distributed amplifiers transform the impedance, but their purpose is to cause peaking at high frequencies and not to improve the power performance. At lower frequencies impedance transformation by line peaking is negligible. Therefore, in first order approximation the load for the output device of each gain cell is 50 Ω.

The signal in the output transmission line gets stronger as it travels toward the amplifier output, and, consequently, the stress on the output transistor of the last gain cell is largest. The voltage swing at this device is approximately equal to the amplifier output voltage swing. Hence, the output power of the amplifier is limited by the output current unless the voltage swing reaches the breakdown voltage of the devices.

The output driving current in this design is increased by using larger transistors. In previous circuits the emitter area of all devices was $4 \times 1\,\mu\text{m}^2$. By using a device of twice the size the output current of the amplifier can be doubled promising higher output power and more gain, if the number of cells is not changed.

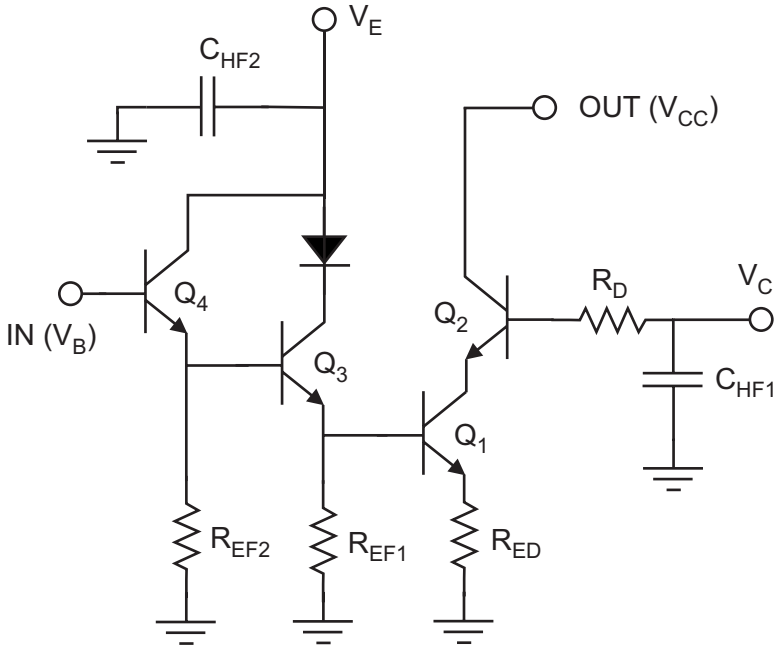


Figure 6.14: Schematic of gain cell optimized for maximum output power. The cascode devices (Q_1, Q_2) have an emitter size of $8 \times 1\,\mu\text{m}^2$. Emitter followers ($Q_3$, Q_4) and diode have an emitter size of $4 \times 1\,\mu\text{m}^2$.

A schematic of the gain cell of this new amplifier is shown in Fig. 6.14. The cascode (Q_1, Q_2) consists of transistors with an emitter size of $8 \times 1\,\mu\text{m}^2$. The implementation of larger transistors increases the input capacitance and reduces the input resistance of the cascode stage considerably. To compensate this, more attenuation compensation is needed and two emitter followers (Q_3, Q_4) with the standard transistor size of $4 \times 1\,\mu\text{m}^2$ are implemented. To ensure equivalent bias conditions for Q_3 and Q_4 a diode is inserted at the collector port of Q_3. As a result the loading of the input transmission line is low enough to design a distributed amplifier with eight cells.

The large cascode devices also cause an increase in output capacitance. This increase can be compensated partly by adjusting the attenuation compensation mechanism of the peaking line L_P (s. Fig. 6.1). However, the input capacitance and output capacitance of the gain cells cannot be made equal. This leads to different impedance levels in input and output artificial transmission lines of the distributed amplifier. As discussed before the matching of the amplifier can be degraded and different electrical delays in the input and output line can degrade the gain performance at high frequencies. However, simulations show, that this degradation is acceptable. It is the price for achieving higher output power levels, while keeping amplifier concepts as simple as possible.

As before, external ports for biasing the emitter followers and the base of the cascode are provided allowing the compensation of design errors and process variations to a certain degree. Shorting capacitances (C_{HF1} and C_{HF2} in Fig. 6.14) are provided to optimize the high-frequency performance. The design and layout concepts of the artificial transmission

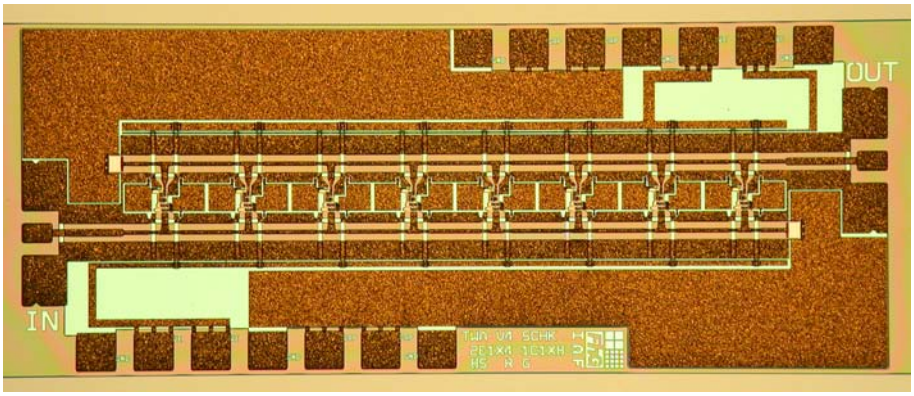


Figure 6.15: Photograph of final distributed amplifier version optimized for output power. The chip size is $900 \times 2280\,\mu\text{m}^2$.

lines and terminations are identical to the previous amplifier versions. A total of 40 devices is used in this amplifier. A chip photograph of the distributed amplifier is given in Fig. 6.15.

Results

The amplifier was measured in different operating conditions. For maximum bandwidth it was biased at moderate voltages and currents. For maximum output power the devices in the amplifier were put under more stress. The bias conditions and corresponding power consumptions are given in Table 6.2. The values correspond to a current density of $1.8...2.7 \times 10^5$ A/cm^2 in each transistor.

Measured scattering parameters are shown in Fig. 6.16. If the amplifier is biased for large bandwidth and moderate peaking, the gain is 12 dB with a peaking of 3 dB and a bandwidth of 94 GHz. The matching of the input and output is better than -10 dB up to 75 GHz and over the full bandwidth, respectively. The amplifier response displays ripples in the gain for frequencies above 80 GHz. The roll-off is fairly steep. The small-signal gain at other operating points is added, too. Dashed lines show the simulated amplifier response

condition	V_{CC} (V)	I_{CC} (mA)	V_C (V)	I_C (mA)
max. bandwidth	3.8	121	2.2	≈ 1
eye diagram	4.0	131	2.4	≈ 1
max. power	4.5	175	2.5	≈ 1

condition	V_E (V)	I_E (mA)	V_B (V)	I_B (mA)	P (mW)
max. bandwidth	3.3	117	2.8	≈ 1	850
eye diagram	3.6	137	2.9	≈ 1	1020
max. power	3.9	165	3.0	≈ 1	1430

Table 6.2: Bias conditions and power consumption for maximum bandwidth, eye diagram measurement, and maximum output power condition of 8-cell distributed amplifier with two emitter followers and large area cascode devices (s. Fig. 6.14).

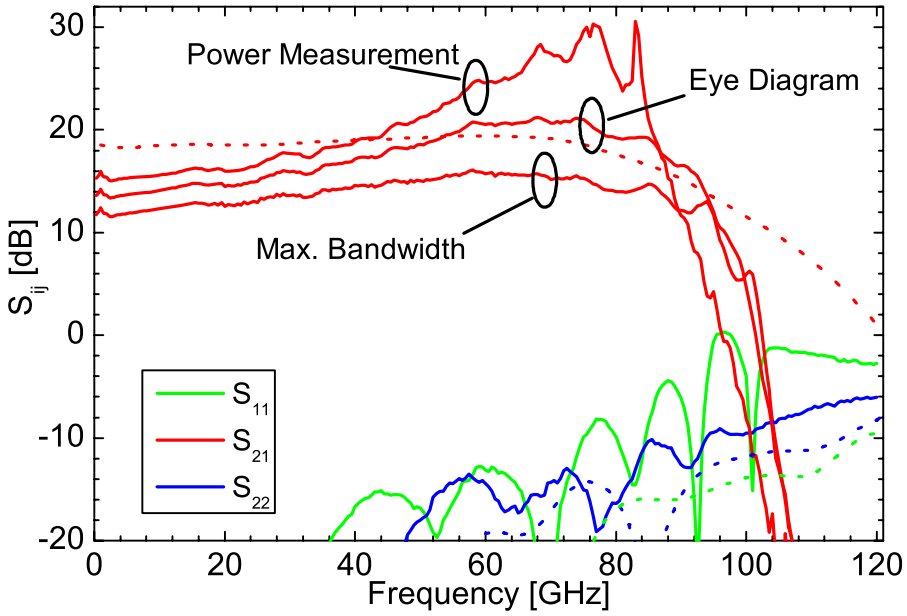


Figure 6.16: Scattering parameters of high-gain distributed amplifier. Solid lines: response at operating point for maximum bandwidth and small-signal gain at operating points for eye diagram and power measurements. Dashed lines: Simulated response for maximum bandwidth condition.

for flat gain condition.

The result of an 80 GBit/s eye diagram measurement is depicted in Fig. 6.17. The output swing is 2.4 V and is limited by the available input swing of the measurement system. The gain for the digital signal is 14.5 dB. The output eye is clearer and the eye opening is much larger in comparison to the input eye. The relative eye opening of the output eye is 44 % in comparison to 22 % for the reference signal. The optimum eye shape was achieved using the tuning capability of the amplifier. The small-signal gain that relates to the eye diagram measurement condition is given in Fig. 6.16.

Power measurements were conducted to investigate the maximum output power of this amplifier. Fig. 6.18 shows the output power against input power at a frequency of 60 GHz. For low and medium output power levels the gain is almost constant around 24 dB and the amplifier behavior is linear. The 1-dB compression point is reached at an output power level of 15.5 dBm. Fig. 6.19 summarizes the output power performance over a frequency range from 5 to 75 GHz. The output power at 1-dB compression is almost constant at 18 dBm up to 40 GHz and decreases less than 3 dB up to 60 GHz. Above this frequency the output power roles off faster to 10.5 dBm at 75 GHz. The large-signal gain of the power measurement corresponds to the small-signal gain in the same operating point (Fig. 6.16).

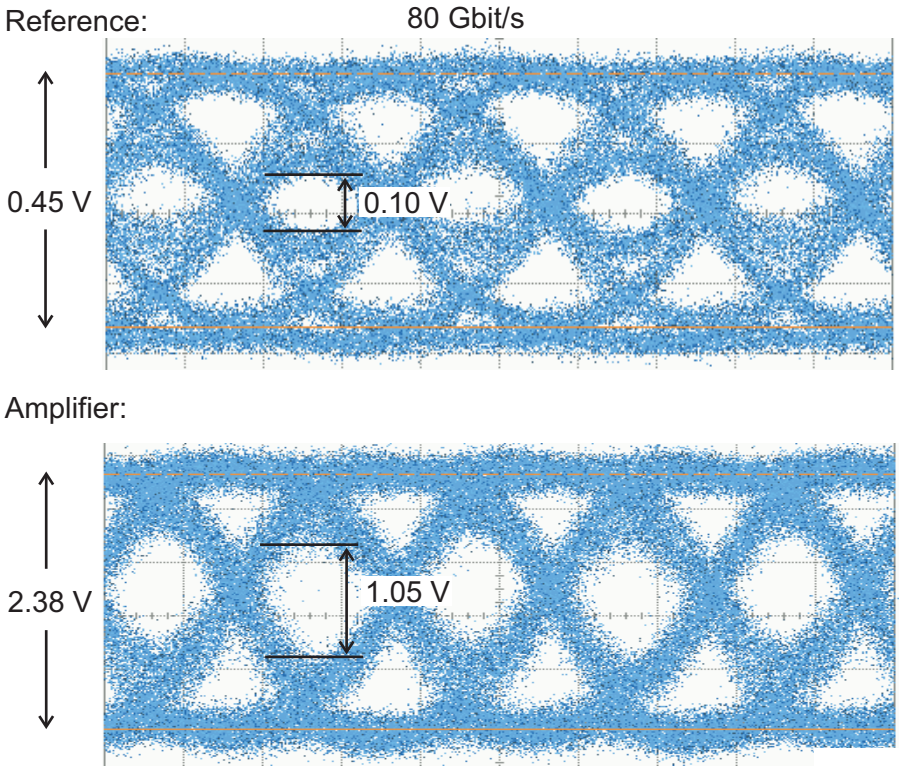


Figure 6.17: Eye diagram of distributed amplifier optimized for maximum output power at 80 GBit/s. The gain of the signal is 14.5 dB. The relative eye openings of reference and amplifier output eye are 22 % and 44 %, respectively.

Discussion

The amplifier's performance meets the expectations. The gain performance exceeds the previous versions and the bandwidth is still larger than 80 GHz. The amplifier tends to oscillate at very high frequencies because excessive peaking had to be used to compensate the change in transistor size. The matching behavior is also degenerated at high frequencies due to this change, but it is still acceptable. However, up to 80 GHz the performance of the amplifier is excellent.

This statement is confirmed by the eye diagram measurement at 80 GBit/s. In comparison to the reference eye the output eye is not only large but also very clear and the opening of the eyes has improved, dramatically. Similar to the previous amplifier version, this is achieved by compensating losses in the measurement system with pronounced peaking of the amplifier. The output eye does not show any limiting effects which strongly suggests that the maximum available output swing could be considerably larger, if a larger input signal would be available.

The power measurements indicate that the maximum output swing at 80 GBit/s is between 2.6 V ($\widehat{=}$ 12 dBm) and 5.1 V ($\widehat{=}$ 18 dBm) with a gain of more than 14 dB. The power spectrum of an 80 GBit/s digital signal gradually decreases from its maximum at

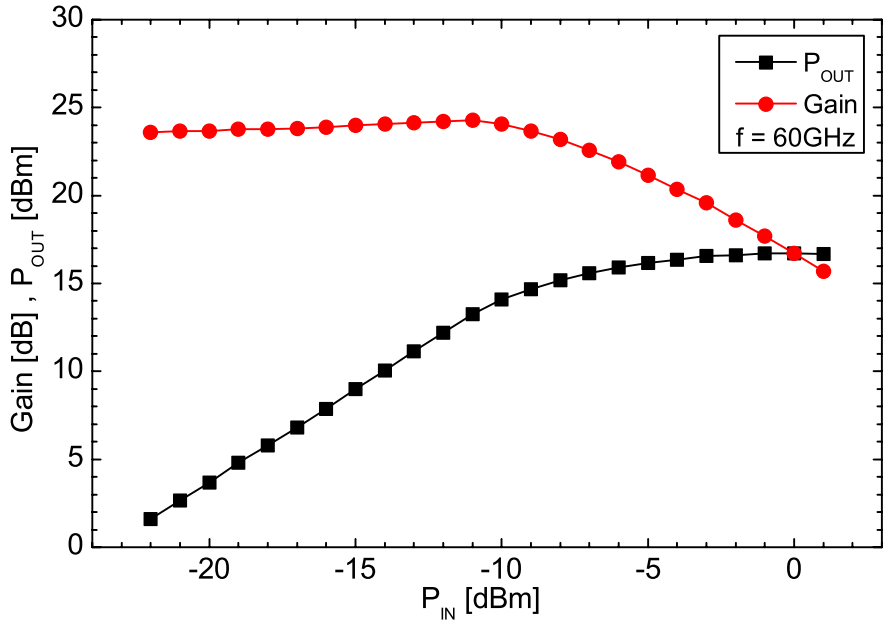


Figure 6.18: Output power against input power of final amplifier version at 60 GHz. At 1-dB compression point the output power is 15.5 dBm.

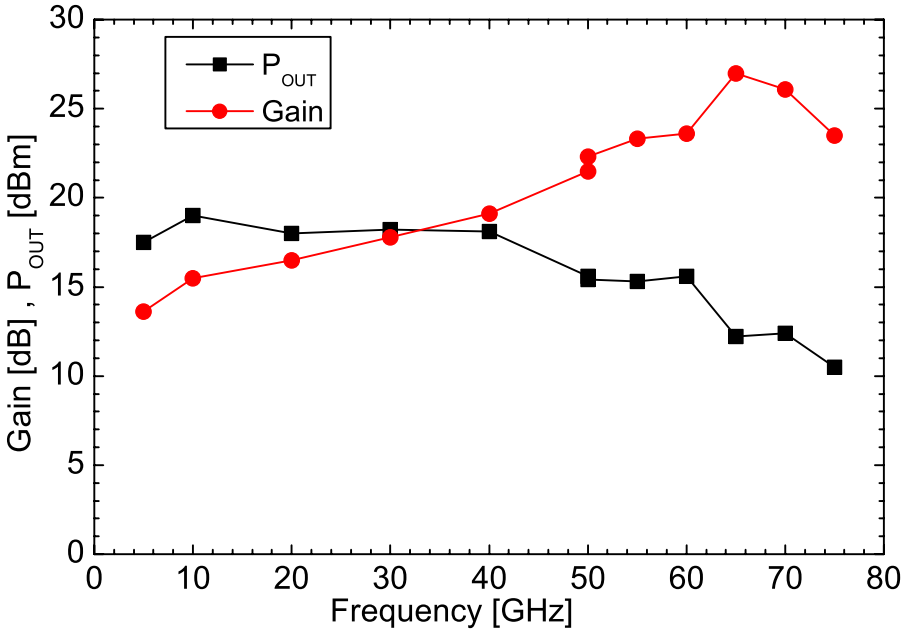


Figure 6.19: Output power at 1-dB compression point and gain of distributed amplifier optimized for maximum output power.

0 Hz to zero at 80 GHz. Hence, the probability is high that a voltage swing can be achieved, that is closer to the upper than to the lower boundary.

The tuning capability of the amplifier is a very useful feature of this amplifier concept. It cannot only be used to optimize the amplifier's performance, but also to compensate losses and deficiencies in other parts of the transmission system. These losses can result from packaging of chips, using bandwidth limited interconnections, or from poor performance of other circuits. These losses typically increase with frequency. How effectively gain peaking can compensate such limitations is demonstrated by the eye diagram measurement (Fig. 6.17).

In summary, the results show, that the gain cell of this amplifier (Fig. 6.14) and the design and layout concept are well suited for realizing modulator drivers at 80 GBit/s at possibly beyond. If this concept is transfered into a differential design voltage swings exceeding 5 V are conceivable.

6.3 Summary

This chapter described the development of modulator driver concepts for high bit rate optical data transmission systems. After a theoretical introduction into distributed amplifier concepts, a first version of such an amplifier demonstrated the usefulness of this concept by reaching 102 GHz 3-dB bandwidth and 10 dB gain. Although the power consumption and gain were comparable to the lumped amplifiers, the distributed amplifier already reached 45 % more bandwidth and 60 % higher output voltage swing.

Using additional attenuation compensation techniques, the second amplifier version further increased gain and output power, while eliminating some deficiencies of the first amplifier version. This amplifier shows excellent matching, very smooth gain, and good power behavior, which would make it suitable for sensitive applications.

Focusing on maximum output power performance, a third amplifier version was realized, which implements transistors of different sizes. This amplifier shows a clear eye diagram at 80 Gbit/s with a voltage swing of 2.4 V and 14.5 dB gain. The voltage swing was limited by the limited power of the digital input signal. As power measurements show voltage swings of 5 V up to 40 GHz with a reasonable roll off to higher frequencies, considerably larger eye diagrams than measured seem possible.

The tuning capability is a useful feature of the last two amplifier versions. It can be used to compensate for losses introduced by other parts of a transmission system. These losses typically increase with frequency and can result from chip packaging, or the use of bandwidth limited components. How effectively gain peaking can compensate such limitations is demonstrated by the eye measurement (Fig. 6.17).

In summary, the results show, that the design and layout concept of the third distributed amplifier version is well suited for realizing modulator drivers at 80 Gbit/s.

Chapter 7

Conclusion and Outlook

Conclusion

The goal of this thesis was the development of broadband amplifiers suitable for ETDM data transmission systems operating at data rates of 80 Gbit/s and beyond. Especially, the modulator driver posed a challenge, because it does not only have to provide a bandwidth of at least 80 GHz but also an output voltage swing of at least 3 V peak-to-peak. The goal has been successfully achieved by having conducted the following development steps:

First, the vertical dimensions of InP DHBTs, i.e., the epitaxial layer structure, was analyzed using dc and hf measurements and subsequently optimized within two iterations. Second, the lateral dimensions, i.e., the transistor layout, was systematically investigated. Critical transistor design issues were identified, experimentally confirmed, and the device layout was optimized. As a result, state-of-the-art transistor devices featuring high-frequency parameters f_{T} and f_{max} of more than 280 GHz and a breakdown voltage exceeding 4 V were achieved. At the same time, the optimized transistors show a high yield in the present technology, which allows the realization of complex circuits using more than 100 devices.

Third, from a variety of modeling approaches, the most suitable one was adopted. An existing preliminary model extraction procedure was refined and optimized with respect to the demands of the project by making a compromise between accuracy and complexity. This resulted in transistor models which were equally suited for analog as well as digital circuit design.

Up to this point, the achievements did not only enable the development of amplifiers, but they also significantly contributed to the successful development of digital circuits for 80 Gbit/s, like frequency divider, multiplexer, demultiplexer, and clock and data recovery circuit.

Fourth, several lumped amplifiers were realized and evaluated. The most suitable lumped amplifier achieved a bandwidth of 70 GHz and a gain of 10.5 dB. It uses only three transistors, has a power consumption of only 120 mW and can be used as amplifier at low power levels. However, it was also shown, that the maximum available output voltage swing of this amplifier is less than 1 V.

Finally, distributed amplifiers were realized and their small-signal response, their output power capabilities, and eye diagrams at 80 Gbit/s were measured. Subsequently, the amplifiers were optimized for modulator driver applications. During this process a distributed amplifier exhibiting 10 dB gain and a 3-dB bandwidth of more than 100 GHz was achieved. So far, this is the highest bandwidth, which has been reported for HBT broadband amplifiers. After optimization with respect to output power, another distributed amplifier was realized, which exhibits a clear output eye diagram at 80 Gbit/s with a gain of 14.5 dB and a voltage output swing of 2.4 V_{pp} limited by the available digital input signal. This result rivals the best reported results with respect to modulator driver applications at 80 Gbit/s. Additional power measurements revealed that the amplifier delivers an output power of 18 dBm (5.1 V_{pp}) at 40 GHz and 1-dB compression. From the power measurement results and taking into account the shape of the power density spectrum of a 80 Gbit/s random bit sequence, considerably larger eye diagrams than measured seem possible. In addition, the distributed amplifier offers a tunable gain peaking, which can be used to optimize circuit performance and to compensate losses in the circuit environment.

In summary, in this work amplifiers, which can be used as modulator drivers at 80 Gbit/s, have been developed from scratch within a period of less than 4 years.

Outlook

The results of this work will contribute to further development steps in the future. In the present transistor technology, also multiplexers for data rates at 80 Gbit/s have been realized. This should enable an integration of a multiplexer and a modulator driver on a single chip.

During the course of the InP project, all necessary electrical components for a 40 Gbit/s transmission system were also realized at Fraunhofer IAF. Systems at such data rates benefit from the outstanding performance of InP DHBT circuits with respect to low-frequency noise and output power.

Finally, the development of circuits for data rates exceeding 80 Gbit/s seems possible, as the current results were achieved in a comparatively short period. Hence, it is likely that continuous research will further exploit the potential of the InP DHBT technology.

Appendix A

Extracted Large-Signal Model Parameters

In this work transistors with an emitter length L_E of 4 and 8 μm were used for circuit design. For these two devices large-signal models were extracted according to the procedure described in Sec. 3.2. The models were used for simulation of amplifier circuits in the ADS software environment by Agilent. The corresponding AgilentHBT model parameters are listed in Table A.1.

Due to the large number of more than 50 model parameters the system of model equations is under-determined. Therefore, an infinite number of different model parameter sets can be extracted all featuring a comparable accuracy. As a consequence, the model parameters scale only within limits, when two devices are compared. The resistance parameters (starting with the letter "R") and some capacitance parameters seem to scale with transistor size. Other parameters, especially parameters describing the delay time, do not seem to scale.

L_E [μm]	4	8	L_E [μm]	4	8	L_E [μm]	4	8
$TNOM$	25	25	VJE	0.64	0.62	$VTR0$	1.38	1.18
RE	6.0	3.0	MJE	0.08	0.08	$VMX0$	162m	1.04
RCI	7.5	3.2	$CEMAX$	60f	100f	$VTCMININV$	1.42f	1.25f
RCX	10	4.9	$VPTE$	1.45	2.62	$VTRMIN$	786m	689m
RBI	12	7.1	$MJER$	0.17	0.01	$VMXMIN$	3.01	2.64
RBX	15	8.5	CJC	20.8f	31.0f	$VTCINV$	3.10μ	14.4μ
IS	3.37f	6.16f	VJC	0.5	0.5	$VTC2INV$	384m	26.4m
NF	1.20	1.18	MJC	0.6	0.6	$FEXTC$	12.7μ	999m
ISR	2.0f	2.0f	$CCMAX$	70f	100f	$TKRK$	209f	227f
NR	1.05	0.99	$VPTC$	1.0	1.0	$IKRK$	16.5m	29.0m
ISE	1.05p	1.21p	$MJCR$	0.1	0.1	$IKRKTR$	1.0μ	1.0μ
NE	1.91	1.85	$ABCX$	0.75	0.75	$VKRK$	617m	638m
ISC	1.78p	312f	TFB	38.2f	40.3f	$VKRK2INV$	1.94f	1.89f
NC	1.28	1.15	$FEXTB$	0	0	$GKRK$	1.59	1.58
ISA	830p	110p	$TFC0$	600f	248f	$VKTR$	142m	170m
NA	2.67	3.16	$TCMIN$	80.4f	72.0f	$VKMX$	525μ	2.94m
ISB	826f	36.2p	ITC	1.79m	0.92m	$FEXKE$	0	0
NB	3.46	2.27	$ITC2$	3.16m	9.03m	TR	1.0n	1.0n
CJE	13.7f	23.2f	$VTC0INV$	3.76	6.44			

Table A.1: Summary of extracted AgilentHBT-model parameters for two different devices with an emitter length L_E of 4 and 8 μm.

Bibliography

[1] B. Agarwal, Q. Lee, D. Mensa, R. Pullela, J. Guthrie, and M. J. W. Rodwell, "80-GHz distributed amplifiers with transferred-substrate heterojunction bipolar transistors," *IEEE Transactions on Microwave Theory and Techniques*, vol. 46, no. 12, pp. 2302–2306, Dec. 1998.

[2] B. Agarwal, A. E. Schmitz, J. J. Brown, M. Matloubian, M. G. Case, M. Le, M. Lui, and M. J. W. Rodwell, "112-GHz, 157-GHz, and 180-GHz InP HEMT traveling-wave amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 46, no. 12, pp. 2553–2559, Dec. 1998.

[3] G. P. Agrawal, *Nonlinear Fiber Optics*. Academic Press, 2001.

[4] Y. Arayashiki, Y. Ohkubo, Y. Amano, A. Takagi, M. Ejima, and Y. Matsuoka, "16 dB 80 GHz InGaP/GaAs HBT distributed amplifier," *Electronics Letters*, vol. 40, no. 4, pp. 244–245, 2004.

[5] Y. Ayasli, S. W. Miller, R. Mozzi, and L. K. Hanes, "Capacitively coupled traveling-wave power amplifier," *IEEE Transactions on Microwave Theory and Techniques*, vol. 32, no. 12, pp. 1704–1709, Dec. 1984.

[6] Y. Ayasli, L. D. Reynolds, R. L. Mozzi, and L. K. Hanes, "2-20-GHz GaAs traveling-wave power amplifier," *IEEE Transactions on Microwave Theory and Techniques*, vol. 32, no. 3, pp. 290–295, Mar. 1984.

[7] Y. Baeyens, R. Pullela, J. P. Mattia, H.-S. Tsai, and Y.-K. Chen, "A 74-GHz bandwidth InAlAs/InGaAs-InP HBT distributed amplifier with 13-dB gain," *IEEE Microwave and Guided Wave Letters*, vol. 9, no. 11, pp. 461–463, Nov. 1999.

[8] Y. Baeyens, G. Georgiou, J. S. Weiner, A. Leven, V. Houtsma, P. Paschke, Q. Lee, R. F. Knopf, Y. Yang, L. Chua, C. Chen, C. T. Liu, and Y.-K. Chen, "InP D-HBT IC's for 40 Gb/s and higher bitrate lightwave transceivers," *IEEE Journal of Solid-State Circuits*, vol. 37, no. 9, pp. 1152–1159, Sept. 2002.

[9] Y. Baeyens, N. Weimann, P. Roux, A. Leven, V. Houtsma, R. F. Knopf, Y. Yang, J. Frackoviak, A. Tate, J. S. Weiner, P. Paschke, and Y.-K. Chen, "High gain-bandwidth differential distributed InP D-HBT driver amplifiers with large ($11.3\,V_{pp}$)

output swing at 40 Gb/s," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 10, pp. 1697–1705, Oct. 2004.

[10] J. L. Benchimol, J. Mba, A. M. Duchenois, P. Berdaguer, B. Sermage, G. L. Roux, S. Blayac, M. Riet, J. Thuret, C. Gonzalez, and P. André, "Improvement of CBE grown InGaAs/InP HBTs using a corbon doped and compositionally graded base," in *Proc. International Conference on Indium Phosphide and Related Materials*, Davos, Switzerland, 1999, pp. 559–562.

[11] J. L. Benchimol, J. Mba, A. M. Duchenois, B. Sermage, P. Launay, D. Caffin, M. Meghelli, and M. Juhel, "CBE growth of carbon doped InGaAs/InP HBTs for 25 Gbit/s circuits," *Journal of Crystal Growth*, vol. 188, no. 1-4, pp. 349–354, June 1998.

[12] H. H. Berger, "Contact resistance on diffused resistors," in *IEEE International Solid-State Circuits Conference*, Pennsylvania, USA, 1969, pp. 160–161.

[13] Y. Betser, D. Ritter, G. Bahir, S. Cohen, and J. Sperling, "Measurement of the minority carrier mobility in the base of heterojunction bipolar transistors using a magnetotransport method," *Applied Physics Letters*, vol. 67, no. 13, pp. 1883–84, Sept. 1995.

[14] J. B. Beyer, S. N. Prasad, R. C. Becker, J. E. Nordman, and G. K. Hohenwarter, "MESFET distributed amplifier design guidelines," *IEEE Transactions on Microwave Theory and Techniques*, vol. 32, no. 3, pp. 268–275, Mar. 1984.

[15] S. Blayac, M. Riet, J. L. Benchimol, P. Berdaguer, and J. Godin, "Geometry effects on RF performances of InP DHBTs for ultrafast integrated circuits," in *Proc. 28th European Solid-State Device Research Conference*, Bordeaux, France, 1998, pp. 540–543.

[16] S. Bousnina, C. Falt, P. Mandeville, A. B. Kouki, and F. M. Ghannouchi, "An accurate on-wafer deembedding technique with application to HBT devices characterization," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 2, pp. 420–424, Feb. 2002.

[17] S. Bousnina, P. Mandeville, A. B. Kouki, R. Surridge, and F. M. Ghannouchi, "Direct parameter-extraction method for HBT small-signal model," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 2, pp. 529–536, Feb. 2002.

[18] L. H. Camnitz, S. Kofol, T. Low, and S. R. Bahl, "An accurate, large signal, high frequency model for GaAs HBTs," in *Proc of* 18^{th} *GaAs IC Symposium*, Orlando, FL, USA, 1996, pp. 303–306.

[19] C. E. Chang and P. M. Asbeck, "Novel HBT stucture for high f_t and low current density," in *Proc. International Electron Devices Meeting*, Washington, DC, USA, 1993, pp. 795–798.

[20] C. Chang, P. Asbeck, P. Zampardi, and K. C. Wang, "Direct measurement of C_{be} and C_{bc} versus voltage for small HBT's with microwave S-parameters for scaled Gummel-Poon BJT models," *IEEE Transactions on Microwave Theory and Techniques*, vol. 47, no. 1, pp. 108–110, Jan. 1999.

[21] M.-C. Chiang, S.-S. Lu, C.-C. Meng, S.-A. Yu, S.-C. Yang, and Y.-J. Chan, "Analysis, design, and optimization of InGaP-GaAs HBT matched-impedance wide-band amplifiers with multiple feedback loops," *IEEE Journal of Solid-State Circuits*, vol. 37, no. 6, pp. 406–409, June 2002.

[22] F.-T. Chien and Y.-J. Chan, "Bandwidth enhancement of transimpedance amplifier by a capacitive-peaking design," *IEEE Journal of Solid-State Circuits*, vol. 34, no. 8, pp. 1167–1170, Aug. 1999.

[23] D. Costa, W. U. Liu, and J. S. Harris, "Direct extraction of the AlGaAs/GaAs heterojunction bipolar transistor small-signal equivalent circuit," *IEEE Transactions on Electron Devices*, vol. 38, no. 9, pp. 2018–2024, Sept. 1991.

[24] N. Dagli, "Wide-bandwidth lasers and modulators for RF photonics," *IEEE Transactions on Microwave Theory and Techniques*, vol. 47, no. 7, pp. 1151–1171, July 1999.

[25] S. D'Agostino and C. Paoloni, "Comparison between MMIC matrix and distributed amplifiers," in *IEE Proceedings - Microwaves, Antennas and Propagation*, vol. 142, Feb. 1995, pp. 75–77.

[26] S. Deibele and J. B. Beyer, "Attenuation compensation in distributed amplifier design," *IEEE Transactions on Microwave Theory and Techniques*, vol. 37, no. 9, pp. 1425–1433, Sept. 1989.

[27] J. Driesen, S. Topaloglu, and F.-J. Tegude, "Optimizing lateral HBT design by utilizing performance estimations," in *11th International Conference on Indium Phosphide and Related Materials*, Davos, Switzerland, 2005.

[28] M. W. Dvorak, O. J. Pitts, S. P. Watkins, and C. R. Bolognesi, "Abrupt junction InP/GaAsSb/InP double heterojunction bipolar transistor with f_t as high as 250 GHz and $BV_{CE0} > 6$ V," in *Proc. International Electron Devices Meeting*, San Francisco, CA, USA, 2000, pp. 178–181.

[29] J. M. Early, "Effects of space-charge layer widening in junction transistors," *Proceedings of the IRE*, vol. 40, pp. 1401–1406, Nov 1952.

[30] T. Enoki, H. Yokoyama, Y. Umeda, and T. Otsuji, "Ultrahigh-speed integrated circuits using InP-based HEMTs," *Jpn. Journal of Applied Physics*, vol. 37, no. 3B, pp. 1359–1364, Mar. 1998, part 1.

[31] O. Esame, Y. Gurbuz, I. Tekin, and A. Bozkurt, "Performance comparison of state-of-the-art heterojunction bipolar devices (HBT) based on AlGaAs/GaAs, Si/SiGe and InGaAs/InP," *Microelectronics Journal*, vol. 35, no. 11, pp. 901–908, Nov. 2004.

[32] R. M. Fano, "Theoretical limitations on the broadband matching of arbitrary impedances," *Journal of the Franklin Institute*, vol. 249, pp. 57–83 and 139–154, Jan. and Feb. 1950.

[33] A. Feygenson, R. K. Montgomery, P. R. Smith, R. A. Hamm, M. Haner, R. D. Yadvish, and M. B. Panish, "InP/InGaAs composite collector heterostructure bipolar transistors and circuits," in *Proc. Indium Phosphide Related Materials (IPRM)*, Paris, France, 1993, pp. 572–575.

[34] A. Garlapati and S. Prasad, "A unified model for single/multifiger HBTs including self-heating effects," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 1, pp. 186–191, Jan. 2001.

[35] Y. Gobert, P. J. Tasker, and K. H. Bachem, "A physical, yet simple, small-signal equivalent circuit for the heterojunction bipolar transistor," *IEEE Transactions on Microwave Theory and Techniques*, vol. 45, no. 1, pp. 149–153, Jan. 1997.

[36] G. Gonzalez, *Microwave Transitor Amplifiers: Analysis and Design*. Prentice-Hall, Inc., 1997.

[37] P. R. Gray and R. G. Meyer, *Analog Integrated Circuits*. John Wiley & Sons, Inc., 1984.

[38] H. K. Gummel and H. C. Poon, "An integral charge control model of bipolar transistors," *Bell Syst. Tech. Journ.*, pp. 827–852, May-June 1970.

[39] W. Hafez and M. Feng, "Lateral scaling of 0.25 μm InP/InGaAs SHBTs with InAs emitter cap," *Electronics Letters*, vol. 40, no. 18, pp. 1151–1153, Sept. 2004.

[40] M. Hafizi, T. Liu, P. A. Macdonald, M. Lui, P. Chu, D. B. Rensch, W. E. Stanchina, and C. S. Wu, "High-performance microwave power AlInAs/GaInAs/InP double heterojunction bipolar transistors with compositionally graded base-collector junction," in *Proc. International Electron Devices Meeting*, Washington, DC, USA, 1993, pp. 791–794.

[41] M. Hafizi, T. Liu, W. E. Stanchina, D. B. Rensch, M. Lui, and Y. K. Brown, "Importance of collector doping in the design of AlInAs/GaInAs/InP double heterojunction bipolar transistors," *Applied Physics Letters*, vol. 64, no. 24, pp. 3261–3263, June 1994.

[42] R. Hoffmann, *Integrierte Mirkrowellenschaltungen*. Berlin: Springer Verlag, 1983.

[43] A. Huber, D. Huber, C. Bergamaschi, T. Morf, and H. Jäckel, "Design and characterisation of a 50 GHz InP/InGaAs HBT amplifier," in *11th International Conference on Indium Phosphide and Related Materials*, Davos, Switzerland, 1999.

[44] C. D. Hull and R. G. Meyer, "Principles of monolithic wideband feedback amplifier design," *International Journal of High Speed Electronics*, vol. 3, no. 1, pp. 53–93, 1992.

[45] M. Ida, K. Kurishima, H. Nakajima, N. Watanabe, and S. Yamahata, "Undoped-emitter InP/InGaAs HBTs for high-speed and low-power applications," in *Proc. International Electron Devices Meeting*, San Francisco, CA, USA, 2000, pp. 854–856.

[46] T. Ishibashi, "Nonequilibrium electron transport in HBTs," *IEEE Transactions on Electron Devices*, vol. 48, no. 11, pp. 2595–2605, Nov. 2001.

[47] M. Iwamoto, D. E. Root, J. B. Scott, A. Cognata, P. M. Asbeck, B. Hughes, and D. C. D'Avanzo, "Large-signal HBT model with improved collector transit time formulation for GaAs and InP technologies," in *IEEE MTT-S Digest*, 2003, pp. 635–638.

[48] B. Jalali and S. J. Pearton, Eds., *InP HBTs: Growth, Processing, and Applications*. Artech House, 1995.

[49] S. J. Jeng, D. R. Greenberg, M. Longstreet, G. Hueckel, D. L. Harame, and D. Jadus, "Lateral scaling of the self-aligned extrinsic base in SiGe HBTs," in *Bipolar/BiCMOS Circuits and Technology Meeting*, Minneapolis, MN, USA, 1996, pp. 15–18.

[50] E. O. Johnson, "A simple physical comparison of amplifying devices of the emitter-control-collector variety: A survey," in *IEEE International Solid-State Circuit Conference Digest*, vol. 2, Feb. 1959, pp. 60–62.

[51] T. Kim and K. Yang, "A new large-signal InP/InGaAs single HBT model including self-heating and impact ionisation effects," in *IEEE MTT-S Digest*, 2002, pp. 2141–2144.

[52] S. Kimura, Y. Imai, Y. Umeda, and T. Enoki, "0-90 GHz InAlAs/InGaAs/InP HEMT distributed baseband amplifier IC," *Electronics Letters*, vol. 31, no. 17, pp. 1430–1431, Aug. 1995.

[53] ——, "Loss-compensated distributed baseband amplifier IC's for optical transmission systems," *IEEE Transactions on Microwave Theory and Techniques*, vol. 44, no. 10, pp. 1688–1693, Oct. 1996.

[54] I. Kipnis, J. F. Kukielka, J. Wholey, and C. P. Snapp, "Silicon bipolar fixed and variable gain amplifier MMICs for microwave and lightwave applications up to 6 GHz," in *IEEE MTT-S Int. Microwave Symp. Dig.*, vol. 1, June 1989, pp. 109–112.

[55] C. T. Kirk, "A theory of transistor cutoff frequency (f_t) falloff at high current densities," *IRE Trans. Electron Dev.*, vol. ED-9, pp. 164–174, Mar. 1962.

[56] W. Ko and Y. Kwon, "High-gain direct-coupled matrix distributed amplifier using active feedback topology," in *IEEE Compound Semiconductor Integrated Circuit Symposium Digest*, Oct. 2004, pp. 309–312.

[57] K. W. Kobayashi, J. Cowles, L. T. Tran, A. Gutierrez-Aitken, T. R. Block, A. K. Oki, and D. C. Streit, "A 50-MHz-55-GHz multidecade InP-based HBT distributed amplifier," *IEEE Microwave and Guided Wave Letters*, vol. 7, no. 10, pp. 353–355, Oct. 1997.

[58] K. W. Kobayashi, R. Esfandiari, and A. K. Oki, "A novel HBT distributed amplifier design topology based on attenuation compensation techniques," in *IEEE MTT-S International Microwave Symposium Digest*, 1994, pp. 447–450.

[59] K. W. Kobayashi, J. Cowles, L. Tran, T. Block, A. K. Oki, and D. C. Streit, "A 2-32 GHz coplanar waveguide InAlAs/InGaAs-InP HBT cascode distributed amplifier," in *IEEE MTT-S International Microwave Symposium Digest*, Orlando, FL, USA, Oct. 1995, pp. 215–218.

[60] K. W. Kobayashi, L. T. Tran, J. C. Cowles, T. R. Block, A. K. Oki, and D. C. Streit, "Extending the bandwidth performance of heterojunction bipolar transisor-based distributed amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 44, no. 5, pp. 739–748, May 1996.

[61] K. W. Kobayashi, R. Esfandiari, M. E. Hafizi, D. C. Streit, A. K. Oki, L. T. Tran, D. K. Umemoto, and M. E. Kim, "GaAs HBT wideband matrix distributed and darlington feedback amplifiers to 24 GHz," *IEEE Transactions on Microwave Theory and Techniques*, vol. 39, no. 12, pp. 2001–2009, Dec. 1991.

[62] K. Krishnamurthy, J. Chow, , M. J. W. Rodwell, and R. Pullela, "0.1-42 GHz InP DHBT distributed amplifiers with 35 dB gain and 15 dBm output," *Electronics Letters*, vol. 39, no. 22, pp. 1594–1595, 2003.

[63] H. Kroemer, "Theory of a wide-gap emitter for transistors," *Proceedings of the IRE*, vol. 45, no. 11, pp. 1535–1537, Nov 1957.

[64] ——, "Heterostructure bipolar transistors and integrated circuits," *Proceedings of the IEEE*, vol. 70, no. 1, pp. 13–25, Jan. 1982.

[65] K. Kurishima, M. Ida, N. Watanabe, and T. Enoki, "Device technology for high-yield and high-performance InP/InGaAs DHBTs," in *Proc. Int. Symp. Compound Semiconductors*, Tokyo, Japan, 2001, pp. 51–56.

[66] K. Kurishima, H. Nakajima, S. Yamahata, T. Kobayashi, and Y. Matsuoka, "Effects of a compositionally-graded $In_xGa_{1-x}As$ base in an abrupt-emitter InP/InGaAs heterojunction bipolar transistor," *Jpn. Journal of Applied Physics*, vol. 34, pp. 1221–1227, Feb. 1995.

[67] K. Kurishima, "An analytic expression of $f_{\max}$ for HBT's," *IEEE Transactions on Electron Devices*, vol. 43, no. 12, pp. 2074–2079, Dec. 1996.

[68] W. Liu, *Handbook of III-V Heterojunction Bipolar Transistors*. John Wiley & Sons, Inc., 1998.

[69] J. M. López-González and L. Prat, "Numerical modelling of abrupt InP/InGaAs HBTs," *Solid-State Electronics*, vol. 39, no. 4, pp. 523–527, 1996.

[70] S. A. Maas and D. Tait, "Parameter-extraction method for heterojunction bipolar transistors," *IEEE Microwave and Guided Wave Letters*, vol. 2, no. 12, pp. 502–504, Dec. 1992.

[71] F. J. MacWilliams and N. J. A. Sloane, "Pseudo-random sequences and arrays," *Proceedings of the IEEE*, vol. 64, no. 12, pp. 1715–1730, Dec. 1976.

[72] R. E. Makon, R. Driad, K. Schneider, M. Ludwig, R. Aidam, R. Quay, M. Schlechtweg, and G. Weimann, "80 Gbit/s monolithically integrated clock and data recovery circuit with 1:2 DEMUX using InP-based DBHTs," in *Proc. 2005 Compound Semiconductor IC Symp.*, Palm Springs, CA, USA, Oct. 2005, pp. 268–271.

[73] R. E. Makon, R. Driad, K. Schneider, H. Maßler, R. Aidam, R. Quay, M. Schlechtweg, and G. Weimann, "A fundamental low phase noise InP-based DHBT VCO operating up to 89 GHz," *Electronics Letters*, vol. 41, no. 17, pp. 37–38, Aug. 2005.

[74] R. E. Makon, M. Lang, R. Driad, K. Schneider, M. Ludwig, R. Aidam, R. Quay, M. Schlechtweg, and G. Weimann, "Over 80 Gbit/s 2:1 multiplexer and low power selector ICs using InP/InGaAs DHBTs," *Electronics Letters*, vol. 41, no. 11, pp. 644–646, May 2005.

[75] S. Masuda, T. Takahashi, and K. Joshin, "An over-100-GHz InP HEMT flip-chip distributed baseband amplifier with inverted microstrip line structure for optical transmission system," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 9, pp. 1479–1484, Sept. 2003.

[76] G. L. Matthaei, L. Young, and E. M. T. Jones, *Microwave Filters, Impedance-matching Networks, and Coupling Structures*. Artech House, 1980.

[77] C. Meliani, G. Post, G. Rondeau, J. Decobert, W. Mouzannar, E. Dutisseuil, and R. Lefèvre, "DC-92 GHz ultra-broadband high gain InP HEMT amplifier with 410

GHz gain-bandwidth product," *Electronics Letters*, vol. 38, no. 20, pp. 1175–1177, Sept. 2002.

[78] R. G. Meyer and R. A. Blauschild, "A 4-terminal wide-band monolithic amplifier," *IEEE Journal of Solid-State Circuits*, vol. SC-16, no. 6, pp. 634–638, Dec. 1981.

[79] C. Nguyen and M. Micovic, "The state-of-the-art of GaAs and InP power devices and amplifiers," *IEEE Transactions on Electron Devices*, vol. 48, no. 3, pp. 472–478, Mar. 2001.

[80] T. Otsuji, K. Murata, T. Enoki, and Y. Umeda, "An 80-Gbit/s multiplexer IC using InAlAs/InGaAs/InP HEMT's," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 9, pp. 1321–1327, Sept. 1998.

[81] S. H. Park, I.-H. Choi, S. W. Choi, M. P. Park, Y. S. Kim, J.-J. Lee, C. S. Park, and H.-M. Park, "Emitter-base geometry dependence of electrical performance of AlGaAs/GaAs HBT," *Journal of the Korean Institute of Telematics and Electronics, South Korea*, vol. 32A, no. 2, pp. 57–65, Feb. 1995.

[82] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipies in Pascal*. Cambridge University Press, 1989, pp. 326–330.

[83] J. G. Proakis, *Digital Communications*. McGraw-Hill, 2000.

[84] J. Pusl, B. Agarwal, R. Pullela, L. D. Nguyen, M. V. Le, M. J. Rodwell, L. Larson, J. Jensen, R. Y. Yu, and M. G. Case, "Capacitive-devision traveling-wave amplifier with 340 GHz gain-bandwidth product," in *IEEE MTT-S Digest*, 1995, pp. 1661–1664.

[85] R. Quay, "Analysis and simulation of high electron mobility transistors," Ph.D. dissertation, Technische Universität Wien, Fakultät für Elektrotechnik und Informationstechnik, 2001.

[86] B. Razavi, *Design of Integrated Circuits for Optical Communications*. McGraw-Hill, 2003.

[87] H.-M. Rein and M. Möller, "Design considerations for very-high-speed Si-bipolar IC's operating up to 50 Gb/s," *IEEE Journal of Solid-State Circuits*, vol. 31, no. 8, pp. 1076–1090, Aug. 1996.

[88] J. S. Rieh, D. Greenberg, M. Khater, K. T. Schonenberg, S.-J. Jeng, F. Pagette, T. Adam, A. Chinthakindi, *et al.*, "SiGe HBTs for millimeter-wave applications with simultaneously optimized f_t and f_{max} of 300 GHz," in *Radio Frequency Integrated Circuits (RFIC) Symposium*, June 2004, pp. 395–398.

[89] M. J. W. Rodwell, M. Urteaga, T. Mathew, D. Scott, D. Mensa, Q. Lee, J. Guthrie, Y. Betser, *et al.*, "Submicron scaling of HBTs," *IEEE Transactions on Electron Devices*, vol. 48, no. 11, pp. 2606–2624, Nov. 2001.

[90] M. Rudolph, R. Doerner, K. Beilenhoff, and P. Heymann, "Scalable GaInP/GaAs HBT large-signal model," *IEEE Transactions on Microwave Theory and Techniques*, vol. 48, no. 12, pp. 2370–2376, Dec. 2000.

[91] A. Samelis and D. Pavlidis, "DC to high-frequency HBT-model parameter evaluation using impedance block conditioned optimization," *IEEE Transactions on Microwave Theory and Techniques*, vol. 45, no. 6, pp. 886–897, June 1997.

[92] U. Schaper and B. Holzapfl, "Analytical parameter extraction of the HBT equivalent circuit with T-like topology from measured S-parameters," *IEEE Transactions on Microwave Theory and Techniques*, vol. 43, no. 3, pp. 493–498, Mar. 1995.

[93] M. Schlechtweg, A. Leuther, A. Tessmann, C. Schwörer, H. Massler, W. Reinert, M. Lang, U. Nowotny, O. Kappeler, M. Walther, and R. Lösch, "Millimeter-wave and mixed-signal integrated circuits based on advanced metamorphic HEMT technology," in *Proc. Indium Phosphide Related Materials (IPRM)*, 2004, pp. 609–614.

[94] K. Schneider, R. Driad, R. E. Makon, H. Maßler, M. Ludwig, R. Quay, M. Schlechtweg, and G. Weimann, "Comparison of InP/InGaAs DHBT distributed amplifiers as modulator drivers for 80 Gbit/s operation," *IEEE Transactions on Microwave Theory and Techniques*, vol. 52, no. 11, pp. 3378–3387, Nov. 2005.

[95] K. Schneider, R. Driad, R. E. Makon, A. Tessmann, R. Aidam, R. Quay, M. Schlechtweg, and G. Weimann, "InP/InGaAs-DHBT distributed amplifier MMICs exceeding 80 GHz bandwidth," in *2005 IEEE MTT-S Int. Microwave Symp. Dig.*, Long Beach, CA, USA, June 2005, pp. 1591–1593.

[96] F. X. Sinnesbichler and G. R. Olbrich, "Accurate large-signal modeling of SiGe HBTs," in *Proc. 2000 IEEE MTT-S Int. Microwave Symp. Dig.*, 2000, pp. 749–752.

[97] M. Sotoodeh, L. Sozzi, A. Vinay, H. A. Khalid, Z. Hu, A. A. Rezazadeh, and R. Menozzi, "Stepping toward standard methods of small-signal parameter extraction for HBT's," *IEEE Transactions on Electron Devices*, vol. 47, no. 6, pp. 1139–1151, June 2000.

[98] T. Suzuki, Y. Nakasha, T. Takahashi, K. Makiyama, T. Hirose, and M. Takikawa, "144-Gbit/s selsctor and 100-Gbit/s 4:1 multiplexer," in *IEEE MTT-S Int. Microwave Symp. Dig.*, 2004, pp. 117–120.

[99] Y. Suzuki, Z. Yamazaki, and H. Hida, "An 80-Gb/s 2.7-V_{p-p} driver IC based on functional distributed circuits for optical transmission systems," in *2005 IEEE MTT-S Int. Microwave Symp. Dig.*, Long Beach, CA, USA, June 2005, pp. 325–328.

[100] D. A. Teeter and W. R. Curtice, "Comparison of hybrid Pi and Tee HBT circuit topologies and their relationship to large signal modeling," in *IEEE MTT-S Digest*, 1997, pp. 375–378.

[101] UCSD HBT model. University of California, San Diego, CA, USA. [Online]. Available: http://hbt.ucsd.edu

[102] I. Vurgaftman and J. R. Meyer, "Band parameters for III-V compound semiconductors and their alloys," *Journal of Applied Physics*, vol. 89, no. 11, pp. 5815–5875, June 2001.

[103] C. Wang and Y. Hsin, "DC and AC characteristics of HBTs with different base width," in *11th International Conference on Indium Phosphide and Related Materials*, Davos, Switzerland, 2005.

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

[105] C.-J. Wei and J. C. M. Hwang, "Direct extraction of equivalent circuit parameters for heterojunction bipolar transistors," *IEEE Transactions on Microwave Theory and Techniques*, vol. 43, no. 9, pp. 2035–2039, Sept. 1995.

[106] R. J. Whittier and D. A. Tremere, "Current gain and cutoff frequency falloff at high currents," *IEEE Trans. Electron Dev.*, vol. ED-16, pp. 39–57, Jan. 1969.

[107] T. Wong, A. Freundorfer, B. Beggs, and J. Sitch, "A 10 Gb/s AlGaAs/GaAs HBT high power fully-differential limiting distributed amplifier for III-V mach-zehnder modulator," *IEEE Journal of Solid-State Circuits*, vol. 31, no. 10, pp. 1388–1393, Oct. 1996.

[108] T. T. Y. Wong, *Fundamentals of Distributed Amplification.* Artech House, 1993.

[109] E. F. Yu, D. G. Hill, L. Zhang, and O. L. Hartin, "The impact of emitter fingers layout and geometry on InGaP HBT thermal resistance," in *Reliability of Compound Semiconductors Workshop*, Monterey, CA, USA, 2004, pp. 21–31.

[110] M. Yu, Z. Lao, S. Lee, V. Radisic, M. Xu, V. Ho, K. Guin, and K. C. Wang, "The development of 40 Gb/s limiting-distributed modulator drivers in InP HBTs," in *IEEE MTT-S Digest*, 2003, pp. 71–73.

[111] Y. Yu, R. Lewen, S. Irmscher, U. Westergren, L. Thylen, U. Eriksson, and W. S. Lee, "80 Gb/s ETDM transmitter with a traveling-wave electroabsorption modulator," in *Optical Fiber Communication Conference Dig.*, vol. 3, Mar. 2005, pp. 85–87.

[112] S. Zhang, G. Niu, J. D. Cressler, A. J. Joseph, G. Freeman, and D. L. Harame, "The effects of geometrical scaling on the RF performance of SiGe HBTs," in *2000 Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems*, Garmisch, Germany, 2000, pp. 48–51.

List of Abbreviations

ADS	Advanced Design System by Agilent Technologies
BCB	Benzocyclobutene
BER	Bit Error Rate
BJT	Bipolar Junction Transistor
BW	Bandwidth
CDR	Clock and Data Recovery Circuit
CPW	Coplanar Waveguide
DA	Distributed Amplifier
DC	Direct Current
DHBT	Double Heterojunction Bipolar Transistor
DUT	Device under Test
EA	Electroabsorbtion
ETDM	Electrical Time Division Multiplex
FET	Field Effect Transistor
FOM	Figure of Merit
HBT	Heterojunction Bipolar Transistor
HEMT	High Electron Mobility Transistor
HF	High Frequency
IAF	Fraunhofer Institute of Applied Solid-State Physics
IMSL	Inverted Microstrip Line
ISI	Intersymbol Interference
MBE	Molecular Beam Epitaxy
MMIC	Monolithic Microwave Integrated Circuit
MOCVD	Metal Organic Chemical Vapor Deposition
MIC	Microwave Integrated Circuit
MIM	Metal-Insulator-Metal
MSL	Microstrip Line
MUX	Multiplexer
MZ	Mach-Zehnder
NRZ	Non-Return to Zero
PRBS	Pseudorandom Bit Sequence

RC	Resistor and Capacitor
RF	Radio Frequency
SCFL	Source Coupled FET Logic
SCR	Space Charge Region
SEM	Scanning Electron Microscope
SHBT	Single Heterojunction Bipolar Transistor
TIA	Transimpedance Amplifier
UCSD	University of California, San Diego
VCO	Voltage Controlled Oscillator

Acknowledgments

This thesis is a result of my work as a scientist at the Fraunhofer Institute of Applied Solid-State Physics (IAF) in Freiburg. The research was funded by the German Federal Ministry of Education and Research.

Most of all, I would like to express my genuine gratitude to my adviser Prof. Dr. G. Weimann, Head of IAF. His guidance, continuous support, and confidence in my capabilities enabled the successful realization of this thesis. Moreover, I thank Prof. Dr. W. Wiesbeck, Head of the Institute of High Frequency Technology and Electronics at the University of Karlsruhe, for his interest in my work and for taking the position of second reviewer.

I would like to thank Dr. M. Schlechtweg, Head of the High Frequency Devices and Circuits Department of IAF, for providing an excellent working environment and for giving me the freedom to realize my ideas. Also, I thank Dr. R. Driad, who led the development of the InP-technology, for invaluable discussions and support. I thank Dr. R. Quay and Dr. J. Rosenzweig for their interest in my work and for helpful advices.

I could not have gotten through this work without the support generously offered by my colleagues in the High Frequency Devices and Circuits Department. In particular, discussions with A. Tessmann, C. Schwörer, R. E. Makon, M. Lang, Dr. F. van Raay and Dr. U. Nowotny solved many problems and gave me new ideas. I benefited from Dr. M. Seelmann-Eggebert's expertise in modeling. H. Maßler, M. Kuri, M. Ludwig, and M. Rießle were very helpful during measuring and testing of devices and circuits.

In addition, I would like to thank Dr. R. Aidam, who was responsible for the growth of epitaxial structures, and all other members of the technology department, who contributed to the successful realization of devices and circuits.

Furthermore, I would like to thank Prof. Peter M. Asbeck and James Chingwei Li from the University of California, San Diego, for their warm welcome. They were very kind and supportive during my visit to their research lab and I learned a great deal from them.

Finally, I would like to thank my family and my friends, especially A. Bullerkotte and U. Schmidt, for cheering me up during the hardest periods of my work.

Curriculum Vitae

Karl Schneider

25. November 1973	Born in Hannover, Germany
1980 – 1984	Primary school (Grundschule Grimselweg, Hannover)
1984 – 1986	Intermediate school (Orientierungsstufe Lüerstraße, Hannover)
1986 – 1993	Secondary school (Kaiser-Wilhelms-Gymnasium, Hannover)
1993	Abitur (university-entrance diploma)
1993 – 1994	Civilian service
1994 – 1997	Studies in Electrical Engineering at the University of Karlsruhe
1997 – 1998	Studies in Physics and English at the Flinders University of South Australia, Adelaide, Australia
1998 – 2001	Continuation of initial studies at the University of Karlsruhe
2001	Dipl.-Ing. degree in Electrical Engineering
2001 – 2002	Visiting scientist at the Massachusetts Institute of Technology, Cambridge, USA
since 2002	Scientist at the Fraunhofer Institute of Applied Solid-State Physics, Freiburg

ISBN 3-86644-021-9

www.uvka.de