An InP-Based Optoelectronic Integrated Circuit for Optical Communication Systems

Research Thesis

Submitted in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Electrical Engineering

Shraga Kraus

Submitted to the Senate of the Technion - Israel Institute of Technology

Iyar 5766 Haifa May 2006

The research thesis was done under the supervision of Prof. Dan Ritter in the Faculty of Electrical Engineering and Microelectronics Research Center

Acknowledgements

I would like to thank many people who helped the research and made it happen:

1. Thanks to the Technion, the Faculty of Electrical Engineering, and the Microelectronics Research Center in which the research was carried out and for putting at my disposal their best facilities and technologies for the research.

2. Thanks to Prof. Dan Ritter for his dedicated supervision and for the support, the help, and the inspiration along the way.

3. Thanks to Prof. Gadi Eisenstein and Prof. Meir Orenstein for their help with the measurements and for putting at my disposal their laboratories and equipment.

4. Thanks to Mr. Ouri Karni and Yohay Swirski for the establishment of the optoelectronic measurement setup, guided by Prof. Gadi Eisenstein, and for the tight support during taking measurements.

5. Thanks to my colleagues to Prof. Dan Ritter's research group: Arkadi Gavrilov, Shimon Cohen, Doron Cohen-Elias, Tsufit Magrisso, and Gennadi Burdo, for thousands of invaluable small assistances.

6. Thanks to the Israeli Ministry of Industry, Trade, and Labor for funding the project under which the research was done.

7. Thanks to my parents, Leah and Jehoshua Kraus, and my parents in-law, Sarah and Elimelech Westreich, for the indispensable help and support along the way.

8. Finally, special thanks to my wife, Naama, and my children — Mattan, Noa, and Yuval — for their patience and solidarity and being a great family, that gave me vitality to do the research and submit it.

Table of Contents

Abstract 1

List of Acronyms 3

List of Symbols 5

1 Introduction 13
- 1.1 Heterojunction Bipolar Transistor . . . 14
- 1.2 Photodiode . . . 15
- 1.3 Monolithic Receivers . . . 16
- 1.4 Lumped Monolithic Receivers . . . 18
- 1.5 Distributed Monolithic Receivers . . . 19
- References for Chapter 1 . . . 20

2 Heterojunction Bipolar Transistor 23
- 2.1 HBT Physical Characteristics and Phenomena . . . 24
  - 2.1.1 DC Characteristics and Phenomena . . . 24
  - 2.1.2 AC Characteristics and Phenomena . . . 31
- 2.2 HBT Fabrication and Structure . . . 34
  - 2.2.1 Fabrication Process . . . 34
  - 2.2.2 Transistor Geometric Structure . . . 41
- 2.3 HBT Measurements and Characterization . . . 46
  - 2.3.1 DC Measurements . . . 46
  - 2.3.2 DC Characterization . . . 49
  - 2.3.3 Voltage–Capacitance Measurement and Characterization . . . 56
  - 2.3.4 RF Measurements . . . 58
- 2.4 Extraction of HBT Small Signal Model . . . 61
  - 2.4.1 Extraction and Peeling of Parasitic Pad Elements . . . 61
  - 2.4.2 Extraction of Collector Elements . . . 63
  - 2.4.3 Extraction of Emitter Elements . . . 64
  - 2.4.4 Extraction of Base Resistance . . . 64
  - 2.4.5 Extraction of Forward Transit Time . . . 66

2.4.6 Extraction of Current Gain Cutoff Frequency 66
2.4.7 Extraction of Mason's Power Gain Cutoff Frequency 67
2.4.8 Actual Extracted Parameters 68
2.5 Construction of HBT VBIC Model 70
2.5.1 Basic Ebers–Moll Model 70
2.5.2 Basic Gummel–Poon Model 72
2.5.3 VBIC Parameters for Constructing Simple Model 72
2.5.4 VBIC Model Validation 74
2.6 Noise in HBT 77
2.6.1 Types of noise 77
2.6.2 Noise Mechanisms in HBT 79
References for Chapter 2 82

3 Photodiode **85**
3.1 Photodiode Physical Characteristics 86
3.1.1 DC Characteristics 86
3.1.2 AC Characteristics 88
3.2 Photodiode Fabrication and Structure 91
3.2.1 Fabrication Process 91
3.2.2 Photodiode Geometric Structure 95
3.3 Photodiode Measurements and Characterization 98
3.3.1 Measurement Equipment and Setup 98
3.3.2 Responsivity Measurement 99
3.3.3 Measurement and Processing 99
3.4 Noise in Photodiode 103
3.4.1 Noise Mechanisms in Photodiode 103
3.4.2 Relations Between Noise and Detected Signal 103
References for Chapter 3 105

4 Optoelectronic Lumped Circuit **107**
4.1 Amplifier Building Blocks 108
4.1.1 Bipolar Single-Stage Amplifier 109
4.1.2 Feedback 116
4.1.3 Noise in Single- and Multiple-Stage Amplifiers 118
4.1.4 Summary 120
4.2 Transimpedance Architecture 122
4.2.1 Operational Amplifier Design 122
4.2.2 Feedback and Stability 122
4.2.3 Input Stage 125
4.2.4 Biasing 126
4.2.5 Noise Performance 127
4.2.6 Power Supply Rejection 130

4.2.7 Layout . . . 133
4.3 Optoelectronic Integration . . . 135
4.3.1 Photodiode Dimensions and Optoelectronic Bandwidth . . . 135
4.3.2 Layout Integration . . . 135
4.3.3 Optoelectronic Noise Performance . . . 135
4.4 Circuit Measurements and Characterization . . . 138
4.4.1 TIA Measurements and Characterization . . . 138
4.4.2 OEIC Measurements and Characterization . . . 138
4.5 Circuit Performance . . . 140
4.5.1 TIA Performance . . . 140
4.5.2 OEIC Performance . . . 140
References for Chapter 4 . . . 143

5 Optoelectronic Distributed Circuit 145
5.1 Distributed Amplifier Fundamentals . . . 146
5.1.1 Transmission Lines . . . 146
5.1.2 Distributed Amplifier . . . 151
5.2 Distributed Amplifier Architecture . . . 154
5.2.1 Amplification Stage . . . 154
5.2.2 Input and Output Characteristic Impedances . . . 155
5.2.3 Number of Stages . . . 156
5.2.4 Biasing . . . 156
5.2.5 Noise Performance . . . 156
5.2.6 Group Delay . . . 157
5.2.7 Layout . . . 158
5.3 Optoelectronic Integration . . . 160
5.3.1 Photodiode Connection . . . 160
5.3.2 Optoelectronic Noise Performance . . . 161
5.4 Circuit Performance . . . 162
5.4.1 Measurements . . . 162
5.4.2 TWA Performance . . . 162
5.4.3 OEIC Performance . . . 162
References for Chapter 5 . . . 166

6 Summary and Conclusions 167
6.1 Photodiode . . . 168
6.2 Lumped Circuits . . . 169
6.3 Distributed Circuits . . . 170
References for Chapter 6 . . . 171

A A Control GUI for C–V Measurements **173**
A.1 Introduction 173
A.2 MATLAB Figures 174
A.3 MATLAB Functions 175

B Miller Effect **193**
B.1 Summary of Miller Effect 193
References for Appendix B 195

List of Figures

1.1 Scheme of a discrete component receiver with wire-bonds 16
1.2 Frequency response of a photoreceiver with 0.55 mm and 0.70 mm wire-bonds 17

2.1 One-dimensional BJT (a) schematic structure and (b) route of electrons in forward active mode 25
2.2 Charge carrier densities vs. position along BJT in forward active mode. . . 26
2.3 Electrical currents marked on a schematic BJT symbol 27
2.4 Energy band diagrams of (a) BJT in equilibrium and (b) in forward active mode, and (c) HBT in equilibrium and (d) in forward active mode 28
2.5 Energy band diagrams of (a) base-graded HBT in equilibrium and (b) in forward active mode 29
2.6 One-dimensional BJT schematic structure with subcollector 29
2.7 One-dimensional HBT (a) schematic structure and (b) small signal elements 32
2.8 Layer structure on wafer as grown by MOMBE system 34
2.9 Wafer layers after emitter metal deposition and liftoff 36
2.10 Wafer layers after emitter etch 37
2.11 Wafer layers after base metal mask 37
2.12 Wafer layers after base etch 38
2.13 Wafer layers after collector etch 38
2.14 Wafer layers after collector metal implementation 39
2.15 Wafer layers after subcollector etch 39
2.16 Wafer layers after transistor passivation 40
2.17 Metal 1 as connected to transistor contacts 41
2.18 Interconnection and capacitor sequence: (a) “metal 1” mask (b) “capacitor” mask (c) “crossover” mask (d) “metal 2” mask 41
2.19 SEM images of a transistor at various fabrication process steps: (a) emitter etch (b) base metal deposition and liftoff (c) emitter protect (d) collector protect (e) isolation (f) emitter expose 44

2.20 SEM images of a transistor and circuit at various fabrication process steps: (a) base–collector via (b) metal 1 deposition and liftoff (c) completed capacitor (d) completed interconnects with crossovers and a resistor 45
2.21 DC measurement setup: (a) measurement system (b) handmade probes . 46
2.22 I–V curve of a PN–junction 47
2.23 Common emitter measurement curves 48
2.24 Common base measurement curves 49
2.25 Gummel plot curves 50
2.26 Extraction of I_0 and n from I–V curve of a PN–junction 51
2.27 β_F and β_{AC} curves — calculated from Gummel plot data 52
2.28 V_{TO} marked on common emitter measurement curves 52
2.29 V_{CEO} marked on (a) α_F and (b) I_B curves 53
2.30 TLM test fixture 54
2.31 Current route in TLM measurement 55
2.32 Extraction of sheet resistance, contact resistance, and transfer length from a TLM curve 55
2.33 C–V measurement system 56
2.34 (a) C–V curve and (b) admittance ratio 57
2.35 Doping profile of a one-sided PN–junction 58
2.36 RF measurement system 59
2.37 Transistor as a two-port network in common emitter configuration 59
2.38 Small signal T model of a HBT 61
2.39 A HBT with pad parasitic elements lumped in a two-port network 62
2.40 Layout of (a) HBT probing pads (b) open test fixture (c) through test fixture 62
2.41 Extraction of C_{BC} 64
2.42 Extraction of (a) L_e (b) $r_e + r_{ee}$ (c) r_b 65
2.43 Extraction of τ_D 66
2.44 Extraction of f_T 67
2.45 Extraction of f_{MAX} 68
2.46 Ebers–Moll model (injection version) 70
2.47 Ebers–Moll model (transport version) 71
2.48 Gummel–Poon model 73
2.49 ADS simulation circuit for VBIC model verification 75
2.50 Comparison between simulated and measured S–parameters 76
2.51 Small signal equivalent circuit of a PN–junction with noise sources 80
2.52 Noise sources in HBT 81

3.1 Process of photon detection by a photodiode 87
3.2 Cross-section of a top-illuminated photodiode 88
3.3 GaInAs Photodiode –3dB frequency due to transit time versus junction width 90
3.4 Small signal model of a photodiode 91

3.5 Layer structure on wafer as grown by MOMBE system 92
3.6 Wafer layers after diode mesa etching 93
3.7 Wafer layers after contact implementation 93
3.8 Wafer layers after "isolation" mask 94
3.9 Wafer layers after "Polyimide cover" mask 94
3.10 Cross section of a completed photodiode with anti-reflective coating . . . 95
3.11 SEM images of a photodiode at various fabrication process steps: (a) contact metal deposition and liftoff (b) polyimide cover — bird-eye view (c) polyimide cover — top view (d) metal 1 deposition and liftoff 97
3.12 Diagram of the optoelectronic measurement system 98
3.13 Pictures of the optoelectronic measurement system: (a) general view (b) optical and electrical probes for PD probing 99
3.14 Extraction of C_j 101
3.15 C–V measurements and processed data of both wafers 102
3.16 Noise source of a photodiode 104

4.1 Bipolar transistor configurations: (a) common emitter (b) common base (c) common collector 108
4.2 Small-signal hybrid-π model of a bipolar transistor 108
4.3 Common emitter stage: (a) circuit topology (b) small signal representation 110
4.4 Common emitter stage with degeneration resistor: (a) circuit topology (b) small signal representation 111
4.5 Common base stage: (a) circuit topology (b) small signal representation . 112
4.6 Common collector stage: (a) circuit topology (b) small signal representation 114
4.7 Circuit topology of a cascode stage 115
4.8 Negative-feedback network scheme 116
4.9 Scheme of opamp with PIPO feedback 118
4.10 Topology of the opamp 123
4.11 Closed-loop TIA topology (*Version B*) 124
4.12 TIA with common base input stage (*Version A*) 125
4.13 TIA version A and the path along which V_x is calculated 127
4.14 V_x versus β and R_{C3} 128
4.15 Simulated noise current 130
4.16 Simulated PSRR 132
4.17 Layout of TIAs: (a) version A (b) version B 134
4.18 TIA version B with 15 μm photodiode and biasing resistor: (a) scheme (b) layout 136
4.19 Probing of a circuit 139
4.20 A curve fitted to OEIC measured data 139
4.21 Simulated and measured performance of various TIAs 141

5.1 Transmission line electrical connection 146

5.2 Models of (a) lossless and (b) lossy transmission lines 147
5.3 Transmission line with periodical capacitive discontinuities 149
5.4 Structure of coplanar waveguide 150
5.5 Z_0 as a function of W and G 150
5.6 Structure of a typical TWA 151
5.7 Two signal paths with different delays 153
5.8 Scheme of an amplification stage 154
5.9 Input impedance of a TWA 155
5.10 Scheme of TWA's biasing 157
5.11 Simulated noise current of 4- and 5-stage amplifiers 158
5.12 Layout of TWAs: (a) strong peaking, 4 stages, electrical version (b) moderate peaking, 5 stage, with PD (c) weak peaking, 4 stages, with PD 159
5.13 Normalized frequency response of a photodiode, TWA, and optoelectronic circuit 160
5.14 Simulated and measured TWA transimpedance gain 163
5.15 Simulated and measured TWA group delay 163
5.16 Simulated and measured output return loss 164

6.1 C–V measurements of a 6000 Å wafer: (a) junction capacitance vs. applied voltage (b) junction width vs. applied voltage 168

A.1 MATLAB figure of cv_main 174
A.2 MATLAB figure of gpib_error 174
A.3 MATLAB figure of save_error 175
A.4 MATLAB figure of save_ok 175

B.1 Voltage amplifier with feedback network 193
B.2 Miller effect with (a) capacitive and (b) resistive feedback networks 194
B.3 C_μ as a feedback capacitor in common emitter stage 195

List of Tables

1.1 Recent works on InP-based lumped photoreceivers 18
1.2 Recent works on InP-based distributed photoreceivers 19

2.1 Layer growth properties 35
2.2 HBT fabrication process steps 43
2.3 HBT extracted parameters 69
2.4 Conversion of measured parameters to VBIC parameters 74
2.5 Summary of noise types 79

3.1 Physical properties of GaInAs and parameters used for calculating the –3dB bandwidth of a photodiode 90
3.2 Layer growth properties 92
3.3 PD fabrication process steps 96
3.4 Measured responsivity and external quantum efficiency of various PDs . . 100
3.5 Extracted small signal parameters of various PDs 101

4.1 Hybrid-π model terms expressed by T model terms 109
4.2 Actual values used in TIA designs 126
4.3 Simulated biasing points of all TIA transistors 128
4.4 TIA Noise performance 130
4.5 NEP of optoelectronic circuits 137
4.6 Simulated and measured performance of TIAs 140
4.7 Optoelectronic bandwidth of TIAs with various photodiodes 141

5.1 Transmission line quantities expressed by its model parameters 148
5.2 Values used for calculation of CPW's characteristics 150
5.3 Actual values of TWA's transmission lines 156
5.4 NEP of optoelectronic distributed circuits 161
5.5 Simulated and measured TWA electrical performances 164
5.6 Simulated and measured TWA optoelectronic performances 165

6.1 Comparison of this work to recent works on InP-based lumped photoreceivers . . . 169
6.2 Comparison of this work to recent works on InP-based distributed photoreceivers . . . 170

Abstract

Heterojunction Bipolar transistors (*HBT*s) based upon the Indium Phosphide material-system exhibit cutoff frequencies in excess of 200 GHz. The same semiconductor material-system also serves for the fabrication of photodiodes for optical communication at 1.55 μm. Therefore, InP-based HBTs play an important role in fiber optics communications. This work presents design of InP-based optoelectronic integrated circuits based on HBTs with InP emitter and GaInAs base and collector. Fiber optics communication strives towards fast rate of 100 Gbps and beyond. In order to achieve this demanding goal monolithic photoreceiver front-ends are required, coupled with state-of-the-art technology.

This research focuses on photoreceiver front-ends that consist of a photodiode and a transimpedance amplifier. Two types of amplifiers were designed: lumped and distributed. The former type excels in small area, low power consumption, and low noise, whilst the latter exhibits considerably higher bandwidth with the same transistor. The lumped amplifiers were designed as transimpedance feedback amplifiers. In the distributed amplifiers, usually designed as power amplifiers, the main goal was to reduce the input impedance. The amplifiers were characterized electrically up to 67 GHz, and the monolithic chips electro-optically up to 40 GHz.

The lumped circuits exhibit electrical bandwidth of 22 GHz and optical bandwidth of 14.2 GHz, whereas the distributed circuits show electrical bandwidth of 60 GHz and optical bandwidth of 20.2 GHz.

List of Acronyms

AC	Alternating Current
ADS	Advanced Design System
ARC	Anti-Reflective Coating
BC	Base–Collector (e.g. *BC junction*)
BE	Base–Emitter (e.g. *BE junction*)
BJT	Bipolar Junction Transistor
CAD	Computer-Aided Design
CB	Common Base
CC	Common Collector
CDR	Clock Data rate
CE	Common Emitter
CPW	CoPlanar Waveguide
DC	Direct Current
DHBT	Double-Heterojunction Bipolar Transistor
EM	Ebers–Moll (esp. EM model)
G-S-G	Ground-Signal-Ground
G-S-G-S-G	Ground-Signal-Ground-Signal-Ground
GP	Gummel–Poon (esp. GP model)
GPIB	General Purpose Interface Bus (IEEE 488)
GUI	Graphical User Interface
HBT	Heterojunction Bipolar Transistor
IIT	Israel Institute of Technology (a.k.a. Technion)
MOMBE	Metal Organic Molecular Beam Epitaxy
NEP	Noise-Equivalent Power
NPN	N-type – P-type – N-type (esp. *NPN transistor*)
NRZ	Non Return-to-Zero
OEIC	OptoElectronic Integrated Circuit
OPAMP	OPerational AMPlifier

PC	Personal Computer
PD	Photodiode
PECVD	Plasma Enhanced Chemical Vapor Deposition
PIN	P-type – Intrinsic – N-type (esp. *PIN–diode*)
PIPO	Parallel-In/Parallel-Out
PISO	Parallel-In/Serial-Out
PN	P-type – N-type (esp. *PN–junction*)
PR	Photo Resist
PSRR	Power Supply Rejection Ratio
RF	Radio Frequency
RZ	Return-to-Zero
RIE	Reactive Ion Etching
SEM	Scanning Electron Microscope
SHBT	Single-Heterojunction Bipolar Transistor
SI	Semi-Insulator
SIPO	Serial-In/Parallel-Out
SISO	Serial-In/Serial-Out
SNR	Signal-to-Noise Ratio
SPA	Semiconductor Parameter Analyzer
TCA	TransConductance Amplifier
TIA	TransImpedance Amplifier
TLM	Transfer Length Method
TWA	Traveling Wave Amplifier
VBIC	Vertical Bipolar Inter-Company (esp. VBIC model)
3D	Three Dimensional

List of Symbols

A	area
A	Amplifier's open-loop gain
a	flicker noise empirical exponential coefficient
A_j	junction area
a_i	small signal current gain
a_v	small signal voltage gain
b	base transport coefficient
b	burst noise empirical exponential coefficient
C	capacitance per unit length (of a transmission line)
c	the speed of light
C_{BC}	base–collector capacitance
C_{BE}	base–emitter capacitance
C_f	feedback capacitor
C_{in}	input capacitor
C_j	junction capacitance
C_{out}	output capacitor
C_{pad}	parasitic pad capacitance
C_{pBC}	capacitance between base and collector pads
C_{pBE}	capacitance between base and emitter pads
C_{pCE}	capacitance between collector and emitter pads
C_{port2}	capacitance measured at port 2 of a two-port network
C_S	source capacitance
$C_{st,in}$	input capacitance of an amplification stage
C_μ	base–collector capacitance (hybrid-π model)
C_π	base–emitter capacitance (hybrid-π model)
CJC	base–collector zero-bias capacitance (VBIC model)
CJE	base–emitter zero-bias junction capacitance (VBIC model)
D_e	electron diffusion coefficient (esp. in the base)

$D_{\pi s}$	voltage division at the input of common emitter stage
E	electrical length
$\overrightarrow{E}$	electrical field
E_c	bottom of conductance energy band
E_g	bandgap energy
E_{ph}	photon's energy
E_v	top of valence energy band
F	feedback transmission function
f	frequency
f_b	burst noise knee frequency
f_{Bragg}	Bragg frequency
f_{MAX}	Mason's power gain cutoff frequency
f_T	current gain cutoff frequency
f_{-3dB}	–3dB frequency
G	parallel conductance per unit length (of a transmission line)
G	distance between signal and round lines (coplanar waveguide)
G_{CL}	closed-loop gain
g_m	transistor's transconductance (hybrid-π model)
g_m	stage's transconductance (in a TWA)
G_p	power gain
h	Planck's constant
I_B	base current
$\overline{i_b^2}$	average base noise current
I_{bg}	lighting current due to background radiation
$\overline{i_{bn}^2}$	average burst noise current
I_C	collector current
$\overline{i_c^2}$	average collector noise current
I_{dark}	photodiode dark current
I_E	emitter current
I_{EE}	input supply current
I_F	forward injection current
$\overline{i_{fn}^2}$	average flicker noise current
I_n	total DC current of a photodiode
$\overline{i_n^2}$	average total noise current
I_{PD}	photodiode lighting current
I_R	reverse injection current
I_S	transistor saturation current

$\overline{i_{sn}^2}$	average shot noise current
$\overline{i_{tn}^2}$	average thermal noise current
I_0	PN–junction saturation current
$IBCN$	non-ideal base–collector saturation current (VBIC model)
$IBEI$	ideal base–emitter saturation current (VBIC model)
IS	transport saturation current (VBIC model)
K	Boltzmann's coefficient
K_b	burst noise empirical coefficient
K_f	flicker noise empirical coefficient
L	inductance per unit length (of a transmission line)
l	transmission line's length
L_b	inductance of base pad
L_c	inductance of collector pad
L_e	inductance of emitter pad
L_e	electron diffusion length
L_T	transfer length (in TLM measurement)
M	avalanche coefficient
N	number of stages (in a TWA)
n	electron density in the base
n	PN–junction ideality factor
n_{air}	diffraction coefficient of air
n_{ARC}	diffraction coefficient of anti-reflective coating
$\overline{n}_B$	electron density in the base under equilibrium conditions
N_d	donor concentration
n_{GaInAs}	diffraction coefficient of GaInAs
N_{opt}	optimal number of stages (in a TWA)
NCN	non-ideal base–collector emission coefficient (VBIC model)
NEI	ideal base–emitter emission coefficient (VBIC model)
NEP	noise-equivalent power of a photodiode
NF	forward emission coefficient (VBIC model)
NR	reverse emission coefficient (VBIC model)
P_{in}	internal perimeter of PD's p contact
P_{opt}	incident optical power
q	electron's electric charge
R	series resistance per unit length (of a transmission line)
$\Re$	photodiode responsivity
R_B	base external resistor

r_b	base resistance
R_C	collector external resistor
R_c	contact resistance (in TLM measurement)
r_c	collector extrinsic resistance
$r_{contact}$	photodiode contact resistance
r_{cr}	differential BC junction resistance in reverse mode
R_D	differential resistance of the three-diode branch at the base of Q3 (in the TIA circuits)
r_d	diode differential resistance
R_E	emitter external resistor
r_e	emitter intrinsic resistance
$R_{E,contact}$	emitter contact resistance
R_{EE}	biasing resistor connected in parallel to a photodiode
r_{ee}	emitter extrinsic resistance
R_f	feedback resistor
R_{in}	input resistance
R_L	load resistance
R_{out}	output resistance
r_{out}	collector intrinsic resistance; photodiode output resistance
R_{pad}	pad resistance
R_S, R_s	source resistance
$R_{st,in}$	input resistance of an amplification stage
R_T	total measured resistance (in TLM measurement)
R_{top}	reflectivity of a photodiode's top surface
r_π	base–emitter resistance (hybrid-π model)
RBI	intrinsic base resistance (VBIC model)
RBX	extrinsic base resistance (VBIC model)
RCI	intrinsic collector resistance (VBIC model)
RCX	extrinsic collector resistance (VBIC model)
RE	emitter resistance (VBIC model)
SNR	power signal-to-noise ratio
T	temperature
T	period
t	metal thickness (coplanar waveguide)
t_{tr}^e	electron transit time
t_{tr}^h	hole transit time
$\tan\delta$	dielectric loss factor
TF	forward transit time (VBIC model)

U	Mason's unilateral power gain
V_A	Early voltage
V_{BE}	base–emitter voltage
V_{CB}	collector–base voltage
V_{CC}	main supply voltage
V_{CE}	collector–emitter voltage
V_{CEO}	avalanche base–collector breakdown voltage of $\alpha_F = 1$
V_{EE}	input supply voltage
V_{ext}	extra supply voltage (in the TWAs)
v_g	group velocity
$\overline{v_n^2}$	average total noise voltage
v_{nCC}	supply voltage noise
v_{oA}	output noise due to variations in V_x (path A)
v_{oB}	output noise due to voltage variations at the base of Q3 (path B)
v_{oC}	output noise due to voltage variations at the output of the cascode stage (path C)
V_{on}	voltage drop on a diode under forward bias
v_p	phase velocity
V_{PDn}	supply voltage of PD's n contact (in the TWAs)
V_{PDp}	supply voltage of PD's p contact (in the TWAs)
v_{sat}	electron saturation velocity
v_{sat}^h	hole saturation velocity
$\overline{v_{tn}^2}$	average thermal noise voltage
V_{TO}	common emitter turn-on voltage
V_x	voltage at the base of Q2 (in the TIA circuits)
W	signal line width (coplanar waveguide)
W_B	base neutral region width
W_C	collector depletion region width
W_{PD}	photodiode depletion region width
$Y_{st,in}$	input admittance of an amplification stage
Z_e	emitter impedance in common collector configuration
Z_f	feedback impedance
Z_{in}	input impedance
$z_{in.a}$	opamp input impedance
Z_L	load impedance
Z_{out}	output impedance
$z_{out.a}$	opamp output impedance
Z_S, Z_s	source impedance

Z_T transimpedance gain
Z_π base–emitter impedance
Z_0 S–parameter/transmission line characteristic impedance
Z_{0in} characteristic impedance of TWA's input transmission line
Z_{0out} characteristic impedance of TWA's output transmission line
α absorption coefficient
α attenuation constant
α_F forward current transmission coefficient
α_R reverse current transmission coefficient
α_0 low frequency forward current transmission coefficient
α_0 forward current transmission coefficient without avalanche
β propagation constant
β_{AC} small signal forward current gain
β_{DC} DC forward current gain
β_F forward current gain
β_R reverse current gain
β_0 low frequency forward current gain
γ emitter emission efficiency
γ complex propagation constant
Γ_{in} voltage input reflectance coefficient
Γ_L voltage output reflectance coefficient
ΔE_c energy difference in the conductance band at a heterojunction
ΔE_v energy difference in the valence band at a heterojunction
Δf bandwidth
ΔT delay difference between the paths through two adjacent stages (in TWA)
ϵ_0 vacuum dielectric coefficient
ϵ_r semiconductor dielectric coefficient
η photodiode external quantum efficiency
λ wavelength
μ_e electron mobility
ρ_s sheet resistance
σ standard deviation
σ material's conductivity
τ_B base transit time
τ_C collector transit time
τ_D transistor forward transit time
τ_e electron lifetime (esp. in the base)

τ_{PD}	photodiode effective transit time
ϕ_{opt}	incident photon flux density (per unit area)
ϕ_0	absorbed photon flux density (per unit area)
ω	angular frequency
ω_p	pole's angular frequency
ω_T	current gain cutoff angular frequency
ω_z	zero's angular frequency
ω_{-3dB}	–3dB angular frequency

Chapter 1

Introduction

This chapter briefly introduces the main concepts of this work. The building blocks of the photoreceivers are the heterojunction bipolar transistor (*HBT*) and the photodiode (*PD*). They are introduced in the first part of this chapter. The second part of the chapter discusses the two types of amplifiers designed in this work — lumped and distributed.

1.1 Heterojunction Bipolar Transistor

Research in the area of InP/GaInAs HBTs is aimed at two main applications: (a) optoelectronic circuits for 1.55 μm fiber optic communications, and (b) fast, low-noise, RF and data conversion circuits. Single-heterojunction bipolar transistors (*SHBT*s) offer the option to use the base and collector layers for the photodetector. Other applications can make use of the double-heterojunction bipolar transistors (*DHBT*s), which exhibit higher breakdown voltage.

Two figures-of-merit determine the frequency performance of a transistor: f_T, the current gain cutoff frequency, and f_{MAX}, the unilateral power gain cutoff frequency. Both frequencies are set by the base resistance, r_b, the emitter resistances, r_e and r_{ee}, base–collector capacitance, C_{BC}, base–emitter capacitance, C_{BE}, and the transit time, τ_D [1]. The emitter resistance, r_e, is a function of the DC biasing of the transistor. Control over f_T and f_{MAX} is therefore achieved by controlling r_b, r_{ee}, C_{BE}, C_{BC}, and τ_D.

The above parameters can be controlled by changing fabrication characteristics. Changing the lateral sizes affects the capacitances and resistances [2], whilst vertical changes (in the layer thicknesses) also affect τ_D, but not r_{ee} [3]. However, achieving high cut-off frequencies at low bias currents is also of great importance as it reduces the power consumption of high-speed circuits [4].

The most relevant parameter for this work is the collector thickness (this will be explained shortly, in the last paragraph). The thicker the collector, the smaller C_{BC} is, but the longer τ_D. Consequently, transistors with thick collector sport low f_T and high f_{MAX}, and transistors with thin collector have extremely high f_T but low f_{MAX}. Optimization for f_T solely achieved record-breaking SHBTs with f_T of 452 [3], 509 [5], and 604 GHz [6], and as low f_{MAX} as 155, 275, and 246 GHz, respectively. The very same transistor, when optimized for f_{MAX}, obtains f_T of 353 GHz and f_{MAX} of 435 GHz [2]. On the DHBT side, InP-based transistors reach f_T of 450 GHz and f_{MAX} of 490 GHz, when optimized for both [7]. Note that optimizing for both f_T and f_{MAX} is the preferred method for high-speed circuits [8].

Fabrication of extremely fast HBTs involves high precision time consuming techniques such as e-beam lithography. As a result, complexity of circuits based on these transistors is limited by the number of transistors. Moreover, the yield of the mentioned processes is unknown, and is believed to be lower than processes consisting upon standard mask lithography.

The HBTs used in this work are SHBTs with 8000 or 6000 Å thick collector. This structure is essential for the fabrication of photodiodes from the base–collector layers, enabling the creation of monolithic optoelectronic circuits. Nevertheless, the transistors have as low f_T as 67 and 90 GHz (for 8000 and 6000 Å, respectively) due to the long transit time, τ_D, originated by thick collectors. The process employs only standard mask lithography on 2″ InP wafers.

1.2 Photodiode

Due to its bandgap of 0.75 eV the $Ga_{0.47}In_{0.53}As$ is the material of choice for photodiodes (*PD*s) intended for 1.55 μm wavelength [9]. This makes the InP/GaInAs material system, and particularly the InP/GaInAs HBT, attractive for monolithic optoelectronic applications. Several structures of photodiode are available [10]:

Top Illuminated — Incident light is perpendicular to the junction plane (usually a PIN junction), that is, absorbtion decays along the junction in the same direction along which charge carriers are drifting. Responsivity and transit time are set by the thickness of the junction. This type of photodiode is the simplest to implement, but one has to tradeoff responsivity to achieve high frequency response, and vice versa. If fabricated monolithically with HBTs, this structure can also share the base–collector layers.

Avalanche — Similar to the above PIN diode, but makes use of avalanche occurrence to amplify the signal. Avalanche photodiodes are difficult to fabricate by mesa technology due to sidewall leakage.

Waveguide (Side Illuminated) — Incident light is parallel to the junction plane, enabling one to optimize both responsivity and transit time. This type of PD is the fastest and most efficient one, but its fabrication is complex and requires high precision precess. In addition, positioning of the fiber is difficult, requiring more processing steps and making the packaging relatively expensive. Also, this type cannot share the base–collector layers with HBTs.

Metal-Semiconductor-Metal — Light incidents a semiconductor area located between two metal contacts. Advantages of this type of PD are relevant only for sub-micron processes.

Back Illuminated/Refractive Facet — The structure is similar to top illuminated, except that the top contact doesn't have optical window, and light is applied from the back side of the wafer (of from its edge, in the refractive facet case) [11]. Light is reflected from the top contact back to the junction, and so responsivity is increased. These photodiodes are complex to implement and expensive to package because of the difficulties concerned with lighting from the back/edge of the wafer.

In order to utilize the advantage of HBTs the photodiodes used in this work are top illuminated. Bandwidth of the top illuminated diode has been widely investigated [12][13]. It was shown that thickness of the diode should be optimized for transit time, junction capacitance, and responsivity. Several improvements in the layer structure were suggested [14][15], but these improvements cannot be applied to PDs that share base–collector layers with HBTs.

As the photodiodes fabricated in this work are monolithically fabricated with HBTs, they were fabricated with 8000 and 6000 Å junctions. Anti-reflective coating was not used in this work but can be added at a latter stage.

1.3 Monolithic Receivers

Photoreceiver front-ends consist of a photodiode and an amplifier. Incident light generates photocurrent in the PD, and this current is amplified (or converted to voltage) by the amplifier. However, the connection between the PD and the amplifier is a major concern.

If the PD and the amplifier are separate discrete components they are placed on a substrate and connected to each other. In most cases this connection is implemented by wire-bonds, as illustrated in figure 1.1. Modeling of wire-bonds has been widely

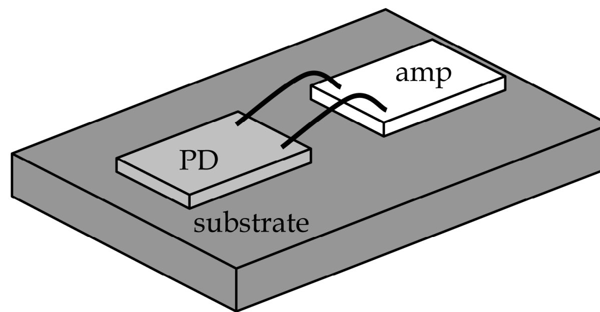


Figure 1.1: Scheme of a discrete element receiver with wire-bonds

investigated, and the main outcome is that the bond acts as an inductor. More advanced models include two capacitors at the ends of the wire-bond (see [16], for example). In mm-wave applications inductances of the wire-bonds can be harmful and distort the frequency response.

In order to understand the wire-bond influence better, it should be indicated that the inductance depends on the bond's length [17]. Figure 1.2 demonstrates the effect of wire-bonds of two different lengths on the frequency response of a photoreceiver [18]. It is evident that long wires significantly reduce the bandwidth and deteriorate the frequency response. The higher the bandwidth of the receiver, the shorter the wire-bonds should be.

On top of the above, fabrication of two separate components, mounting them on a substrate, and connecting them by wire-bonds, increase manufacturing costs and complexity. To overcome these problems monolithic design, which incorporates both the photodiode and the electronic circuits on a single die, is required.

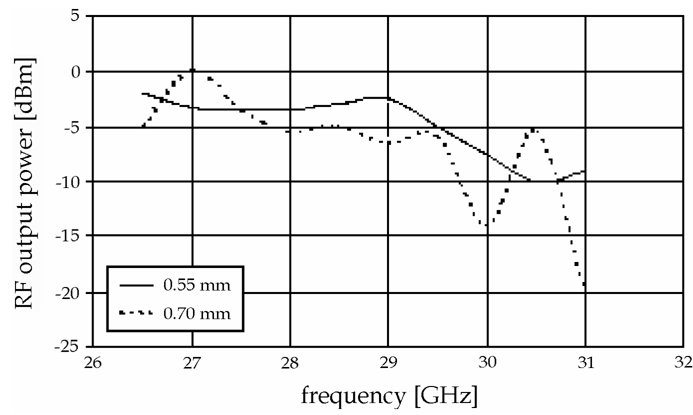


Figure 1.2: Frequency response of a photoreceiver with 0.55 mm and 0.70 mm wire-bonds (from [18])

In this work two monolithic receivers were designed and fabricated: lumped and distributed. Both are introduced hereinafter.

1.4 Lumped Monolithic Receivers

Lumped circuits exhibit stable performance, flat frequency response, high linearity, low power consumption, and small area. Their main drawback is the bandwidth — it is limited to about third of the transistor's f_T. Being relatively slow, the top illuminated PD is a good match for lumped preamplifiers, especially when HBT is considered. The work done in this area on InP/GaInAs technology in the last years is summarized in table 1.1.

Team ⟶	ETHZ Switzerland [19]	ETHZ Switzerland [20]	Notre Dame IN, USA [21]	KAIST Korea [11]
Transistor	HBT	HBT	HEMT	HBT
Photodiode	top illuminated	top illuminated	top illuminated	refractive facet
Anti-reflective Coating	yes	yes	yes	yes
Optical Bandwidth [GHz]	30	50	8.3	6.9
Optoelectronic Gain [V/W]	48	N/A	410	85
Gain-Bandwidth Product $\left[\mathrm{THz} \cdot \frac{\mathrm{V}}{\mathrm{W}}\right]$	1.44	N/A	3.40	0.59

Table 1.1: Recent works on InP-based lumped photoreceivers

Notes: In [20] 4000 Å collector was used in the transistors, and thus the optoelectronic gain is assumed low. In [21] the transistors are HEMTs, enabling their optimization separately from the PD. In [11] high gain is achieved due to the refractive facet diode. Only [19] can be fairly compared to the lumped circuits of this work.

The lumped amplifiers of this work are based on a previous design [22] with major modifications, in order to enhance bandwidth and achieve more stable and reliable performance. Both schematics and layout were completely redesigned.

1.5 Distributed Monolithic Receivers

Distributed amplifiers, also knows as traveling wave amplifiers (*TWAs*), achieve high bandwidth in comparison to lumped amplifier — they almost reach f_T [8]. However, they consume more power, more area, have lower linearity and wavier response. In this work the wavy response was utilized for enhancement of the optoelectronic bandwidth [23]. During the last years research in InP-based optoelectronic distributed photoreceivers concentrated in HEMTs and waveguide photodiodes. The relevant works are listed in table 1.2.

Team $\longrightarrow$	NTT Corp. Japan [24]	HHI Berlin Germany [25]
Transistor	HEMT	HEMT
Photodiode	waveguide	waveguide
Anti-reflective Coating	yes	yes
Optical Bandwidth [GHz]	46.5	72.0
Optoelectronic Gain [V/W]	55	45
Gain-Bandwidth Product $\left[\text{THz} \cdot \frac{\text{V}}{\text{W}}\right]$	2.56	3.24

Table 1.2: Recent works on InP-based distributed photoreceivers

The Technion's technology has proven suitable for broadband TWAs in the form of power amplifiers [8][26]. These amplifiers achieved bandwidth of 75 GHz and record-breaking bandwidth-to-f_T ratio in excess of 0.92. This work, in it distributed circuits part, was targeted at modifying the design to transimpedance amplifier and improving the yield, which in practice required an all-new design.

References for Chapter 1

[1] Solon Jose Spiegel, Dan Ritter, R. A. Hamm, A. Feygenson, and P. R. Smith, "Extraction of the InP/GaInAs heterojunction bipolar transistor small-signal equivalent circuit", *IEEE Transactions on Electron Devices*, vol. 42, no. 6, pp. 1059–1064, June 1995.

[2] 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, September 2004.

[3] Walid Hafez, Jie-Wei Lai, and Milton Feng, "Vertical scaling of 0.25-μm emitter InP/InGaAs single heterojunction bipolar transistors with f_T of 452 GHz", *IEEE Electron Device Letters*, vol. 24, no. 7, pp. 436–438, July 2003.

[4] Walid Hafez, Jie-Wei Lai, and Milton Feng, "Low-power high-speed operation of submicron InP–InGaAs SHBTs at 1 mA", *IEEE Electron Device Letters*, vol. 24, no. 7, pp. 427–429, July 2003.

[5] W. Hafez, Jie-Wei Lai, and M. Feng, "InP/InGaAs SHBTs with 75 nm collector and $f_T > 500$ GHz", *Electronics Letters*, vol. 39, no. 20, pp. 1475–1476, October 2003.

[6] Walid Hafez and Milton Feng, "Experimental demonstration of pseudomorphic heterojunction bipolar transistors with cutoff frequencies above 600 GHz", *Applied Physics Letters*, vol. 86, no. 15, pp. 152101, April 2005.

[7] Zach Griffith, Mark J. W. Rodwell, Xiao-Ming Fang, Dmitri Loubychev, Ying Wu, Joel M. Fastenau, and Amy W. K. Liu, "InGaAs/InP DHBTs with 120-nm collector having simultaneously high f_T , $f_{max} \geq 450$ GHz", *IEEE Electron Device Letters*, vol. 26, no. 8, pp. 530–532, August 2005.

[8] E. Cohen, Y. Betser, B. Sheinman, S. Cohen, S. Sidorov, A. Gavrilov, and D. Ritter, "75 GHz InP HBT distributed amplifier with record figures of merit and low power dissipation", *IEEE Transactions on Electron Devices*, vol. 53, no. 2, pp. 392–394, February 2006.

[9] Hong Wang and Geok Ing Ng, "Electrical properties and transport mechanisms of InP/InGaAs HBTs operated at low temperature", *IEEE Transactions on Electron Devices*, vol. 48, no. 8, pp. 1492–1497, August 2001.

[10] Mukunda B. Das, "Optoelectronic detectors and receivers: Speed and sensitivity limits", *Proceedings of the 1998 Conference on Optoelectronic and Microelectronic Materials Devices*, pp. 15–22, December 1988.

[11] Bangkeun Lee, Yongjoo Song, and Kyounghoon Yang, "InP-based OEIC photoreceivers using shared layer integration technology of heterojunction bipolar transistors and refracting-facet photodiodes", *2003 International Conference on Solid State Devices and Materials*, pp. 182–183, September 2003.

[12] D. Wake, R. H. Walling, I. D. Henning, and D. G. Parker, "Planar-junction, top-illuminated GaInAs/InP pin photodiode with bandwidth of 25 GHz", *Electronics Letters*, vol. 25, no. 15, pp. 967–969, July 1989.

[13] M. Agethen, D. Keiper, G. Janssen, A. Brennemann, P. Velling, C. van der Berg, and R. M. Bertenburg, "InGaAs PIN detectors for frequencies above 100 GHz", *14th Conference on Indium Phodphide and Related Materials*, pp. 673–676, May 2002.

[14] Chong-Long Ho, Meng-Chyi Wu, Wen-Jeng Ho, and Jy-Wang Liaw, "Bandwidth enhancement for p-end-illuminated InP/InGaAs/InP p-i-n photodiodes by utilizing symmetrical doping profiles", *IEEE Journal of Lightwave Technology*, vol. 17, no. 5, pp. 912–917, May 1999.

[15] Y. Muramoto and T. Ishibashi, "InP/InGaAs pin photodiode structure maximising bandwidth and efficiency", *Electronics Letters*, vol. 39, no. 24, pp. 1749–1750, November 2003.

[16] Federico Alimenti, Paolo Mezzanotte, Luca Roselli, and Roberto Sorrentino, "Modeling and characterization of the bonding-wire interconnection", *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 1, pp. 142–150, January 2001.

[17] L. Gomez-Rojas, N. J. Gomes, X. Wang, and P. A. Davies, "Simple technique for fine tuning of impedance matching circuits for the mm-wave region", *IEE Seminar on Packaging and Interconnects at Microwave and mm-Wave Frequencies*, pp. 7/1–7/12, 2000.

[18] *http://transmanche.kent.ac.uk/pin.htm*.

[19] Martin Bitter, Raimond Bauknecht, Werner Hunziker, and Hans Melchior, "Monolithic InGaAs–InP p-i-n/HBT 40-Gb/s optical receiver module", *IEEE Photonics Technology Letters*, vol. 12, no. 1, pp. 74–76, January 2000.

[20] D. Huber, R. Bauknecht, C. Bergamaschi, M. Bitter, A. Huber, T. Morf, A. Neiger, M. Rohner, I. Schnyder, V. Schwarz, and H. Jäckel, "InP–InGaAs single HBT technology for photoreceiver OEICs at 40 Gb/s and beyond", *IEEE Journal of Lightwave Technology*, vol. 18, no. 7, pp. 992–1000, July 2000.

[21] P. Fay, C. Caneau, and I. Adesida, "High-speed MSM/HEMT and p-i-n/HEMT monolithic photoreceivers", *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 1, pp. 62–67, January 2002.

[22] Benny Sheinman, *Modeling and Performance of Ultrafast InP Based Heterojunction Bipolar Transistors and Circuits*, PhD thesis, Technion – IIT, October 2004.

[23] J. Bellon and M. J. N. Sibley, "Frequency response compensation of transit time limited pin photodiode", *Electronics Letters*, vol. 36, no. 14, pp. 1222–1223, July 2000.

[24] Kiyoto Takahata, Yoshifumi Muramoto, Hideki Fukano, Kazutoshi Kato, Atsuo Kozen, Shunji Kimura, Yuhki Imai, Yutaka Miyamoto, Osaake Nakajima, and Yutaka Matsuoka, "Ultrafast monolithic receiver OEIC composed of multimode waveguide p-i-n photodiode and HEMT distributed amplifier", *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 6, no. 1, pp. 31–37, January/February 2000.

[25] Gebre Giorgis Mekonnen, Heinz-Gunter Bach, Andreas Beling, Reinhard Kunkel, Detlef Schmidt, and Wolfgang Schlaak, "80-Gb/s InP-based waveguide-integrated photoreceiver", *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 11, no. 2, pp. 356–360, March/April 2005.

[26] Emanuel Cohen, "Design of wide-band distributed amplifier based on InP/GaInAs hetrojunction bipolar transistors", Master's thesis, Technion – IIT, 2002.

Chapter 2

Heterojunction Bipolar Transistor

The heterojunction bipolar transistor (*HBT*) is the basic building block of the integrated circuits designed in this work. This chapter deals with all aspects concerned with the design and fabrication of HBTs required for the circuit design. It opens with an overview on the physics and phenomena that characterize a HBT, followed by full description of the HBT fabrication process. The next sections deal with measurements and characterization of the transistor, and the extraction of a small signal model. The results obtained by the measurement, characterization, and extraction are used for the construction of a VBIC simulation model. This procedure is described in the following section. A short introduction to noise mechanisms that reside in the HBT complete the chapter.

2.1 HBT Physical Characteristics and Phenomena

This section briefly outlines the physics underlies the HBT device. Physical view on the transistor is essential for properly optimizing the fabrication process, constructing a simulation model, and designing an integrated circuit.

2.1.1 DC Characteristics and Phenomena

Operation under DC conditions is the basic mode of the transistor in which the most typical physical phenomena occur. The principles and concepts required for understanding transistor DC operation are overviewed below in close relationship to the HBT device.

Physics of Normal Forward Operation

The HBT, like bipolar junction transistors (*BJT*s), consists of an emitter, base, and collector. This work will focus on NPN transistors, in which the emitter and the collector are n-type semiconductors, and the base is p-type. One can refer to this structure as two PN–junctions: base–emitter junction (*BE junction*) and base–collector junction (*BC junction*). In the *forward active mode* the BE junction is forward biased, whilst BC junction is reverse biased.

The forward biased BE junction emits electrons to the base and holes to the emitter. Since the emitter doping concentration is significantly higher than that of the base, electron emission into the base is much stronger than hole emission back to the emitter (note that in HBT device this injection ratio in BE junction is achieved by a heterojunction, not by doping concentration ratio).

The base is considerably shorter than the electron diffusion length in the base, L_e. Diffusion transport the electrons to the depletion region of BC junction. Once an electron reaches the depletion region it drifts to the collector neutral region due to the strong electrical field in the depletion region.

As mentioned above, the base is p-type doped. The fact that both electrons and holes exist in the base brings about a recombination process in a rate determined by electron lifetime in the base, τ_e.

Figure 2.1a shows a schematic view of the structure of one-dimensional BJT. Figure 2.1b shows the route of electrons from the emitter neutral region right to the collector neutral region. Note the electron that recombines with a hole in the base.

Excess Charge Carrier Density

Making the assumption that electrons are weakly injected from the emitter to the base, they become minority charge carriers in their new p-type environment. Figure 2.2 shows charge carrier densities vs. position along a one-dimensional BJT (see [1]).

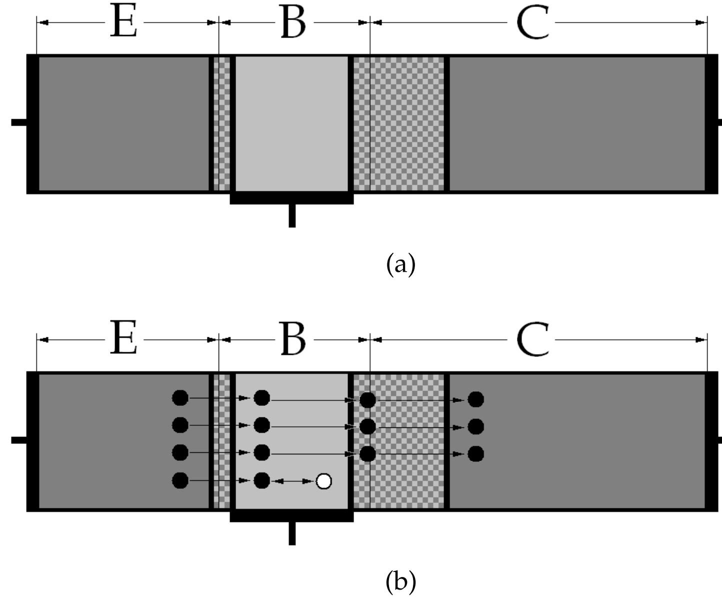


Figure 2.1: One-dimensional BJT (a) schematic structure and (b) route of electrons in forward active mode.
N-type areas are marked in dark gray, p-type in light gray. Checkered areas represent depletion regions, black circles represent electrons, and white circle represents hole. The letters E, B, and C stand for emitter, base, and collector, respectively.

Assuming that the quasi-Fermi energy levels are constant throughout BE depletion region, electron density at $x = 0$ is given by:

$$n(x = 0) = \overline{n}_B \cdot e^{\frac{qV_{BE}}{KT}} \tag{2.1}$$

where $\overline{n}_B$ is the electron density in the base under equilibrium conditions, q denotes the electron's electric charge, K is Boltzmann's coefficient, and T is temperature. On the other side of the base, i.e. at $x = W_B$, electrons are quickly collected by the BC depletion region. This causes their density at this location to tend to zero. In addition, the fact that base neutral region width, W_B, is much shorter than electron diffusion length, L_e, results in linear density profile along the base as it behaves like a short diode.

Making another assumption that electron current is limited by diffusion (thanks to the short base), one concludes that emitter current due to electrons is determined by the slope of electron density at $x = 0$, and can be written as [2]:

$$J_e = q \cdot D_e \cdot \frac{n(x = 0)}{W_B} \tag{2.2}$$

where D_e is the electron diffusion coefficient in the base.

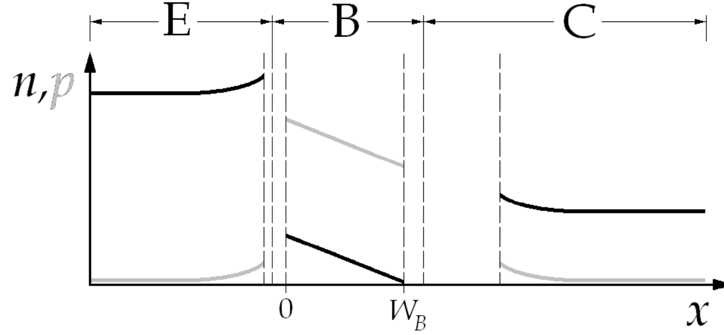


Figure 2.2: Charge carrier densities vs. position along BJT in forward active mode.
Electron density is marked in black line, hole density in gray. Dashed lines denote metallurgic junctions with depletion region borders from both sides. (Not to scale.)

Hole Current and Current Gain

A small fraction of the total current flowing through the transistor consists of hole current. Holes are injected from the base to the emitter, though BE junction is designed to keep this injection as small as possible. The ratio between electron and total (electron + hole) currents injected in BE junction is called *emitter emission efficiency*, and marked as γ ($0 \leq \gamma \leq 1$).

In addition, electrons and holes recombine in the base, as mentioned before. The ratio between number of electrons injected from the emitter and number of electrons arriving BC depletion region is called *base transport coefficient*, and marked as b ($0 \leq b \leq 1$).

These two processes result in a current that enters the transistor from the base contact – contrary to electron current that flows from the emitter contact to the collector contact. Since electron current consists of negative charge carriers, the electrical current direction of emitter and collector currents is from the collector to the emitter, as illustrated in figure 2.3. For a given biasing point, electron-related currents behave according to the following relationship:

$$I_C = I_E \cdot \gamma b \tag{2.3}$$

The expression γb that appears in (2.3) can be marked as α_F, thus

$$\alpha_F = \gamma b \tag{2.4}$$

Substituting (2.4) into (2.3) gives

$$I_C = I_E \cdot \alpha_F \tag{2.5}$$

Obviously, all transistor currents follow Kirchoff's current law, namely

$$I_E = I_C + I_B \tag{2.6}$$

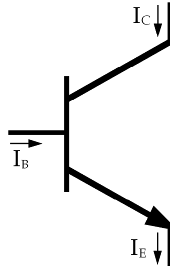


Figure 2.3: Electrical currents marked on a schematic BJT symbol

which, when combined with (2.5), can be written as

$$I_C = I_B \left(\frac{\alpha_F}{1 - \alpha_F} \right) \tag{2.7}$$

The parenthetic expression in (2.7) is known as the *forward current gain*, denoted by β_F:

$$\beta_F = \frac{\alpha_F}{1 - \alpha_F} \tag{2.8}$$

Finally, (2.7) can be expressed as

$$I_C = I_B \cdot \beta_F \tag{2.9}$$

which is the equation that best describes the transistor forward current gain. Note that β_F is bias independent, thus it has constant value as long as no other physical effects occur.

By examining (2.8) one concludes that the closer α_F to 1 the higher the current gain β_F. In practice this is done by designing γ and b to be as close to 1 as possible, by means of fabrication optimization.

Energy Bands and Transistor Parameters

Emitter emission efficiency and base transport coefficient are critical parameters of any BJT. Controlling these parameters is possible through careful design of the transistor materials and doping concentrations. The behavior of the parameters as a result of the design is well understood by studying the energy band model of the BJT.

A band diagram of a BJT under equilibrium conditions is shown in figure 2.4a. It can be observed in the diagram that the emitter doping level is higher than that of the base. This one-sided junction, when forward biased, causes electron injection from the emitter to the base to be much stronger than hole injection from the base to the emitter, making γ closer to 1. A similar diagram in forward active mode is illustrated in figure 2.4b. In this case that BE junction is forward biased, and BC junction is reverse biased.

So far only "regular" BJT, in which only one semiconductor is exploited, was discussed. A HBT features a heterojunction as BE junction that offers more flexibility in the design of other transistor parameters. Figure 2.4c shows a band diagram of a HBT under equilibrium conditions. A diagram in forward active mode is shown in figure 2.4d.

The energy difference in the valence band, ΔE_v, entraps holes and avoids them from being injected to the emitter, whilst the difference in the conductance band, ΔE_c, acts as a Schottky contact that emits electrons to the base. This mechanism sets γ to 1 regardless of the emitter and base doping concentrations.

Another improvement implemented in HBTs is base composition grading. This method takes advantage of the compound semiconductor the base in constructed from. Altering the composition ratio, x, provides control over the band gap of the semiconductor. This is used for creating a band diagram shown in figure 2.5. Since the base is p-type semiconductor it holds E_v horizontal. The graded band gap, caused by the graded composition, inclines E_c, so that electrons (but not holes) are accelerated by the band slope — which is so-called *quasi– electrical field*. The acceleration causes the electrons to stay shorter time in the base, which improves transistor AC performance.

In this work only Single–Heterojunction Bipolar Transistor (*SHBT*) is discussed. In SHBT there's only one heterojunction — BE junction. Another type of HBT, the Double–Heterojunction Bipolar Transistor (*DHBT*), implements two heterojunctions — both BE and BC junctions.

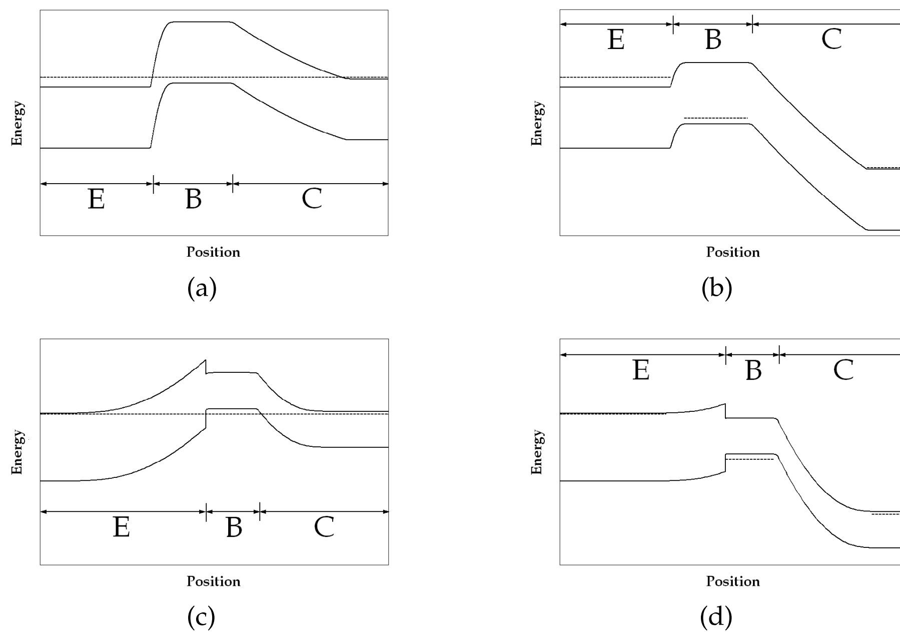
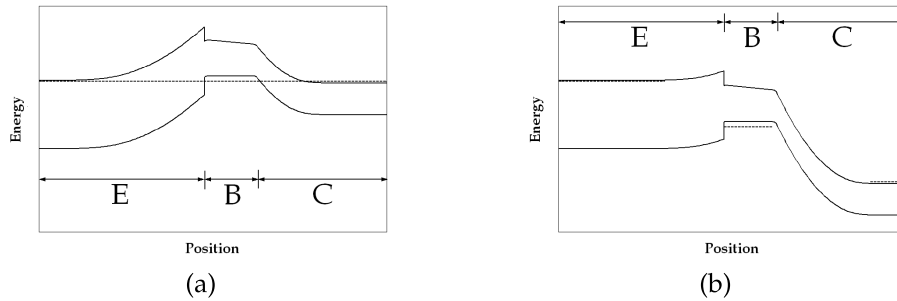


Figure 2.4: Energy band diagrams of (a) BJT in equilibrium and (b) in forward active mode, and (c) HBT in equilibrium and (d) in forward active mode. Valence and conductance bands in solid lines, Fermi (or quasi-Fermi) levels in dashed lines.



(a) (b)

Figure 2.5: Energy band diagrams of (a) base-graded HBT in equilibrium and (b) in forward active mode

Kirk Effect

In fact, the neutral region of the collector has no effect on transistor operation, i.e. on current gain. This means that the collector neutral region acts as a serial resistor that contributes nothing to the current gain and consumes power (there are more problems caused by this resistance, described later in section 2.1.2). The only part of the collector that makes any difference to the gain is BC depletion region. For this reason BC junction is usually designed as a PIN–diode, so that BC depletion region is located at the intrinsic part of the diode. The n-type side is featured by high doping level, resulting in low serial resistance. This n-type region is called *subcollector*.

Figure 2.6 shows a schematic view of the structure of one-dimensional BJT, including the subcollector. As with every PIN–diode, the intrinsic region requires no bias (or a very small bias) to become completely depleted.

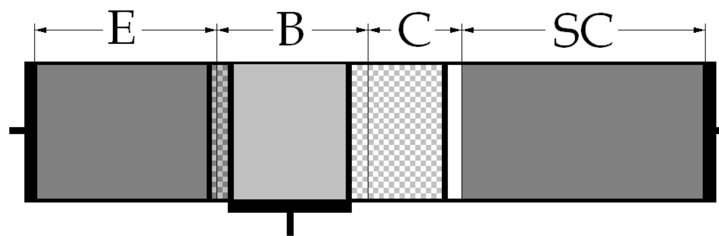


Figure 2.6: One-dimensional BJT schematic structure with subcollector. Intrinsic collector area is marked in white. The letters SC stand for subcollector.

In actual fabrication processes growing a purely intrinsic semiconductor layer is not feasible. Hence, a minimal doping level should be added to the collector layer, making BC junction a $P^+N^-N^+$ diode.

If the current flowing through the transistor is considerably large, electron density in BC depletion region rises and cannot be neglected. When electron density becomes

greater than n-dopant concentration of the collector, the entire BC depletion region moves from BC metallurgic junction towards collector-subcollector metallurgic junction. The high doping concentration of the subcollector and electron density in the collector make the new depletion region significantly narrow.

The process mentioned above is called *Kirk effect*. It affects forward current gain β_F (since there's recombination in the formerly– BC depletion region which makes b smaller) as well as AC performance to be discussed later (section 2.1.2). In most cases Kirk effect is negligible in DC terms due to other dominant effect — avalanche base–collector breakdown — described hereinafter.

Avalanche Base–Collector Breakdown

So far we have assumed that electron scattering in BC depletion region doesn't result electron–hole generation. For extreme electrical fields this proposition becomes incorrect. Strong electrical field in BC depletion region, a result of high voltage applied to BC junction, brings about the electrons to generate electron–hole pairs. These electron–hole pairs drift to the base and the collector due to the strong electrical field that exists in BC depletion region.

This occurrence is called *avalanche base–collector breakdown*. It increases collector current I_C and leaves emitter current I_E unchanged, dictating an appropriate change in base current I_B. In terms of transistor parameters, avalanche base-collector breakdown increases α_F to be closer to 1, and even greater than 1. In some nomenclatures α_F is expressed as $M \cdot \alpha_0$, where M is the avalanche coefficient and α_0 is actually α_F when no avalanche occurs.

At the point of $\alpha_F = 1$ equation (2.5) becomes

$$I_C = I_E \tag{2.10}$$

thus it can be concluded from (2.6) that

$$I_B = 0 \tag{2.11}$$

and (2.8) yields:

$$\beta_F \longrightarrow \infty \tag{2.12}$$

The collector–base voltage V_{CB} applied to the transistor when $\alpha_F = 1$ is denoted by V_{CEO}.

As I_C increases α_F exceeds 1 and equation (2.5) can be written as

$$I_C > I_E \tag{2.13}$$

so that I_B changes its direction, i.e.

$$I_B < 0 \tag{2.14}$$

Avalanche base–collector breakdown can be clearly observed in common base measurement, described later (in section 2.3.1).

Early Effect

As mentioned before, forward current gain is considered to be bias independent. Recall that the parameters constructing β_F are γ, emitter emission efficiency, and b, base transport coefficient. γ is determined by BE junction design and has nothing with biasing, and b is related to the base width W_B, which was assumed so far to be constant. Another β_F–related correlation between W_B and I_E is the slope of electron density in the base, that is inversely proportional to W_B (see equation (2.2)).

The assumption of constant W_B is imprecise. Doping level of the base side of BC junction is considerably higher than that of the collector side. As a result BC depletion region shrinks and stretches (according to BC bias) mainly, but not only, at the expense of the collector. However, the small changes in BC depletion region on the base side are not negligible, and they find their way to the forward current gain.

The final result of this effect, called *Early effect*, is a rise in β_F when BC reverse bias is increased in its absolute value. In different words, I_C varies with collector–emitter voltage, V_{CE}, while I_B remains constant [1]. Early effect, if existent, is evident in common emitter measurement, described later (in section 2.3.1).

2.1.2 AC Characteristics and Phenomena

The high frequency response is a major consideration in HBT design. This section overviews the small signal model elements that affect AC performance of the HBT.

Basically, the small signal elements that affect AC behavior are resistors and capacitors, constructing RC poles in the transistor transmission function. In addition, any delay in the charge carriers transit through the transistor will leave its footprints in the transmission function in the form of phase or delay element.

Process nature and limitations dictate a particular HBT structure. A schematic cross-sectional view of a simplified HBT is illustrated in figure 2.7a. The small signal elements, as placed on a HBT structure in figure 2.7b, are detailed hereinafter.

Emitter Resistances

The emitter consists of two parts: BE junction and the neutral region between BE junction and emitter contact. The neutral region is simply an ohmic resistor whose resistance is determined by donor doping concentration, electron mobility μ_e, contact quality, and emitter geometric dimensions. This resistance is denoted by r_{ee} in figure 2.7b.

BE junction, like every PN–junction under forward bias, can be represented by a small signal scheme of a resistor in parallel to a capacitor. The dynamic resistance of BE junction is marked by r_e in figure 2.7b. Its value is given by

$$r_e = \frac{KT}{qI_E} \tag{2.15}$$

just like a PN–junction. Note that the higher the current flowing in the emitter, the smaller the r_e.

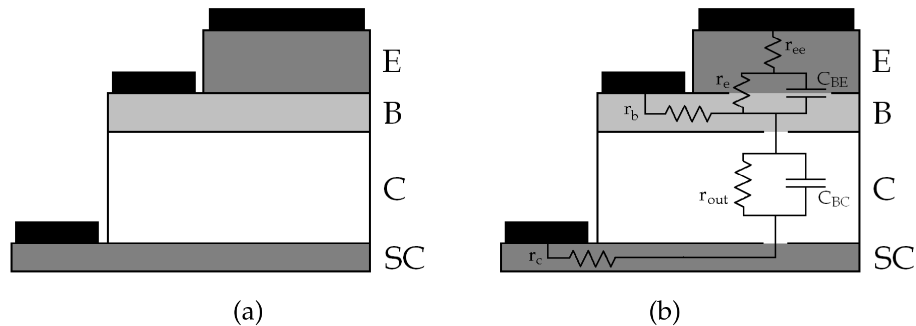


Figure 2.7: One-dimensional HBT (a) schematic structure and (b) small signal elements.
N-type areas are marked in dark gray, p-type in light gray, and intrinsic in white.

Collector Resistances

Similarly to emitter resistances, collector resistances are subcollector ohmic resistance between BC junction and collector contact (r_c in figure 2.7b) and BC junction dynamic resistance (r_{out} in figure 2.7b). The latter's name is taken from common emitter configuration in which r_{out} is the output resistance. In forward active mode BC junction is under reverse bias, thus one would have expected that $r_{out} \longrightarrow \infty$. In practical transistors Early effect and avalanche breakdown are accentuated as a decrease in r_{out}.

Base Resistance

The transistor operation is based on electron injection from the emitter to the collector through the base. This happens only in the area geometrically located under the emitter. The current flowing from the base to its contact has to pass through a neutral region between the base active area and base contact. This region behaves like an ohmic resistor denoted by r_b in figure 2.7b.

Base–Collector Capacitance

Like BE junction, BC junction contributes a resistor and a capacitor to the small signal model. Since BC junction is reverse biased, its diffusion capacitance can be neglected. This leaves the scheme with r_{out} and the junction capacitance, C_{BC}. The fact that BC junction is a PIN–diode causes that a very small reverse bias (in its absolute value) is required for depleting the entire intrinsic region. As a result, constant capacitance over various voltages is easily achieved.

In terms of performance in frequency domain, the existence of both capacitors and resistors in the small signal model appears as RC time constants that reduce the characteristic frequencies of the transistor — f_T and f_{MAX} — that will be discussed later on in

this chapter (sections 2.4.6 and 2.4.7).

Base–Emitter Capacitance

This capacitance is a byproduct of BE junction. As mentioned before, the junction becomes a resistor and a capacitor when coming to small signal model. Hence, the BE capacitance of the model equals to the junction capacitance.

Forward Transit Time

The time it takes to an electron to pass through the transistor contributes a delay (or a phase) to the frequency response of the transistor. This transit time consists of base transit time and collector transit time.

Base transit time, τ_B, is determined by the base width, W_B, and electron velocity in the base. Assuming diffusive transport, the transit time is given by [2]:

$$\tau_B = \frac{{W_B}^2}{2D_e} \tag{2.16}$$

Although fabrication of an extremely narrow base is feasible, it increases r_b to unacceptable values. Base composition grading shortens τ_B, though.

Collector delay, τ_C, is an outcome of the time required for an electron to go through BC depletion region. It can be estimated as follows [3]:

$$\tau_C \approx \frac{W_C}{2v_{sat}} \tag{2.17}$$

where W_C is BC depletion region width (i.e. collector width) and v_{sat} is the electron saturation velocity. Transistor layer design should take τ_C into account. This is done by optimizing W_C for the best C_{BC} and τ_C integration.

The total transistor transit time consists of the transit times of the base and the collector, namely

$$\tau_D = \tau_B + \tau_C \tag{2.18}$$

τ_D is the delay in the expression of α_F, described later in section 2.4.

2.2 HBT Fabrication and Structure

The HBT discussed in this work is base upon the InP material system. This section summarizes the process steps as well as the final product with emphasis on the aspects that affect device physics and performance.

2.2.1 Fabrication Process

HBT fabrication consists of two major levels: layer growth and wafer processing. Once all layers are grown on an InP wafer, wafer processing starts and no further growth is executed.

Transistor and circuit fabrication is based on *mesa* process. In such process a device is constructed from several layers grown on each other. The highest layer is etched to be the smallest in area, next layer is slightly bigger, and so on — like the rings of the Hanoi Towers game. When looking at two adjacent layers, one can think of them as a table resting on a big room's floor. Hence the name *mesa*, which is the Greek word for table.

The entire fabrication process to be mentioned below, i.e. both layer growth and wafer processing, is carried out at the labs and clean rooms of the Microelectronics Research Center, Technion – IIT.

Wafer Layer Structure

Layers are grown bottom to top on a semi-insulator (*SI*) InP 2″ wafer. They are grown lattice matched by a MOMBE (Metal Organic Molecular Beam Epitaxy) system. Layer structure is illustrated in figure 2.8.

Figure 2.8: Layer structure on wafer as grown by MOMBE system

As far as InP-based SHBT is concerned the emitter layer is made of InP and the base, the collector, and the subcollector layers are made of GaInAs. Another GaInAs layer is grown above the emitter layer for better electrical contact to the metal layer deposited later during wafer processing. Finally, an InP cap layer is grown to protect the wafer from moisture and oxidation damage.

All layers and their properties are summarized in table 2.1. Doping level and thickness of each layer are designed to optimize transistor characteristics. The various considerations involved in the design are detailed below.

It could be expected that the emitter layer is highly doped in order to minimize the serial resistance, r_{ee} (figure 2.7b). In fact, this is the case for the upper 300 Å of the emitter. The rest of the layer is moderately doped to reduce base–emitter capacitance, C_{BE} [4]. Note that HBTs don't require high emitter doping level to obtain good emission efficiency (γ).

The base is heavily doped to reduce base resistance r_b. Another way to reduce r_b is increasing the base layer thickness. Since base transport coefficient is proportional to W_B and base transit time is proportional to ${W_B}^2$ (2.16) a compromise should be made between b, τ_B and r_b. A 280 Å layer yields sheet resistance of about 700 $\Omega/\square$ — more than satisfying for the needs of the HBT discussed in this work.

The collector layer is intrinsic up to the growth process cleanness capabilities. Intrinsic collector is essential for maintaining constant C_{BC} over wide range of biasing points. The thickness of the collector determines C_{BC} value as well as collector transit time, τ_C. For optoelectronic monolithic circuits it also sets the quantum efficiency of the photodiode detector, η, as the detector consists of the base and collector layers.

The subcollector is made of the same material as the collector (GaInAs), but has to be

Layer	Thickness [Å]	Composition	Type	Dopant	Doping Concentration [cm^{-3}]
Cap Layer	500	InP	intrinsic	—	—
Emitter Contact	400	GaInAs	N^+	Si or Sn	$5 \cdot 10^{19}$
Emitter	300 2200	InP InP	N^+ N	Si or Sn Si or Sn	$5 \cdot 10^{19}$ $2 \cdot 10^{17}$
Base	280	graded $Ga_{0.53}In_{0.47}As \rightarrow Ga_{0.43}In_{0.57}As$	P^+	C	$\sim 4 \cdot 10^{19}$
Collector	$2000 \div 8000$	GaInAs	N^-	Si or Sn	$\sim 7 \cdot 10^{15}$
Etch Stop	200	InP	intrinsic	—	—
Subcollector	4000	GaInAs	N^+	Si or Sn	$\sim 4 \cdot 10^{19}$
Substrate (Wafer)	500[μm]	InP	semi–insulator	Fe	—

Table 2.1: Layer growth properties

etched by another lithography to create larger mesa. A thin layer of InP enables to stop the collector selective etch by switching to a different etchant. The edges of the subcollector protruding over the collector mesa are the place on which the collector contact is deposited later on. The subcollector is thick and heavily doped to minimize r_c.

At this point wafer is transferred to the clean rooms and wafer processing begins. No additional layer will be grown by the MOMBE system.

Emitter Process

Once the wafer arrived the clean room environment the cap layer is removed. A photo resist (*PR*) is applied to the wafer and patterned by photolithography. Next step is emitter metal deposition. The metal is basically Au (Gold) with thin Pt (Platinum) and Ti (Titanium) layers for better electrical contact. The deposited metal covers the wafer together with the PR that remains on the wafer. Now the PR is removed by means of ultrasonic liftoff, leaving bare areas behind it.

This process leaves $1 \times 10\ \mu\text{m}^2$ rectangular metals on the wafer and the mask that defines them is called "emitter metal". The wafer layers at this level are illustrated in figure 2.9.

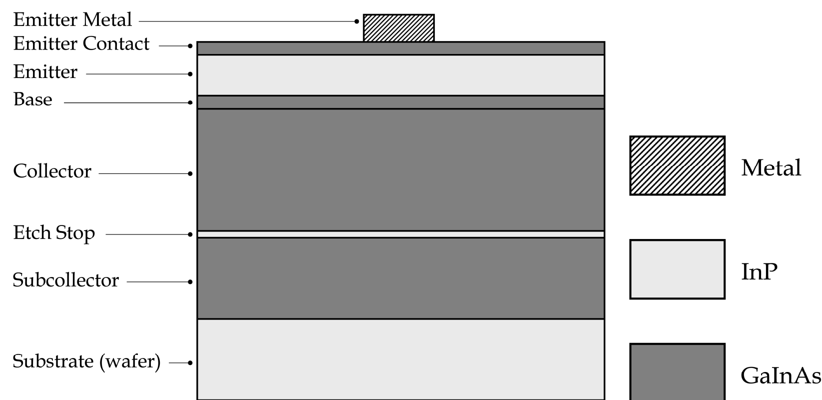


Figure 2.9: Wafer layers after emitter metal deposition and liftoff

Now that emitter metal is done the emitter layers (emitter contact and emitter) are wet etched. No lithography is necessary for this etch since the metal acts as an etch mask. The process allows undercut of the emitter layer resulting an emitter that is slightly smaller than emitter metal. This undercut is critical for the success of the transistor fabrication because of the self–aligned base contact process described later. Figure 2.10 shows the layers at this level.

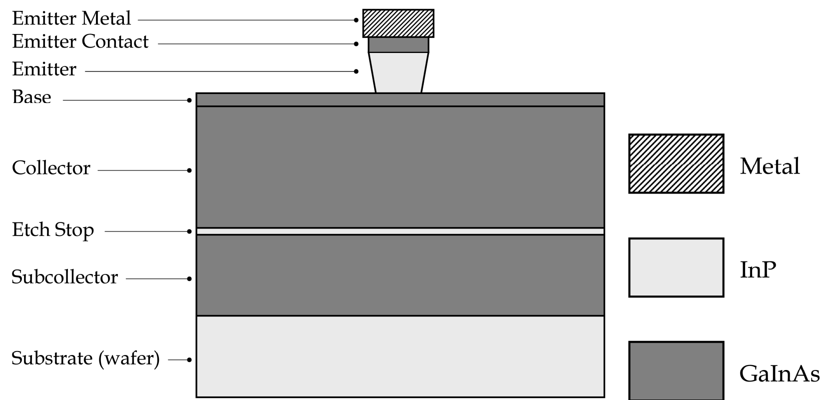


Figure 2.10: Wafer layers after emitter etch

Base and Collector Process

Photo resist is now applied and exposed through "base metal" mask. This mask defines the area of base contact metal. Another metal layer is deposited on the wafer and shaped by liftoff. The base metal area contains the emitter area so that the new metal is heaped up on the emitter metal. The base contact is thus self aligned to the emitter, resulting in low r_b. Layers at this level are shown in figure 2.11.

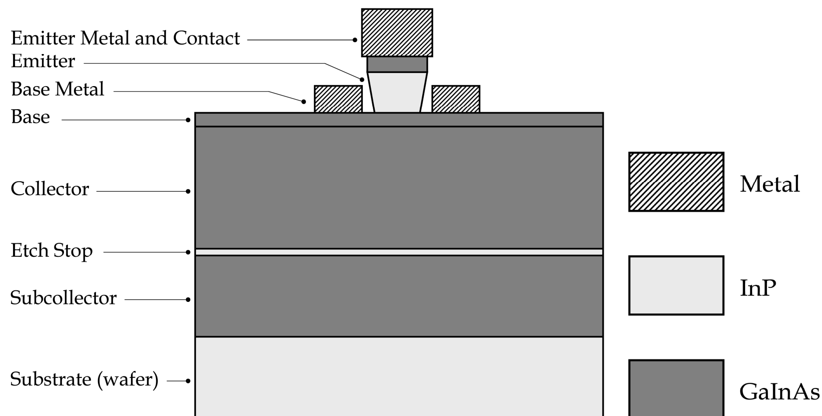


Figure 2.11: Wafer layers after base metal mask

Polyimide is spread on the wafer and dry etched by Oxygen plasma RIE (Reactive Ion Etching) using photolithography of "emitter protect" mask. Now that emitter layer is covered tightly under Polyimide protection base mesa can be created by means of wet etch. The mesa is defined by the base metal and emitter protection, thus no lithography is required. Etch is stopped when it reaches the collector layer by simply counting the

time it takes to the solution to etch a thickness equals to the base thickness (adding 20% to be on the safe side). Layers after base etch are illustrated in figure 2.12.

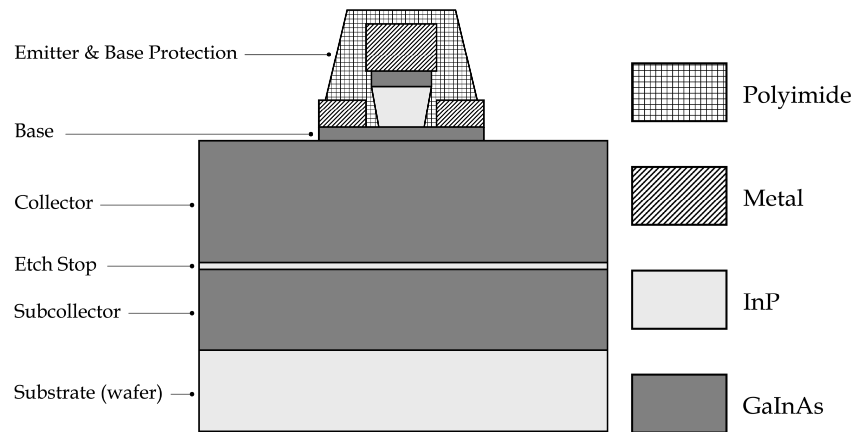


Figure 2.12: Wafer layers after base etch

Next mask defines the collector mesa, and is called "collector protect". The collector (made of GaInAs) is wet etched down to the etch stop layer (made of InP), and then the etch stop is etched by a different selective solution. The collector mesa is slightly wider than the base mesa to enable undercut during collector etch as well as fixing certain process faults. Figure 2.13 shows the layers at this level.

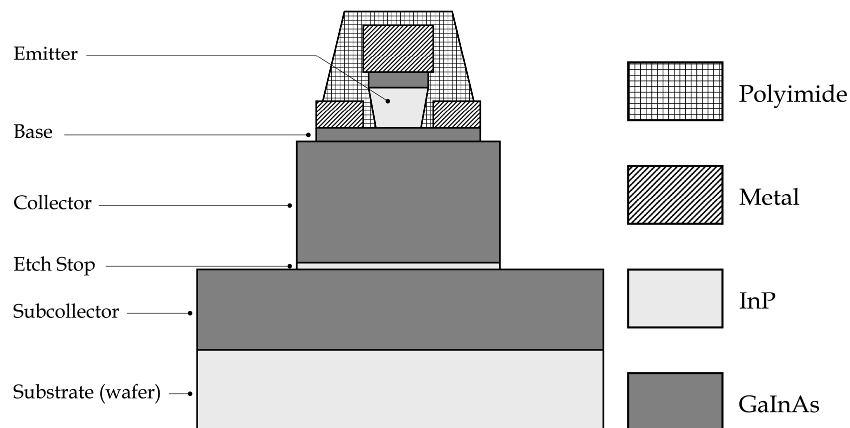


Figure 2.13: Wafer layers after collector etch

Subcollector Process

Metal used for collector contacts is now deposited and shaped by "collector metal" mask. This process leaves metal areas on the subcollector layer for collector contacts. Results can be seen in figure 2.14.

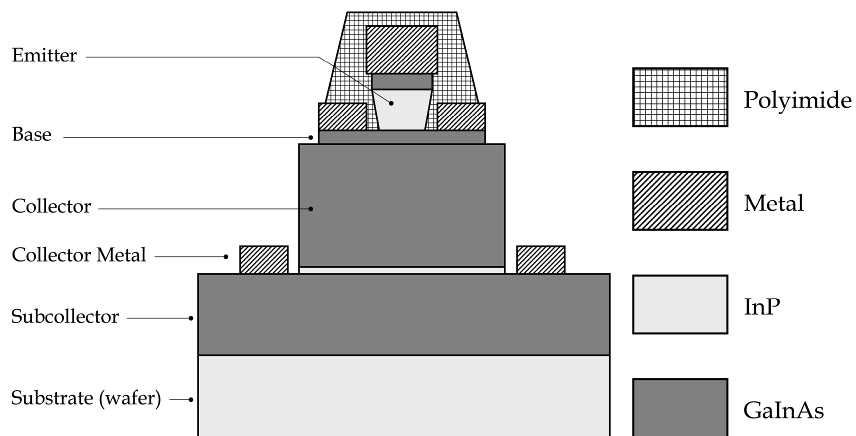


Figure 2.14: Wafer layers after collector metal implementation

At this level the subcollector is etched to isolate the various transistors on the wafer from each other and to expose the insulator wafer, as illustrated in figure 2.15. This is carried out by exploiting the "isolation" mask.

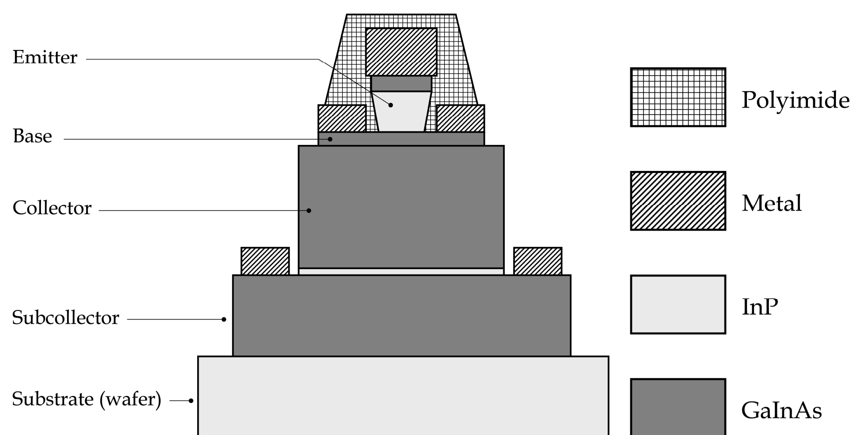


Figure 2.15: Wafer layers after subcollector etch

Contact and Interconnect Process

The contact and interconnect process includes opening transistor connections to the outer circuit and interconnecting the individual components to an integrated circuit. The term *individual components* refers to transistors and capacitors, as well as resistors, inductors, and input/output pads. While the last three are simply geometric shapes made of conducting layers, transistor connections and capacitors require dedicated fabrication methods to obtain the desired results.

Prior to contacts creation the wafer is coated with Polyimide followed by RIE dry etch, using photolithography of "emitter expose" mask. After this lithography the transistor is encapsulated by Polyimide passivation cover except of emitter and collector vias that expose some areas of the emitter and collector metals (hence the name of the mask). Additional mask, called "base–collector via", is required to open base via because of Polyimide topography.

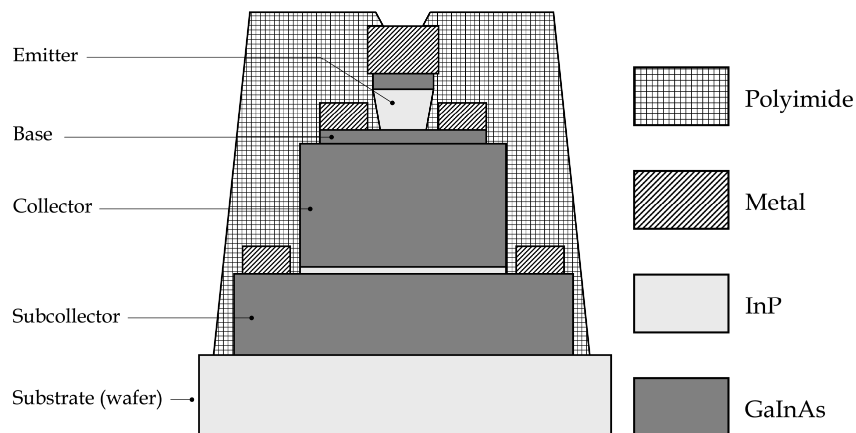


Figure 2.16: Wafer layers after transistor passivation. Only emitter via is shown in this cross–section

Now we come to interconnect. "Metal 1" is the very first mask in this sequence from which most of the interconnect lines, inductors, transmission lines, and pads are produced. Figure 2.17 demonstrates how metal 1 connects to transistor contacts.

In fact, metal 1 is the base layer for capacitors. A capacitor consists of Si_3N_4 dielectric film sandwiched between two metal sheets: metal 1 and metal 2. This Si_3N_4 is deposited by means of PECVD (Plasma Enhanced Chemical Vapor Deposition). Metal 2 is also used for interconnection, but since there's no insulator above metal 1 Polyimide crossover layer is applied in between. Resistors are implemented in an additional NiCr layer. The entire process from metal 1 through metal 2 (excluding the resistor layer) is illustrated in figure 2.18.

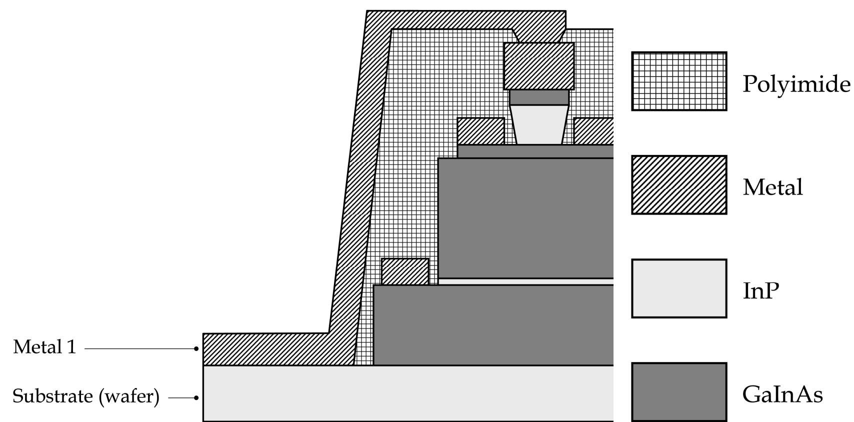


Figure 2.17: Metal 1 as connected to transistor contacts. Only emitter connection is shown in this cross–section

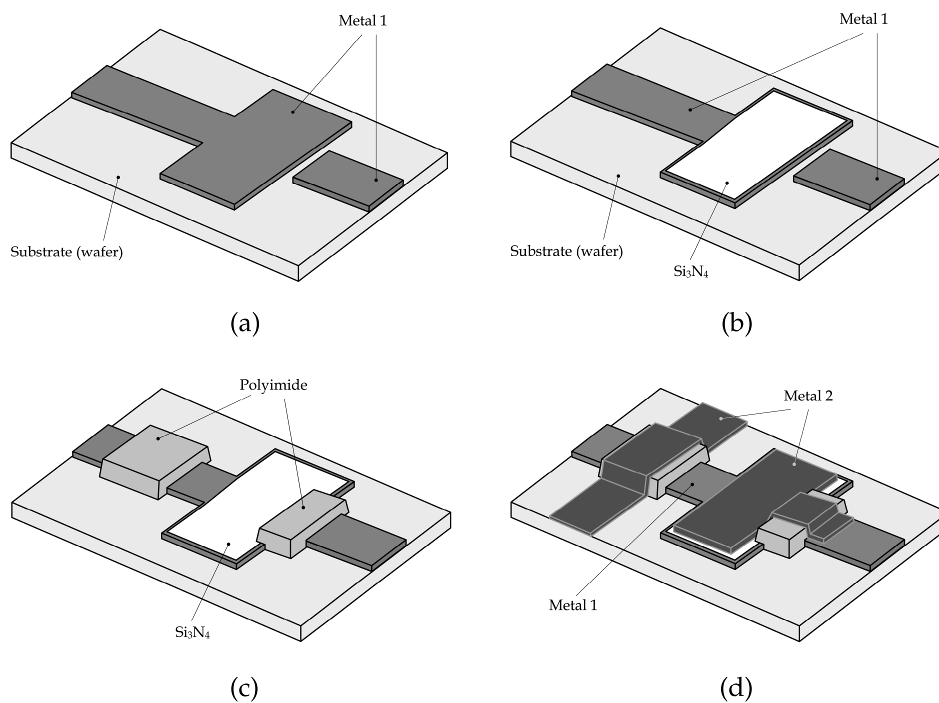


Figure 2.18: Interconnection and capacitor sequence: (a) “metal 1” mask (b) “capacitor” mask (c) “crossover” mask (d) “metal 2” mask

2.2.2 Transistor Geometric Structure

The final structure of a HBT is discussed in this section. Structure is a direct result of the fabrication process described in the previous section. Layer structure, together with

geometric structure, determine transistor's behavior and characteristics which in turn affect the performance of the integrated circuit in which the transistor is embedded.

Before we go on to HBT structure the process is summarized step by step in a table to give an idea about the various levels and what is created in each level. Summary is presented in table 2.2 with special notations for masks and references to the appropriate figures.

Emitter Geometric Structure

The emitter is designed as a long narrow rectangle with dimensions of $1 \times 10\ \mu m^2$. The advantage of this structure is long perimeter for small emitter area, which results low C_{BE} as well as low r_b. A SEM (Scanning Electron Microscope) image of emitter metal with an etched emitter under it is shown in figure 2.19a.

Base Geometric Structure

The base gets its shape from the base metal, therefore both base and base contact constraints are taken into account when base is designed. The base itself must be as small in area as possible to reduce C_{BC}. Being a Miller capacitance in common emitter configuration, C_{BC} is a significant impediment for transistor performance. On the other hand, base metal should include a $5 \times 5\ \mu m^2$ pad to provide enough space for via (step Q of the process, figure 2.20a) and contact with metal 1 (step S). This results in the shape shown in figure 2.19b.

The gap between emitter and base contacts is prone to chemical etch processes and has to be filled with Polyimide. This is done in step H of the process, as shows figure 2.19c.

Collector Geometric Structure

Since C_{BC} value is determined by base area, the collector is created slightly larger than the base (step J); this doesn't affect C_{BC}. Although controlled undercut in collector etch, so that the collector is smaller in area than the base, could reduce C_{BC} [5] — this is not practical for integrated circuits that contain dozens of transistors and demand high yield and reproducibility. Collector geometric structure is shown in figure 2.19d.

Subcollector Geometric Structure

The subcollector has low sheet resistance and contributes nothing to C_{BC}. This is the reason for making the subcollector large enough to contain big collector contacts (step N). The big contacts (step M) ensure good electrical contact without compromising transistor performance. See figure 2.19e.

Step Index	Mask Number	Mask/Process Name	Description	Sketch Figure	SEM Figure
A		Cap Layer Removal	Cap layer removal by wet etch		
B	1	*Emitter metal*	Lithography for emitter contacts		
C		Metal Deposition	Emitter metal deposition and liftoff	2.9	
D		Emitter Etch	Emitter wet etch with undercut	2.10	2.19a
E	2	*Base Metal*	Lithography for base contacts		
F		Metal Deposition	Base metal deposition and liftoff	2.11	2.19b
G		Polyimide Spread	Polyimide spread by spinning for emitter protection		
H	3	*Emitter Protect*	Lithography and Polyimide dry etch to protect emitter zone against base etchant		2.19c
I		Base Etch	Base wet etch	2.12	
J	4	*Collector Protect*	Lithography and collector wet etch		2.19d
K		Etch Stop	Wet etch of etch stop layer	2.13	
L	5	*Collector Metal*	Lithography for collector contacts		
M		Metal Deposition	Collector metal deposition and liftoff	2.14	
N	6	*Isolation*	Lithography and subcollector wet etch	2.15	2.19e
O		Polyimide Spread	Polyimide spread by spinning for transistor protection		
P	7	*Emitter Expose*	Lithography and Polyimide dry etch for transistor protection and isolation; emitter via is opened	2.16	2.19f
Q	8	*Base–Collector Via*	Base and collector via open		2.20a
R	9	*Metal 1*	Lithography for metal 1		
S		Metal Deposition	Metal 1 deposition and liftoff	2.17 2.18a	2.20b
T	10	*Capacitor*	Lithography for capacitor dielectric film		
U		SiN Deposition	Capacitor SiN deposition and liftoff	2.18b	
V		Polyimide Spread	Polyimide spread by spinning for crossover covering		
W	11	*Crossover*	Lithography and Polyimide dry etch for crossover covering	2.18c	
X	12	*Resistor*	Lithography for resistors		
Y		NiCr Deposition	Resistor NiCr deposition and liftoff		
Z	13	*Metal 2*	Lithography for metal 2		
ZZ		Metal Deposition	Metal 2 deposition and liftoff	2.18d	2.20c 2.20d

Table 2.2: HBT fabrication process steps

Interconnect Structure

As the backbone of all connections in a circuit metal 1 is thick enough to be capable of probe gliding and to overcome transistor topography (step S). It should adhere to the

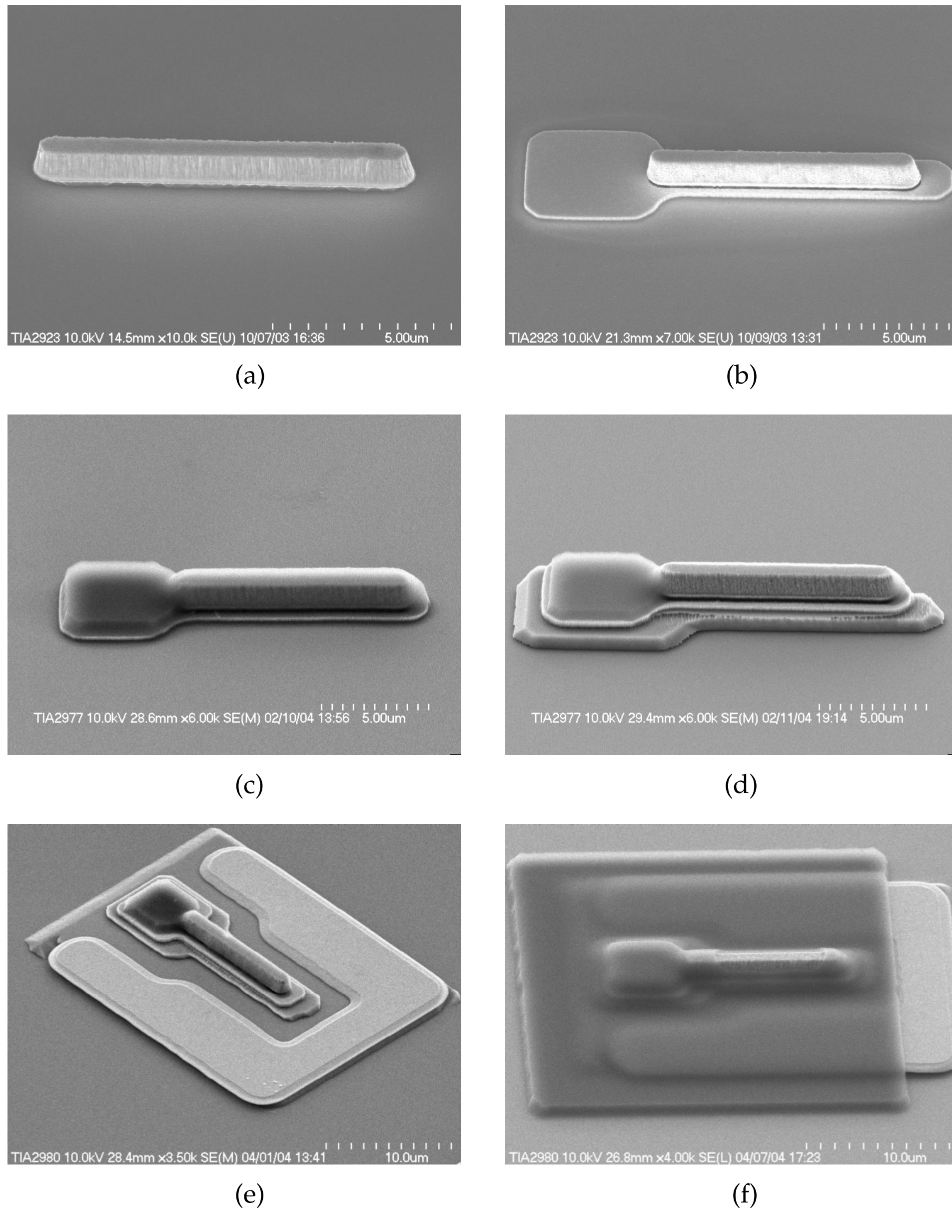
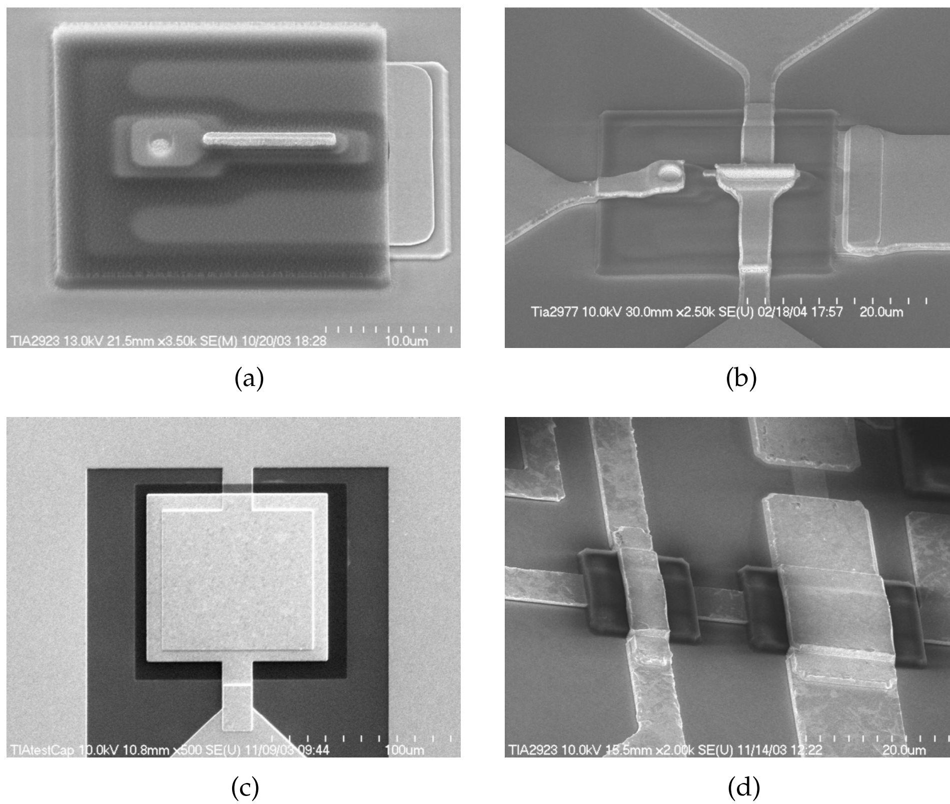


Figure 2.19: SEM images of a transistor at various fabrication process steps: (a) emitter etch (b) base metal deposition and liftoff (c) emitter protect (d) collector protect (e) isolation (f) emitter expose

emitter contact by sinking in its via (figure 2.19f). The same is true for the base contact (figure 2.20a). A completed transistor together with its metal 1 connections is shown in figure 2.20b.

Figure 2.20c shows a capacitor made of metal 1 (step S), SiN (step U), and metal 2 (step ZZ) layers. Contrary to figure 2.18 this capacitor doesn't implement Polyimide to isolate metal 2 contact from metal 1 plate. Alternatively, SiN layer is used for this.

Conductor crossovers with Polyimide bridges (step W) are shown in figure 2.20d. A NiCr resistor (step Y) is evident in the upper right corner of this image.



(a) (b) (c) (d)

Figure 2.20: SEM images of a transistor and circuit at various fabrication process steps: (a) base–collector via (b) metal 1 deposition and liftoff (c) completed capacitor (d) completed interconnects with crossovers and a resistor

2.3 HBT Measurements and Characterization

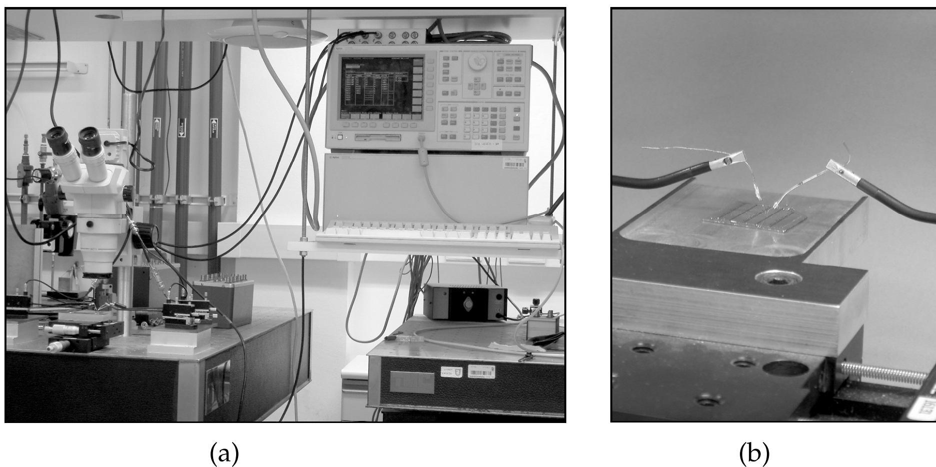
Characterization is the very first step towards construction of a simulation model. It also enables one to ensure that nothing has gone wrong during layer growth and wafer processing. In order to extract the transistor's characters and figures of merit measurements of three categories are carried out: DC, voltage–capacitance, and RF small signal.

2.3.1 DC Measurements

DC measurements are usually performed on large area transistors. The process of large area devices consists of 3 masks and is much shorter than that of small area transistors and circuits, making it an efficient manner of layer growth calibration and other experiments. Obviously, all DC measurements are also taken from small area transistors for accurate characterization.

Measurement Equipment and Setup

The system used for DC measurements of this work includes an Agilent 4155B semiconductor parameter analyzer (*SPA*) and handmade DC probes. A handmade probe is simply a short thin gold wire (Ø 0.5 mil) soldered to a metal pin. Probes are maneuvered by means of 3D micrometer positioners. A picture of the measurement system is shown in figure 2.21a, and a closeup on handmade DC probes in figure 2.21b.



(a) (b)

Figure 2.21: DC measurement setup: (a) measurement system (b) handmade probes

Measurement results are stored in text files and transferred to a PC, and then read from files and processed by MathWorks' MATLAB 6.5 software.

Junction Measurement

Measuring I–V curves of BC and BE junctions is an initial test that can reveal the occurrence of defects in a transistor. In junction measurement voltage sweeps from reverse bias to forward bias, and the current flowing in the junction is plotted versus the voltage. Initial glance at an I–V curve of a junction can indicate whether the junction behaves like a diode at all or severe current leakage occurs. A typical I–V curve of PN–junction is illustrated in figure 2.22.

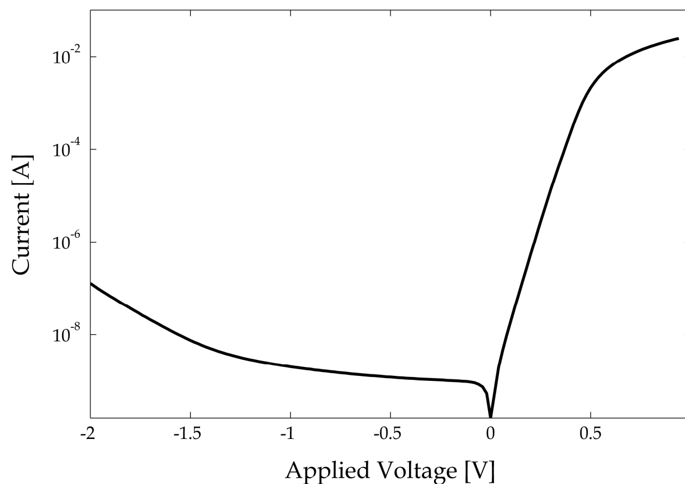


Figure 2.22: I–V curve of a PN–junction (measured on BC junction of a large area device)

As can be observed in the positive bias zone of figure 2.22, I–V curve rises exponentially up to 0.5 V, then the serial resistances of the junction layers and pads become dominant and the curve bends. In the reverse bias zone slight leakage is observable. This leakage occurs due to mesa edge imperfections besides other reverse conductance mechanisms. The leakage is easily diagnosed by comparing two sequential measurements: mesa edge leakage is inconsistent and therefore yields different curves, whilst bulk mechanisms result in identical curves.

Common Emitter Measurement

In common emitter measurement I_B is set to constant value and V_{CE} is swept from 0 to sufficient positive voltage. This measurement realizes the current gain aspect of tran-

sistor operation. For low V_{CE} both BC and BE junctions are forward biased, hence the transistor is in saturation mode. As V_{CE} rises the transistor enters forward active mode and I_C is determined by the supplied I_B and β_F of the transistor, regardless of V_{CE} (equation (2.9)).

The procedure mentioned above is repeated for several I_B values to ensure proper behavior under wide variety of biasing points. Moreover, increasing I_B in fixed steps should yield fixed gaps between the various curves, as demonstrated in figure 2.23.

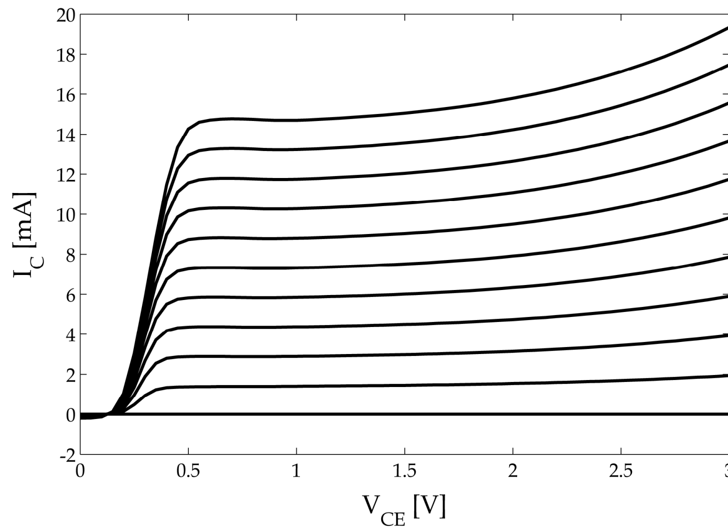


Figure 2.23: Common emitter measurement curves (measured on a large area device). I_B varies from 0 to 200 μA in 20 μA steps

According to equation (2.9) one could expect I_C curves to be straight and horizontal in forward active mode. In practice, the curves are convex due to avalanche base–collector breakdown. Note that Early effect is negligible in HBT thanks to the high doping level of the base. Also changes in r_c aren't something to worry about because BC junction is a PIN–diode with constant depletion region width. Should Early effect occur in a BJT, its common emitter curves would bend even when no avalanche takes place.

Common Base Measurement

Common base measurement is useful for exploring avalanche base–collector breakdown, as explained in the next section. I_E, the input current, is set to constant value and I_C is plotted versus V_{CB}. Repeating the above for several I_E values yields the curves shown in figure 2.24. Once I_E and I_C are known, extraction of I_B and α_F is apparent.

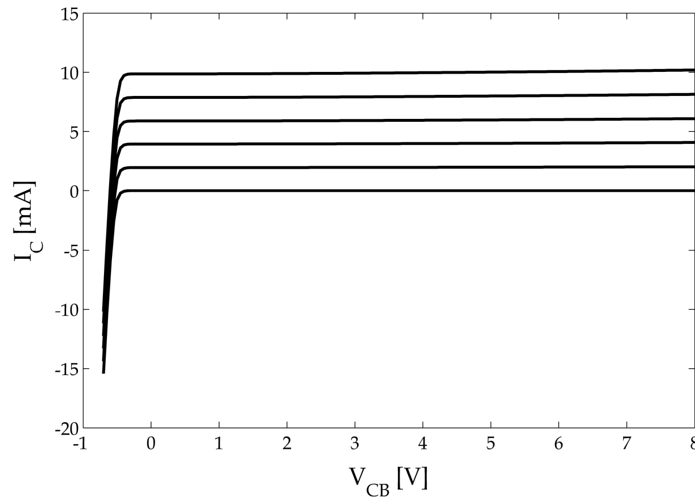


Figure 2.24: Common base measurement curves (measured on a large area device). I_E varies from 0 to 10 mA in 2 mA steps

Gummel Plot Measurement

In Gummel plot measurement BC junction is shorted and V_{CE} ($= V_{BE}$) is swept over a forward bias range (in BE junction terms). I_B and I_C are plotted versus V_{CE} to learn about how current is distributed between the base and the collector, i.e. what is β_F value. Typical Gummel plot curves are illustrated in figure 2.25.

Sheet Resistance Measurement

Sheet resistance, as well as pad and contact parameters, are extracted from transfer length method (*TLM*) measurement. Rectangular pads are ordered in a straight line with different gaps between each other. The resistance between two adjacent pads is measured for every pair. TLM measurement and theory will be explained in detail in section 2.3.2.

2.3.2 DC Characterization

As mentioned before, data of DC measurements is stored in text files and processed by MATLAB. This section describes the process of DC parameter extraction and characterization of a transistor.

From junction measurements ideality factor, n, and saturation current, I_0, are derived for both BE and BC junctions. Common emitter measurement contributes turn-on voltage, V_{TO}, Early voltage, V_A, and β_F dependency on temperature. V_{CEO}, the value of V_{CB}

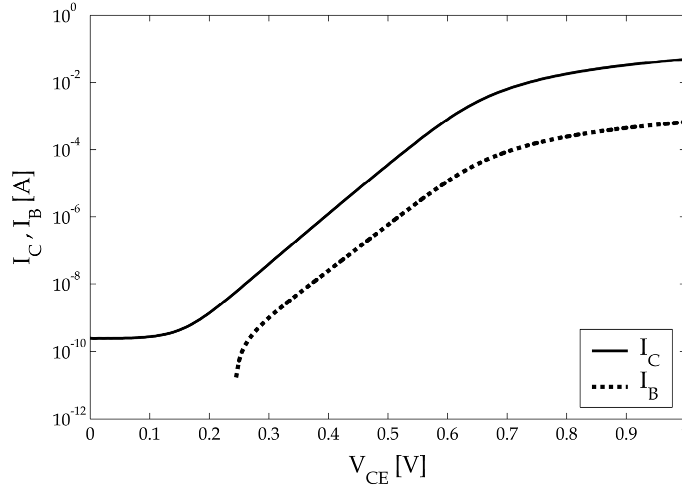


Figure 2.25: Gummel plot curves (measured on a large area device)

at which avalanche base–collector breakdown brings α_F to be equal to 1, is derived from common base measurement. Finally, Gummel plot measurement helps to get information about β_F, the forward current gain, and β_{AC}, the small signal differential current gain.

Junction Ideality Factor and Saturation Current

The I–V curve of a PN–junction consists of the following equation [2][1]:

$$I = I_0 \cdot \left(e^{\frac{qV}{nKT}} - 1 \right) \tag{2.19}$$

where I_0 is the saturation current of the junction and n is its ideality factor. Plotting $\ln(I)$ versus V yields linear curve in the positive region of V, as the unity becomes negligible in comparison to the exponential term:

$$\ln(I) = \ln(I_0) + \frac{qV}{nKT} \tag{2.20}$$

As evident in (2.20), this linear curve has a slope of $\frac{q}{nKT}$, and it intercepts the I axis at $\ln(I_0)$ (refer to figure 2.26).

Forward Current Gain

Forward current gain, β_F, can be extracted either by processing common emitter measurement data (dividing I_C by I_B) or from Gummel plot data. Gummel plot data is processed as follows: A couple of curves are calculated from the given I_C and I_B data

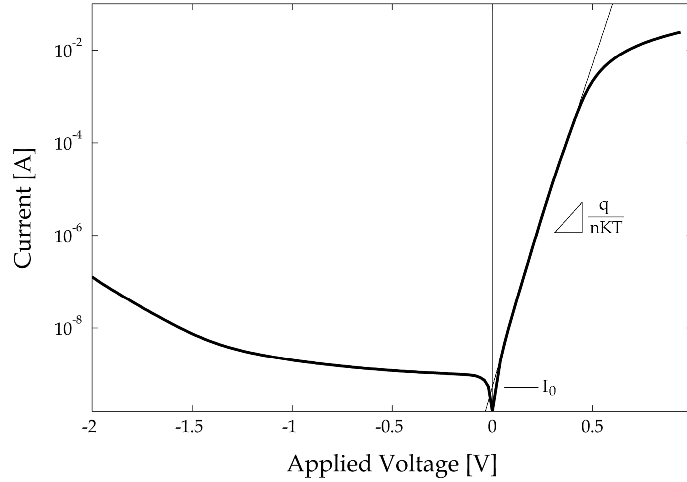


Figure 2.26: Extraction of I_0 and n from I–V curve of a PN–junction

according to the following relationships:

$$\beta_F = \frac{I_C}{I_B} \tag{2.21}$$

$$\beta_{AC} = \frac{\partial I_C}{\partial I_B} \tag{2.22}$$

Equation (2.21) consists directly on (2.9), while (2.22) is a differential version for small signal current gain. Since BE junction is forward biased in forward active mode, one can assume that $V_{BE} \approx 0.9$ V (at least in small area devices). Hence, the most significant results of β_F and β_{AC} are those of $V_{BE} = 0.9$V. Refer to figure 2.27; keep in mind that in Gummel plot measurement $V_{CE} = V_{BE}$.

Common Emitter Turn–On Voltage

The turn–on voltage, V_{TO}, is the V_{CE} value at which the transistor starts to conduct positive collector current (according to the notations on figure 2.3) in common emitter configuration. Obviously, this can be observed in the common emitter curves, as shown in figure 2.28.

Avalanche Breakdown Voltage

The occurrence of avalanche base–collector breakdown is evident both in common emitter and common base measurements. As the avalanche is generated by the electrical field in BC depletion region, it has nothing with Gummel plot measurement, in which $V_{CB} = 0$.

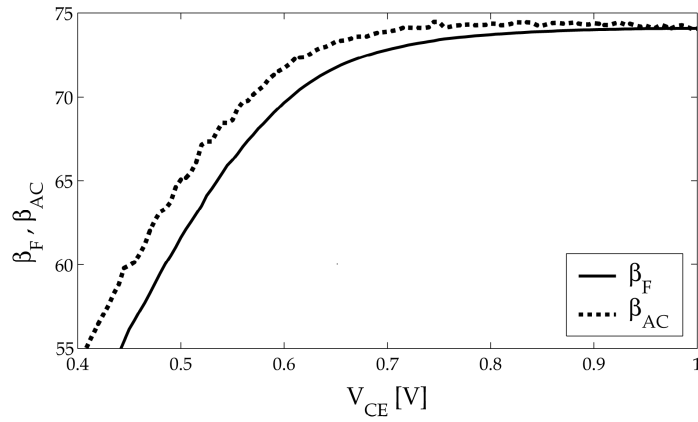


Figure 2.27: β_F and β_{AC} curves — calculated from Gummel plot data (measured on a large area device)

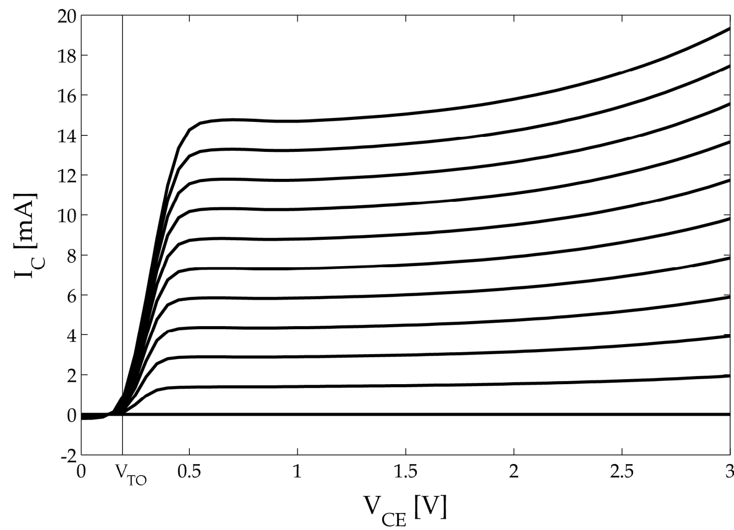


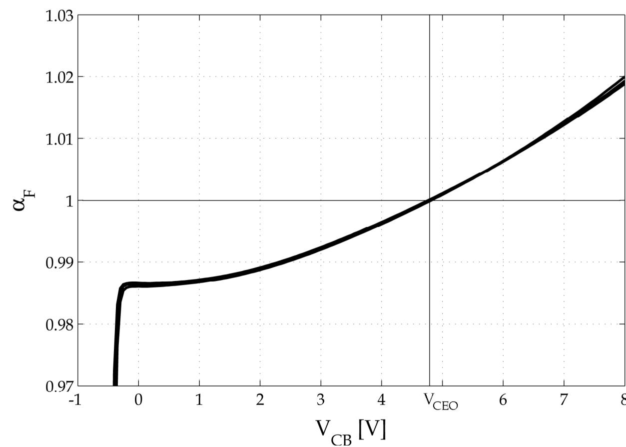
Figure 2.28: V_{TO} marked on common emitter measurement curves (measured on a large area device)

As mentioned above, β_F can be calculated from common emitter data by simply dividing I_C by I_B on every measurement point. According to equation (2.12), V_{CE} is limited as follows:

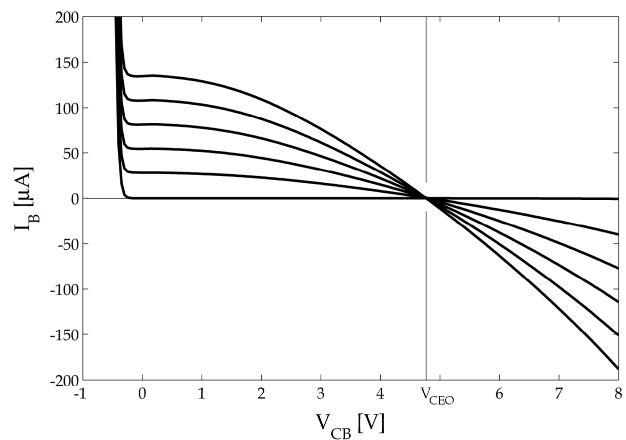
$$V_{CE} < V_{CEO} - 0.9\text{V} \tag{2.23}$$

to avoid I_C from extremely high values. Any attempt to apply too high V_{CE} will end up with burnt-out transistor. Although this has great importance in practice, it doesn't provide much information about avalanche base–collector breakdown and V_{CEO}.

A better method for learning avalanche breakdown is by manipulating common base measurement data. There are two parameters that can be derived directly: α_F, by dividing I_C by I_E, and I_B, by subtracting I_C from I_E. V_{CEO} is defined by the value of V_{CB} at which $\alpha_F = 1$ (or $M \cdot \alpha_0 = 1$), therefore the voltage at which α_F curves intercept $\alpha_F = 1$ line is what we are looking for, as presented in figure 2.29a. In addition, one concludes from equation (2.11) that I_B curves intercept $I_B = 0$ line at V_{CEO}. All interceptions occur at the same voltage because avalanche base–collector breakdown depends on V_{CB}; this is shown in figure 2.29b.



(a)



(b)

Figure 2.29: V_{CEO} marked on (a) α_F and (b) I_B curves (measured on a large area device)

Base Sheet Resistance and Pad Parameters

TLM is used for characterizing both base sheet resistance and emitter contact resistance. A series of measurements is performed on $Z \times L$-sized pads organized in a row, as illustrated in figure 2.30. The result is a list of the resistances measured between each pair of

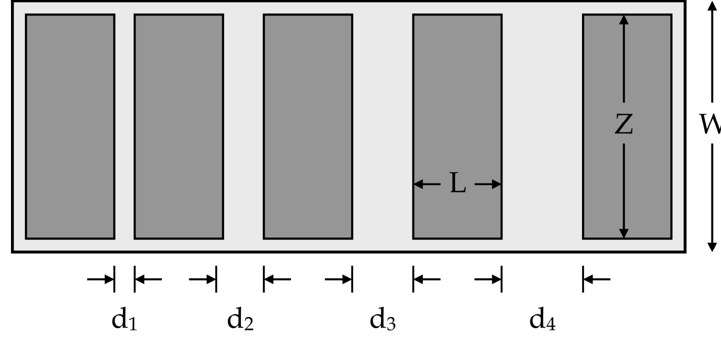


Figure 2.30: TLM test fixture

adjacent pads. The total resistance between 2 pads is given by

$$R_T = \rho_s \frac{d_i}{Z} + 2R_c \tag{2.24}$$

where ρ_s is the layer sheet resistance, d_i is the gap between the i-th pads pair, Z is the pad's width, and R_c is the resistance of each contact. It is apparent from (2.24) that plotting R_T versus d_i yields a linear curve with a slope of $\frac{\rho_s}{Z}$. This curve intercepts the R_T axis at $2R_c$.

The current paves its way across the layer, from one pad to another. It is emitted from the edge of the source pad facing towards the target pad. At the target side, the current enters the pad through the very symmetric edge. Consider the route of current in figure 2.31. The current concentrates across the edge in a L_T–wide strip (Actually, the current has an exponential concentration profile near the edge with a characteristic length of L_T). For $Z \approx W$ (alternatively, for $Z \gg d_i$) and $L \gg L_T$ one obtains [6]:

$$R_c \approx \rho_s \frac{L_T}{Z} \tag{2.25}$$

Using (2.25) in (2.24) gives

$$R_T = \rho_s \frac{d_i}{Z} + 2\rho_s \frac{L_T}{Z} \tag{2.26}$$

It follows that the previously mentioned R_T versus d_i curve intercepts the d_i axis at $-2L_T$. Figure 2.32 summarizes the extraction of ρ_s, R_c, and L_T from a TLM curve. The trend line presented in this figure was derived from the measured TLM data by means of linear regression.

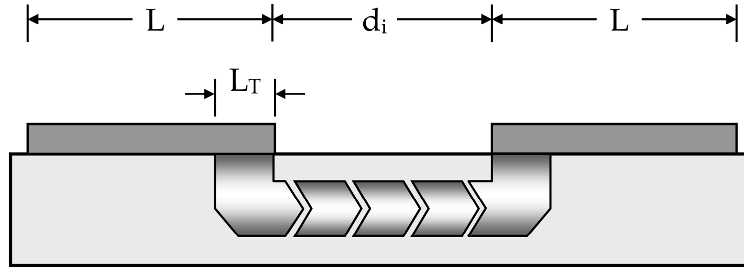


Figure 2.31: Current route in TLM measurement

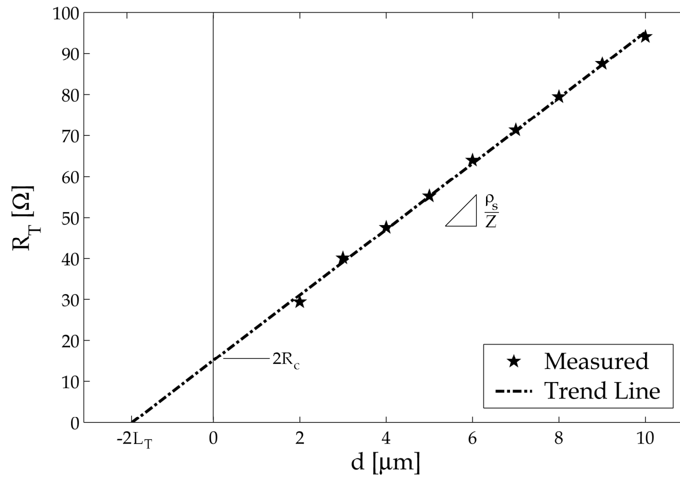


Figure 2.32: Extraction of sheet resistance, contact resistance, and transfer length from a TLM curve (measured with $Z = 100\ \mu$m, $L = 70\ \mu$m)

Sheet resistance, ρ_s, has great importance for the base resistance, r_b, but not for the emitter as the current flows in the emitter vertically; contact resistance, R_c, is crucial both for the emitter and for the base; and L_T, the transfer length, is a critical parameter in the design of the base contact. Note that the emitter contact resistance is given by

$$R_{E,contact} = L_T Z \frac{R_c}{A} \tag{2.27}$$

where A is the area of the emitter contact, and $L_T Z \cdot R_c$ is the pad resistance, usually denoted by R_{pad}.

2.3.3 Voltage–Capacitance Measurement and Characterization

The voltage–capacitance measurement (also referred to as *C–V measurements*) is essential for analyzing doping concentration profile and estimating the capacitance of a small area device for its simulation model. Knowing the doping concentration profile is essential for layer growth control and it provides useful information about which voltage is required for depleting the entire collector and achieving minimal C_{BC}.

Measurement Equipment and Setup

The C–V measurement system consists of a Hewlett Packard 4280A C–V Plotter and a PC, connected to each other by means of GPIB interface. The measured PN–junction is probed by handmade probes (just like the ones used for DC measurements). All HP 4280A operations and data acquisition are done remotely from the PC using a dedicated program written in MATLAB. This program takes advantage of the MATLAB capabilities to provide GUI and handle GPIB connection (The program will be described in detail in appendix A). Acquired data is saved in text files and processed by MATLAB. A picture of the measurement system is shown in figure 2.33.

C–V Measurement

In C–V measurement bias is applied to a junction and capacitance is measured. The instrument applies 1 MHz small signal voltage to the tested device and reads the magnitude and phase of the current generated by this signal. Then the instrument derives the capacitance and conductivity between its probes from the imaginary and real parts of the current, correspondingly. Due to the unavoidable series resistance, the accuracy

Figure 2.33: C–V measurement system

of this measurement is limited by the conductivity parallel to the capacitance. Low conductance is essential for precise results. Hence, PN–junction measurement yields reliable results only when the junction is reverse biased.

PN–junctions are usually measured over a range of voltages. Voltage is swept by the HP 4280A instrument and capacitance and conductivity are measured and saved in text files. This results the curve shown in figure 2.34a. In addition, the conductance G is compared to the imaginary admittance $j\omega C$ to ensure dependable results. Admittance ratio is shown in figure 2.34b and is expected to be $\ll 1$.

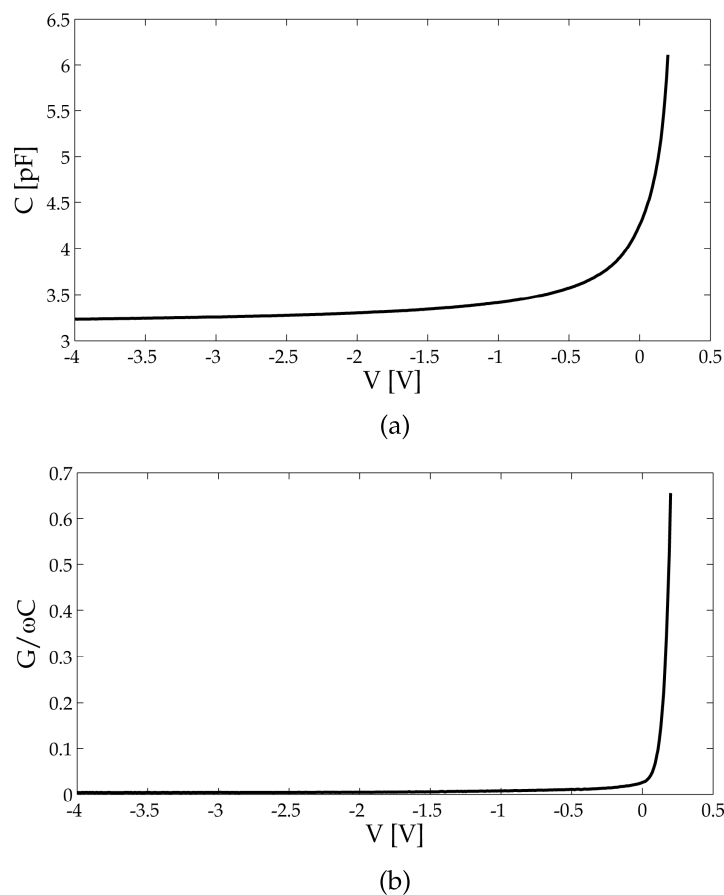


Figure 2.34: (a) C–V curve and (b) admittance ratio (measured on a large area device)

Extraction of Collector Doping Profile

C–V measurement is a powerful tool to investigate PN–junctions. It enables one to know what voltage is required for achieving specific capacitance. Furthermore, the doping

profile can be extracted from the data for better layer growth control. Firstly, junction width d as a function of applied voltage V should be calculated. This is done by using the following equality [2]:

$$d(V) = \epsilon_0 \epsilon_r \frac{A}{C(V)} \tag{2.28}$$

where ϵ_0 is the vacuum dielectric coefficient, ϵ_r is the semiconductor dielectric coefficient, and A is the junction area. Now that junction width is known doping profile at the lower doped side of the junction is given by [2]

$$N_d(V) = \frac{2}{A^2 \epsilon_0 \epsilon_r q \cdot \frac{\partial\left(\frac{1}{C^2(V)}\right)}{\partial V}} \tag{2.29}$$

N_d in (2.29) is the donor concentration at the lower doping side. Since in NPN-HBT the base has the highest doping level, only collector and emitter doping profiles can be extracted — and both are n-type semiconductors. BC junction is a one-sided junction, hence d also represents position along the collector (where $d = 0$ is the metallurgic junction). A doping profile curve derived from C–V measurement data is shown in figure 2.35. This curve was measured on BC junction of a large area transistor and was used for layer growth calibration.

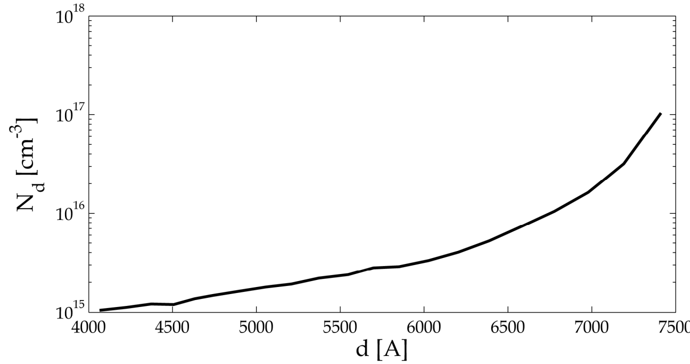


Figure 2.35: Doping profile of a one-sided PN–junction (measured on a large area device)

2.3.4 RF Measurements

RF measurements test transistor behavior in the frequency domain. Transistor biasing point is set to a desired value prior to the measurement, and then measurement is carried out. RF measurement results, together with DC and C–V results, provide sufficient information for constructing a simulation model of a transistor. Measurement procedure is detailed below.

Measurement Equipment and Setup

Measurement system consists of Hewlett Packard 8722C network analyzer, Karl Suss probe station, Picoprobe 40A-GSG-150-P RF probes, and Picoprobe CS-5 calibration substrate. This comes up to the system shown in figure 2.36.

Figure 2.36: RF measurement system

S–Parameter Measurement

For RF measurement purposes transistor is referred to as a two-port network. It is connected in common emitter configuration, forming the network illustrated in figure 2.37. The network analyzer measures the scattering matrix (*S–parameters*) of the two-port network over a range of frequencies. When measurement has finished data is acquired to a PC via GPIB connection and stored in a text file.

The transistor is biased by means of either DC power supplies or SPA's outputs, set-

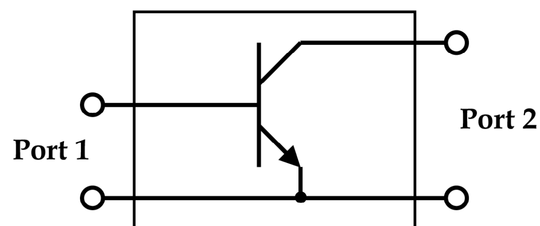


Figure 2.37: Transistor as a two-port network in common emitter configuration

ting I_B and V_{CE} to the desired biasing point. DC is fed through the internal bias tees of the HP 8722C network analyzer.

Prevalently a set of S–parameter measurements at different biasing points is necessary. This is also done by controlling the entire measurement system by a PC, using MATLAB or National Instruments' LabView software.

2.4 Extraction of HBT Small Signal Model

Small signal model is extracted from the S–parameter data for each individual biasing point. The model used for extraction in this work is the T model [7], whose components have been described previously (in section 2.1.2. See also figure 2.7b). A full T model is shown in figure 2.38 in common emitter orientation. The extraction process begins

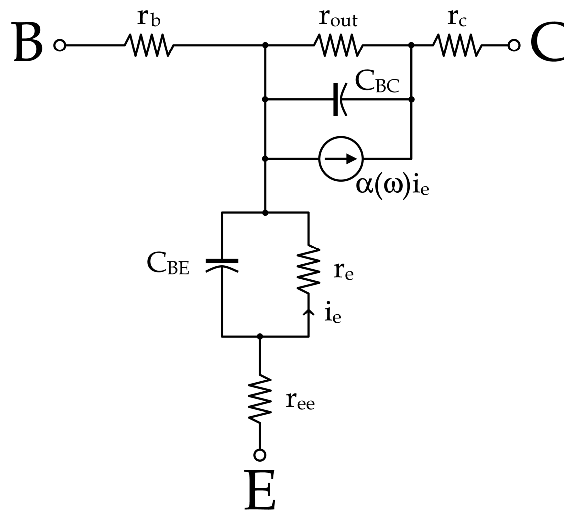


Figure 2.38: Small signal T model of a HBT

with the parasitic capacitances and inductances of the probing pads. Once these are known they are peeled off the data so that only the transistor's S–parameters remain. Then the collector, emitter, and base elements are extracted. Next step is the extraction of the forward transit time, τ_D, current gain cutoff frequency, f_T, and Mason's power gain cutoff frequency, f_{MAX}. The entire extraction process is detailed below step by step.

2.4.1 Extraction and Peeling of Parasitic Pad Elements

In order to be able to measure a transistor, it has to be surrounded with pads. The pads, made of *metal 1* layer, are designed to have a characteristic impedance of 50 Ω to avoid reflections from the measurement system (these subjects will be discussed later in chapter 5). The pads can be modelled as a set of shunt capacitors and serial inductors lumped in with the transistor in the two-port network measured by the network analyzer. Figure 2.39 shows the equivalent circuit of the pads. As S–parameter measurement includes the pad effect, it should be eliminated from the data. For this purpose two additional test fixtures are fabricated on the wafer — *open* and *through* fixtures — which are geometrically similar to the pads attached to the transistor (figure 2.40). S–parameters

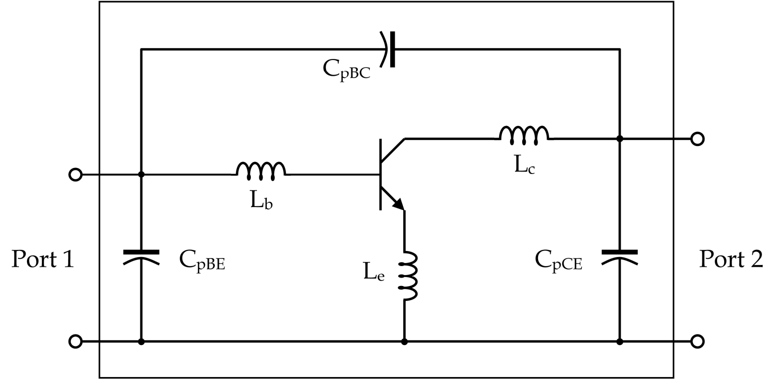
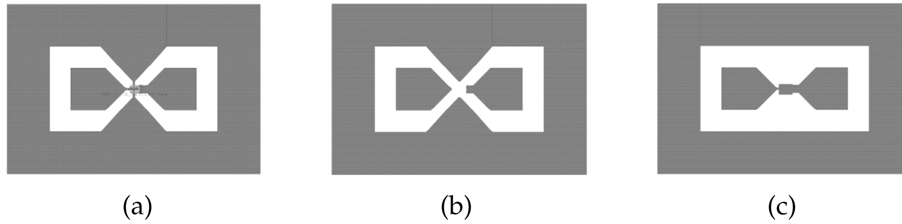


Figure 2.39: A HBT with pad parasitic components lumped in a two-port network



(a) (b) (c)

Figure 2.40: Layout of (a) HBT probing pads (b) open test fixture (c) through test fixture

measured on those fixtures provide enough information to derive 5 out of the 6 parasitic elements.

The through fixture pads can be modelled as two shunt capacitors (C_{pBE} and C_{pCE}) and a serial inductor (L_{thru}). Using its S–parameters and converting them to Y–parameters, it can be shown that [8]:

$$L_{thru} = -\frac{1}{\omega} \cdot \mathrm{Im}\left(\frac{1}{Y_{21}}\right) \tag{2.30}$$

$$C_{pBE} = \frac{1}{\omega} \cdot \mathrm{Im}\left(Y_{11} + Y_{21}\right) \tag{2.31}$$

$$C_{pCE} = \frac{1}{\omega} \cdot \mathrm{Im}\left(Y_{22} + Y_{12}\right) \tag{2.32}$$

The similarity between the test fixture and transistor's pads yields the following equality:

$$L_{thru} = L_b + L_c \tag{2.33}$$

where the ratio between L_b and L_c can be derived from the geometrical shape of the contacts. For the transistors and pads used in this work assuming that $L_b \approx L_c$ is adequate.

The model of the open fixture includes only capacitors — C_{pBE}, C_{pCE} and C_{pBC}. Since the shunt capacitances are known, the latter is given by either

$$C_{pBC} = \frac{1}{\omega} \cdot \mathrm{Im}\left(Y_{11}\right) - C_{pBE} \tag{2.34}$$

or

$$C_{pBC} = \frac{1}{\omega} \cdot \mathrm{Im}\left(Y_{22}\right) - C_{pCE} \tag{2.35}$$

Actual measurements show that C_{pBC} has a very low value and can be neglected in most cases. Typical values are 25 fF for C_{pBE} and C_{pCE}, and 3 fF for C_{pBC}.

Now that capacitor values are known, they can be peeled off the two-port network data by using the following relations (yet neglecting C_{pBC}):

$$y_{11} = Y_{11} - j\omega C_{pBE} \tag{2.36}$$

$$y_{12} = Y_{12} \tag{2.37}$$

$$y_{21} = Y_{21} \tag{2.38}$$

$$y_{22} = Y_{22} - j\omega C_{pCE} \tag{2.39}$$

y_{nn} denote the Y–parameters of the two-port network after peeling.

2.4.2 Extraction of Collector Elements

From now on the extraction procedure follows the work of S. J. Spiegel *et al* [9]. C_{BC} and r_b are assumed to be lumped rather that distributed elements to keep the model as simple as possible. In addition, L_b and L_c values are known at this level, thus no optimization is necessary (contrary to [9]). The extraction procedure makes use of the Z–parameter representation so Y–parameters are converted to Z.

C_{BC} is extracted from the flat zone of the graph of the following expression:

$$C_{BC} = -\frac{1}{\omega} \cdot \frac{1}{\mathrm{Im}\left(z_{22} - z_{21}\right)} \tag{2.40}$$

At low frequencies a capacitor acts as an open circuit which results in an inaccurate noisy extraction. As frequency increases the graph converges to the measured capacitance value and becomes "flat", then at high frequencies inductive elements get into the picture (see figure 2.41). r_c cannot be extracted and is thus estimated to be 2.5 Ω — a negligible resistor for most circuits. Since Early effect doesn't occur in HBTs we assume that $r_{out} \longrightarrow \infty$. A value of 30 kΩ can be considered as an open circuit for most simulation purposes.

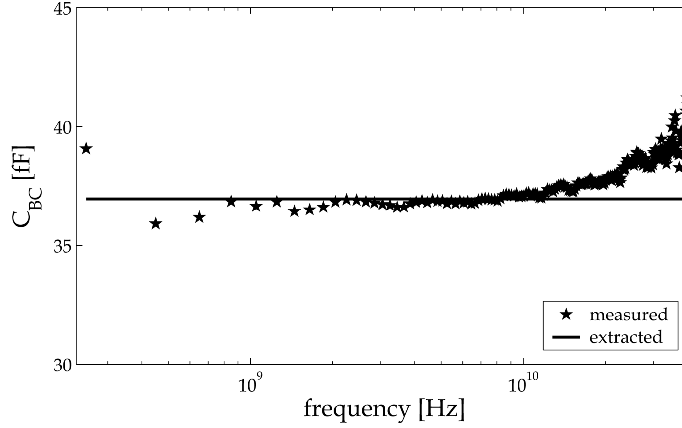


Figure 2.41: Extraction of C_{BC}

2.4.3 Extraction of Emitter Elements

Emitter elements include L_e, r_e, r_{ee}, and C_{BE}. L_e is extracted from the flat zone of the graph of the following expression:

$$L_e = \frac{1}{\omega} \cdot \mathrm{Im}\,(z_{12}) \tag{2.41}$$

Here the flat zone exists at high frequencies, where capacitive influences become negligible (refer to figure 2.42a). The total emitter resistance, $r_e + r_{ee}$, is derived by using the following equality:

$$r_e + r_{ee} = \mathrm{Re}\,(z_{12}) \tag{2.42}$$

This curve is flat at low frequencies. At high frequencies the inductive elements dominate z_{12}, as illustrated in figure 2.42b. The above result can be separated into r_e and r_{ee} by simply evaluating r_e using

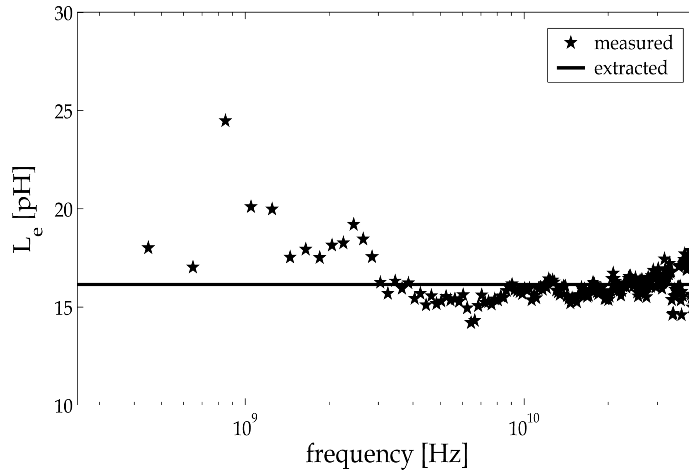
$$r_e = \frac{qI_E}{KT} \tag{2.43}$$

The last element, C_{BE}, is estimated according to the junction width and area.

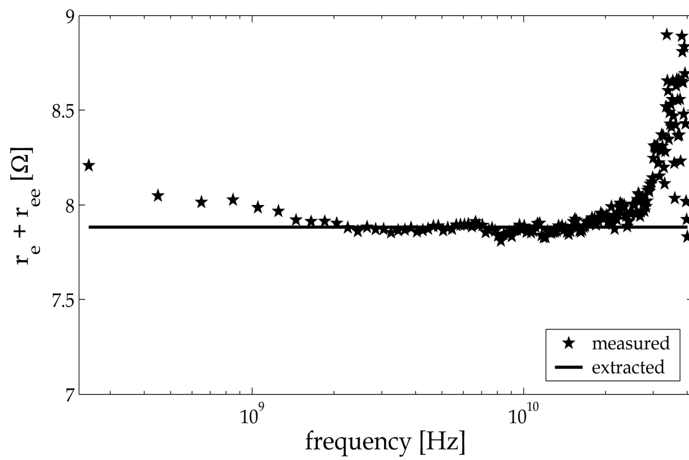
2.4.4 Extraction of Base Resistance

As evident in figure 2.38, r_b is the sole element existing in the base part of the T model. Base resistance has great importance in governing transistor AC performance in common emitter configuration, because it is connected in serial to a Miller capacitor (this issue will be discussed in detail in chapter 4). The fact that r_b is connected to a high value element (Miller capacitance) results in a noisy extraction, as evident in figure 2.42c. However, the resistance is given by:

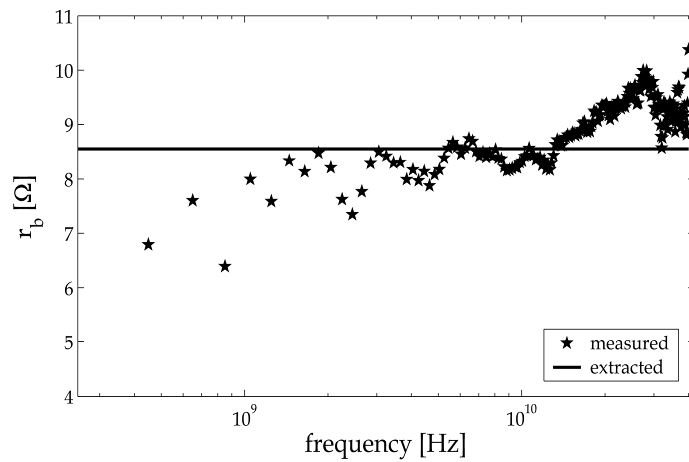
$$r_b = \mathrm{Re}\,(z_{11} - z_{12}) \tag{2.44}$$



(a)



(b)



(c)

Figure 2.42: Extraction of (a) L_e (b) $r_e + r_{ee}$ (c) r_b

Because of the large capacitance involved in this extraction, as well as the inductances, the graph of r_b versus frequency is flat in the intermediate frequencies.

2.4.5 Extraction of Forward Transit Time

In terms of T model, the forward transit time, τ_D, is hidden in the collector current source. It can be expressed by either a pole or phase element in $\alpha(\omega)$, namely

$$\alpha(\omega) = \left(\alpha_0 \cdot e^{j\omega\tau_D}\right) \tag{2.45}$$

The transit time itself is given by

$$\tau_D = -\frac{1}{\omega} \cdot \text{Im}\left(\frac{z_{12} - z_{21}}{z_{22} - z_{21}}\right) - (r_e C_{BE} + r_c C_{BC}) \tag{2.46}$$

Also here extraction is valid in the intermediate frequencies, as shown in figure 2.43. It should be noted that the term $r_c C_{BC}$ in (2.46) is based on the estimation of r_c. However, the term $r_e C_{BE}$ usually dominates the right parenthesis.

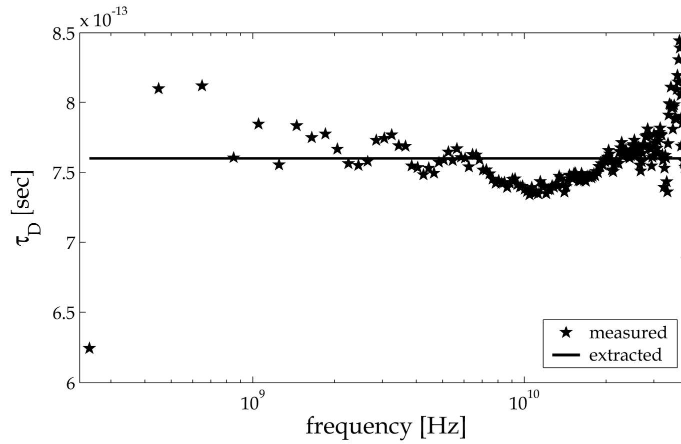


Figure 2.43: Extraction of τ_D

2.4.6 Extraction of Current Gain Cutoff Frequency

At this level pad inductances are peeled off the two-port parameters as follows:

$$Z_{11} = z_{11} - j\omega(L_e + L_b) \tag{2.47}$$
$$Z_{12} = z_{12} - j\omega L_e \tag{2.48}$$
$$Z_{21} = z_{21} - j\omega L_e \tag{2.49}$$
$$Z_{22} = z_{22} - j\omega(L_e + L_c) \tag{2.50}$$

Then Z–parameters are converted to h–parameters. Current gain is given by h_{21}, and the current gain cutoff frequency, f_T, is the frequency at which h_{21} becomes unity. f_T is usually higher than measurement limits, so it can be evaluated by either extrapolating h_{21} in a –20 dB/dec slope or by fitting a single-pole function to the measured data. Figure 2.44 demonstrates these two fashions of f_T extraction.

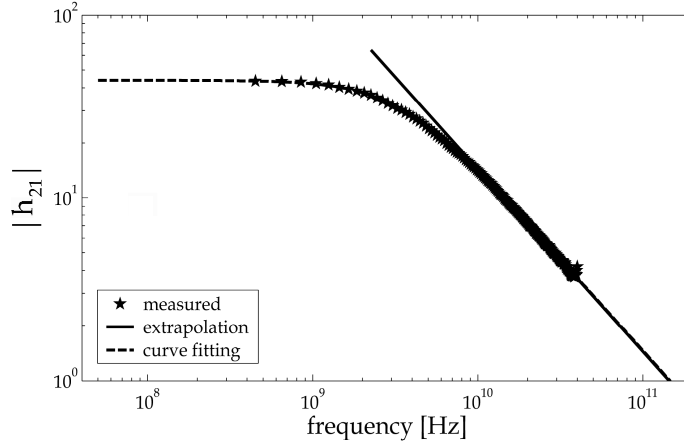


Figure 2.44: Extraction of f_T

The result obtained from measured data can be compared with the theoretical value of f_T, which is [9]

$$f_T = \frac{1}{2\pi\left[\tau_D + r_e C_{BE} + (r_e + r_{ee} + r_c)\, C_{BC}\right]} \tag{2.51}$$

Measured and calculated values must agree to ensure proper extraction of the small signal elements.

2.4.7 Extraction of Mason's Power Gain Cutoff Frequency

A transistor, when connected in a certain configuration that includes a matching network that meets some specific requirements [10], has a unilateral power gain given by [10]

$$U = \frac{|Z_{12} - Z_{21}|}{4\left(\mathrm{Re}\,(Z_{11}) \cdot \mathrm{Re}\,(Z_{22}) - \mathrm{Re}\,(Z_{12}) \cdot \mathrm{Re}\,(Z_{21})\right)} \tag{2.52}$$

where U is known as Mason's unilateral power gain. The frequency at which U equals to unity has several practical meanings [11], such as maximum frequency of activity or oscillation. This frequency is denoted by f_{MAX}, and it can be evaluated by the following expression [5]:

$$f_{MAX} = \frac{f_T}{8\pi \cdot r_b C_{BC}} \tag{2.53}$$

Consider the transistor illustrated in figure 2.39. The elements describing the pads amount to a network that meets all Mason's requirements for a matching network. This means that U is peeling independent, i.e. U can be extracted from the two-port parameters either before or after peeling the pad parasitics.

Also here, f_{MAX} is usually higher than measurement limits, and it's evaluated by both extrapolating U in a –20 dB/dec slope and fitting a dual-pole function to the measured data, as shown in figure 2.45. Note that U is in power units, therefore the decibels should be calculated appropriately.

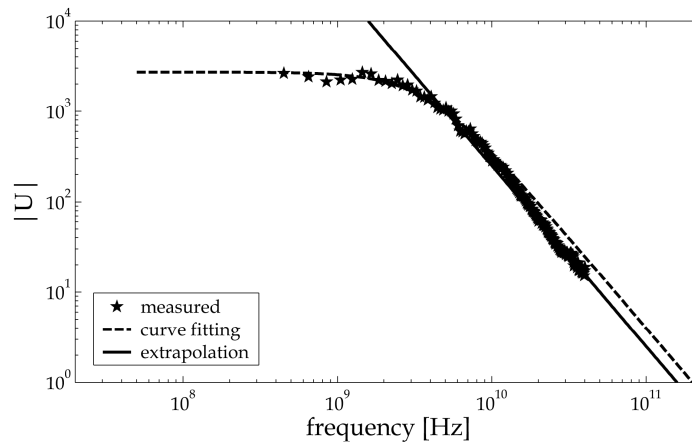


Figure 2.45: Extraction of f_{MAX}

2.4.8 Actual Extracted Parameters

Two different wafers, from which transistors and circuits have been fabricated, were used for this work. The wafers differ in collector width, where 6000 and 8000 Å were grown. All other layers are identical in both wafers. Table 2.3 summarizes the parameters extracted from measurements of transistors from these wafers.

Parameter	Parameter Description	6000 Å Collector Wafer	8000 Å Collector Wafer
β_0	Low frequency current gain	31	27
n_e	BE junction ideality factor	1.22	1.54
n_c	BC junction ideality factor	1.71	1.64
I_{ES}	BE junction saturation current	2.09 [fA]	15.5 [fA]
I_{CS}	BC junction saturation current	1490 [pA]	703 [pA]
τ_D	Device forward transit time	1.3 [psec]	1.6 [psec]
r_{ee}	Emitter external resistance	3.0 [Ω]	3.0 [Ω]
r_b	Base resistance	3.0 [Ω]	3.0 [Ω]
r_c	Collector external resistance	2.5 [Ω]	2.5 [Ω]
C_{BE}	BE junction capacitance	29 [fF]	29 [fF]
C_{BC}	BC junction capacitance	20 [fF]	17 [fF]
V_{CEO}	Avalanche breakdown voltage	4.8 [V]	6.2 [V]

Table 2.3: HBT extracted parameters

2.5 Construction of HBT VBIC Model

When both DC and RF measurements are done construction of a simulation model is obtainable. The VBIC model, widely used in many circuit CAD softwares, is the model utilized in this work in a simplified guise, which is adequate for small signal simulations with some limitations. The guideline of the model construction is to keep it as simple and comprehensible as possible, leaving no empirical parameters for numerical parameter fitting.

In order to derive VBIC parameters, the BJT model development is outlined from the pioneering Ebers–Moll model, through Gummel–Poon model, up to VBIC model.

2.5.1 Basic Ebers–Moll Model

Two versions of Ebers–Moll (*EM*) model are available: injection and transport version. The former is intuitive and provides the simplest illustration of transistor operation, whereas the latter facilitates the inclusion of large signal phenomena in the model.

Injection Version

This model is based on injection of charge carriers from the emitter to the base, and the forward current transmission coefficient, α_F. Due to the possibility to connect the transistor in an opposite direction — in which the collector acts as an emitter, and vice versa — the model is symmetric yet employing different parameters for reverse mode (such as α_R and r_{cr}). The following figure depicts this model.

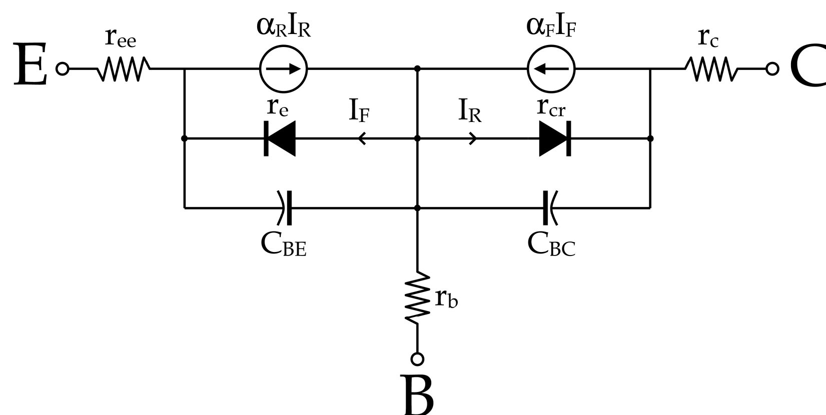


Figure 2.46: Ebers–Moll model (injection version)

The resistors r_e and r_{cr} model the differential resistances of BE and BC junctions, respectively. In large signal terms they behave like PN–junctions according to their I–V curves. Additional equations are required to complete the model, known as the *Ebers–Moll equations* [12]:

$$I_F = I_{ES} \cdot e^{\left(\frac{qV_{BE}}{KT} - 1\right)} \tag{2.54}$$

$$I_R = I_{CS} \cdot e^{\left(\frac{qV_{BC}}{KT} - 1\right)} \tag{2.55}$$

$$I_E = I_F - \alpha_R I_R \tag{2.56}$$

$$I_C = -I_R + \alpha_F I_F \tag{2.57}$$

$$I_B = (1 - \alpha_F) I_F + (1 - \alpha_R) I_R \tag{2.58}$$

where I_{ES} and I_{CS} are BE and BC junction saturation currents, respectively.

This model is easy to realize and includes basic large signal behavior. However, it lacks modeling of Early effect, Kirk effect, avalanche breakdown, and transistor heating.

Transport Version

Taking advantage of the reciprocity property of the model, the following equation can be written [12]:

$$\alpha_F I_{ES} = \alpha_R I_{CS} \triangleq I_S \tag{2.59}$$

where I_S is the entire transistor saturation current. Using some network theorems we obtain the model illustrated in figure 2.47. The Ebers–Moll equations ((2.54) — (2.58))

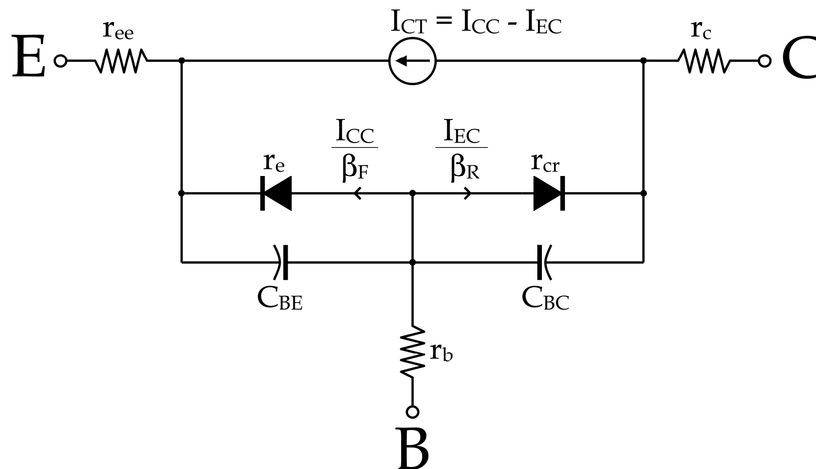


Figure 2.47: Ebers–Moll model (transport version)

should now be rewritten as follows:

$$I_{CC} = I_S \cdot e^{\left(\frac{qV_{BE}}{KT}-1\right)} \quad (2.60)$$

$$I_{EC} = I_S \cdot e^{\left(\frac{qV_{BC}}{KT}-1\right)} \quad (2.61)$$

$$I_{CT} = I_{CC} - I_{EC} \quad (2.62)$$

$$I_E = -I_{CT} - \frac{I_{CC}}{\beta_F} \quad (2.63)$$

$$I_C = I_{CT} - \frac{I_{EC}}{\beta_R} \quad (2.64)$$

$$I_B = \frac{I_{CC}}{\beta_F} + \frac{I_{EC}}{\beta_R} \quad (2.65)$$

In this form of the model Early effect can be added to the model by modifying equations (2.62) and (2.65) and using an additional parameter, V_A, that denotes Early voltage (V_A is the voltage at which all extrapolated common emitter curves, in forward active mode, intercept the $I_C = 0$ axis) [12]. Yet Kirk effect, avalanche breakdown, and heating are not modeled.

2.5.2 Basic Gummel–Poon Model

The original Gummel–Poon (*GP*) model comprises three additional phenomena over EM model:

- Recombination in the space charge region or other effects that affect current gain
- Modeling of Early effect
- Strong injection in BE junction

The last two phenomena are irrelevant for HBT. In order to incorporate these changes each diode is split into two diodes — one models the normal operation of the transistor (r_e in figure 2.48), and the other models other currents that don't contribute to current gain, such as recombination in the depletion regions (r_{ES} in figure 2.48). The equations of I_S and transistor currents are modified as well to reflect the behavior under all regimes. Contrary to EM model, GP equations are derived from electric charge starting point, instead of EM's current approach.

In an improved version of GP model BC junction is modeled by two parallel junctions: external and internal junction. This division is intended for modeling the distributed nature of the junction. A current source is connected in parallel to each portion of BC junction to model avalanche breakdown effect (I_{avch}). All of the changes amount to the topology shown in figure 2.48 (see [12], [13]).

2.5.3 VBIC Parameters for Constructing Simple Model

The vertical bipolar inter-company (*VBIC*) model has been developed in the mid 90's in order to overcome several limitations and drawbacks of GP model. The idea was to

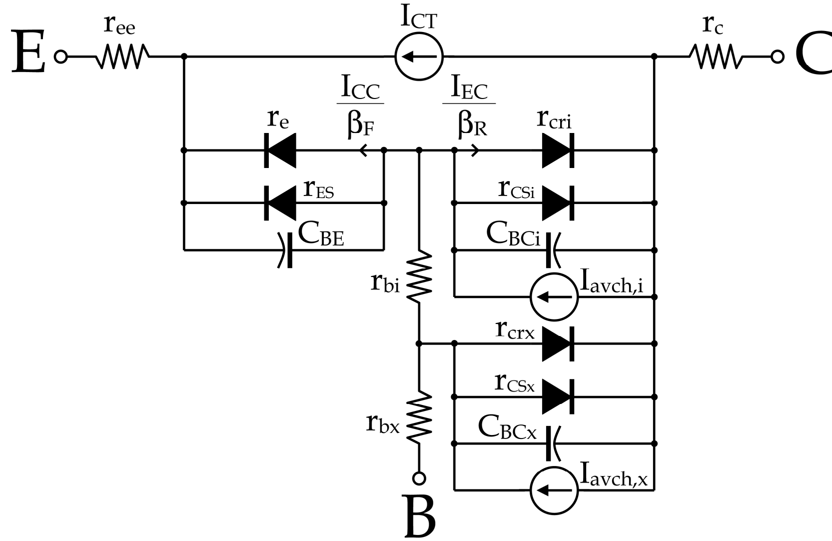


Figure 2.48: Gummel–Poon model

stay as close to GP model as possible and therefore make only the necessary changes in the model [14]. VBIC model provides modeling of many phenomena, of which some are exclusive to silicon BJT technology and irrelevant to the InP technology of this work. From an InP-based HBT viewpoint, the most important phenomena modeled by VBIC are Kirk effect, improved avalanche breakdown modeling, and temperature dependence (including self heating). Besides GP model, VBIC model includes an additional substrate parasitic transistor and another heating modeling sub-circuit. Although comprehensive modifications were applied to all equations, electric charge approach has still been retained.

The construction of VBIC model in this work was targeted towards a simple one-to-one correspondence to the measured small signal model elements. For this reason BC junction hasn't been split to internal and external junctions, low current and avalanche breakdown effects have been ignored, as well as Kirk effect and temperature variations. These approximations have proved to be precise enough — well below process tolerances — for circuit simulation. Moreover, decent circuit design must not depend on sensitive parameters. As a consequence, only basic parameters were set to meaningful values whereas all the others were reset to ineffective defaults. Being of prime importance, the values substituted in the effective VBIC parameters are listed in table 2.4 (according to [15]). Note that measured parameter nomenclature coincides with that of table 2.3.

VBIC Parameter	Value calculated from measured parameters	Parameter Description [16]
RCX	r_c	extrinsic collector resistance
RCI	0	intrinsic collector resistance
RBX	r_b	extrinsic base resistance
RBI	0	intrinsic base resistance
RE	r_{ee}	emitter resistance
IS	$I_{ES} \cdot \alpha_0$	transport saturation current
NF	n_e	forward emission coefficient
NR	n_c	reverse emission coefficient
CJE	C_{BE}	base–emitter zero-bias junction capacitance
CJC	C_{BC}	base–collector zero-bias capacitance
$IBEI$	$\frac{IS}{\beta_0}$	ideal base–emitter saturation current
NEI	n_e	ideal base–emitter emission coefficient
$IBCN$	I_{CS}	non-ideal base–collector saturation current
NCN	n_c	non-ideal base–collector emission coefficient
TF	τ_D	forward transit time

Table 2.4: Conversion of measured parameters to VBIC parameters

2.5.4 VBIC Model Validation

When VBIC model is constructed it's essential to verify it and ensure that parameters have been extracted correctly. The validation sequence is described hereinafter, followed by descriptions of the simulation environment and results.

Validation Program

Since the model constructed in this work is minimal and intended for small signal simulations, the VBIC model's S–parameters are compared to the measured S–parameters in various biasing points. Biasing point range extends over all important points — from collector currents as small as 1 mA to large currents that launch Kirk effect; and from the smallest V_{CE} that still retains forward active mode to large voltages just before avalanche breakdown and transistor burnout.

The model is simulated in Agilent's Advanced Design System (*ADS*) 2003A by using its S–parameter simulation tool. Biasing point is set manually for each test, and then S–parameter simulation is invoked.

Simulation Circuit

The simulation circuit includes a transistor with the extracted VBIC model, and a circuit that represents the small signal model of the biasing point under test. Both are surrounded by the pad parasitic elements extracted from the test fixture measurements. Each is simulated by S–parameter simulation, and results are compared to the measured S–parameters.

Figure 2.49 shows the schematic of the simulation circuit in ADS. The upper circuit simulates the small signal model, and the lower circuit simulates a transistor with the constructed VBIC model. Another module, evident in the left side of figure 2.49, loads the measured S–parameter file for comparison purposes.

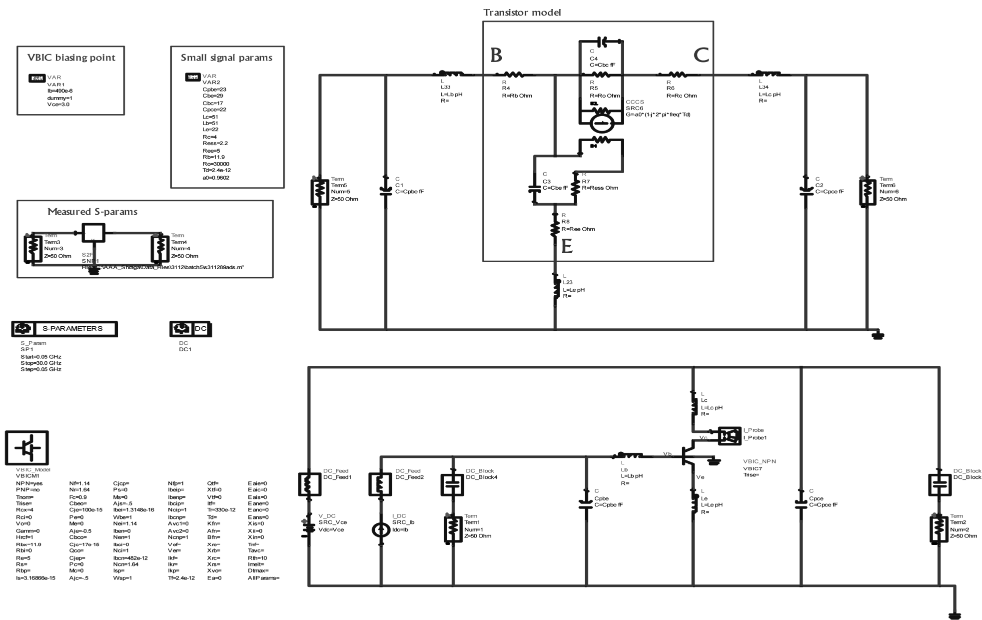


Figure 2.49: ADS simulation circuit for VBIC model verification

Validation Results

The results obtained from the comparison simulations are presented in graphs, as can be seen in figure 2.50. Measured data is displayed in gray stars, and simulated data of the VBIC model in solid black line. Both real and imaginary parts of each parameter are displayed. A comparison between the S–parameters of the VBIC model and the small signal model , not presented here, shows negligible differences.

The magnitude and phase of each parameter are compared by calculating an error function — the ratio between measured and simulated data in the case of magnitude, and difference in the case of phase. Model is adjusted to keep the magnitude error below 1 dB and phase error below 8 degrees in all biasing points.

During model validation we have noticed that for an unknown reason VBIC model does not take into account NEI, the ideal base–emitter emission coefficient, in the calculation of r_e. This is in contrast to GP model.

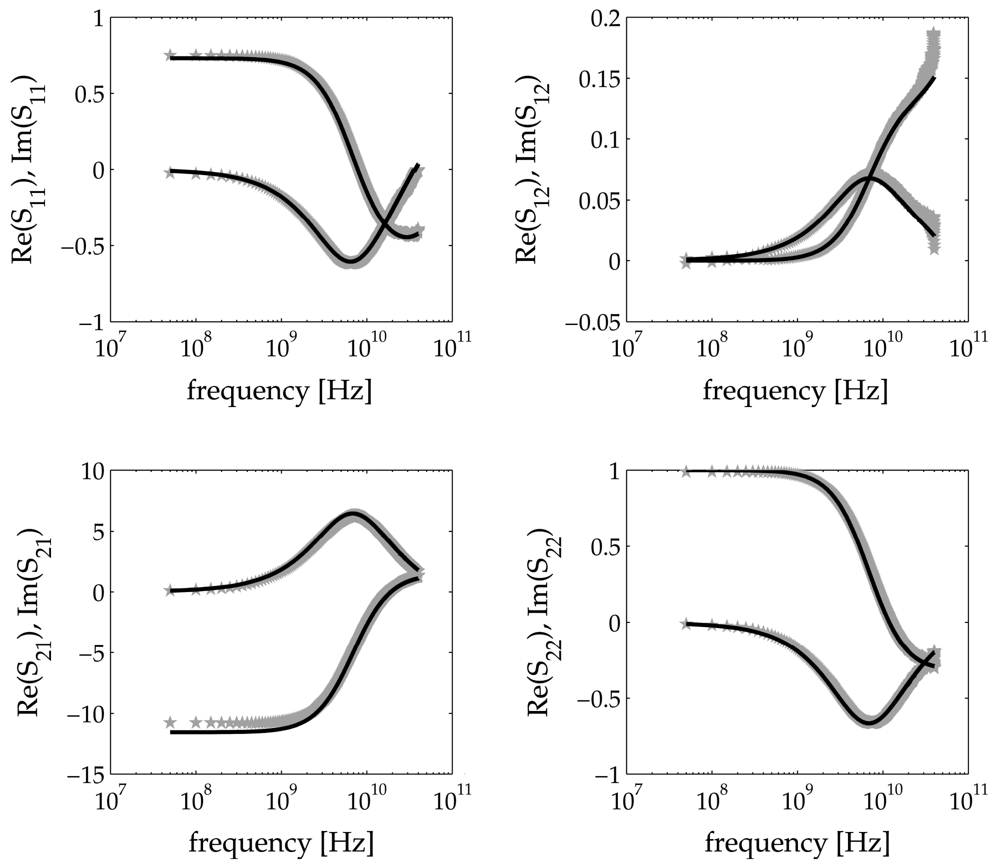


Figure 2.50: Comparison between simulated and measured S–parameters

2.6 Noise in HBT

The aim of this section is to provide a short description of noise types and their impact on transistor noise performance. The equations presented here are for reference in the next chapters, but they are mentioned here due to their direct connection with the elements of the HBT model. A treatment of circuit noise performance is available in chapter 4.

2.6.1 Types of noise

Electrical noise is a consequence of the fact that electrical charge, as well as electrical current, consists of discrete charge carriers rather than continues charge. The differences in the behavior of each particular charge carrier result in fluctuations in the electrical current which are translated to random noise current or voltage.

Shot Noise

Shot noise is a result of DC current flowing through a PN–junction. Every charge carrier has a probability to pass across the junction determined by its energy and direction of velocity. The average square-current resulted by this noise mechanism is [1]

$$\overline{i_{sn}^2} = 2qI_D\Delta f \tag{2.66}$$

where q is the electron's electric charge, I_D is the DC current of the diode, and Δf is the bandwidth of interest. It can also be expressed as a noise-current spectral density,

$$\frac{\overline{i_{sn}^2}}{\Delta f} = 2qI_D \tag{2.67}$$

which is independent of frequency. The total diode current has therefore a normal distribution function with a Gaussian around I_D and standard deviation of

$$\sigma = \sqrt{\overline{i_{sn}^2}} = \sqrt{2qI_D\Delta f} \tag{2.68}$$

The noise is modeled as a current source connected in parallel to the diode. Note that shot noise degrades significantly at frequencies of the order of $\frac{1}{\tau_D}$, where τ_D is the junction transit time.

Flicker Noise (1/f Noise)

The source of flicker noise is not well known. Some of the possible explanations are contaminations and crystal defects that create traps for charge carriers. Carriers are randomly trapped in and released from these traps, resulting in a noise current that is given by [1]

$$\overline{i_{fn}^2} = K_f I_D^a \frac{\Delta f}{f} \tag{2.69}$$

where K_f is an empirical constant unique for a particular device, I_D is the DC current flowing in the device, a is a an empirical constant ($0.5 \leq a \leq 2$), and Δf is a small bandwidth around f. Flicker noise's distribution is usually non-Gaussian.

Due to its inverse proportion to f flicker noise is so-called 1/f noise.

Johnson Noise (Thermal Noise)

Johnson noise, also referred to as thermal noise, is generated by the thermal motions of charge carriers in every resistor. This noise mechanism can be represented by a voltage source in series to the resistor, in which [1]

$$\overline{v_{tn}^2} = 4KTR\Delta f \tag{2.70}$$

or by a shunt current source,

$$\overline{i_{tn}^2} = 4KT\frac{1}{R}\Delta f \tag{2.71}$$

where R is the resistor's resistance. Like in shot noise, thermal noise is frequency independent, but is linear with temperature. In addition, current distribution is Gaussian with standard deviation of

$$\sigma = \sqrt{\overline{i_{tn}^2}} = \sqrt{4KT\frac{1}{R}\Delta f} \tag{2.72}$$

similarly to shot noise.

Burst Noise

The origin of burst noise is not fully recognized. The noise itself appears as a square wave current in several frequencies. It is strong in low frequencies up to a knee frequency, f_b, as can be seen in the following expression [1]:

$$\overline{i_{bn}^2} = K_b I_D^b \cdot \frac{\Delta f}{1 + \left(\frac{f}{f_b}\right)^2} \tag{2.73}$$

where K_b, b, and f_b are empirical elements, and I_D is the DC current flowing in the device. The constant b satisfies $0.5 \leq b \leq 2$.

Avalanche Noise

When avalanche occurs in a PN–junction the random creation of electron–hole pairs generates noise of a similar quality to shot noise. Since high energy is involved in the process of electron–hole generation, the noise is significantly stronger than shot noise. Furthermore, the noise is amplified by the avalanche mechanism, making it even stronger.

It is difficult to predict avalanche noise, and it doesn't behave in a Gaussian manner. The only recommendation is to avoid it.

Summary

All noise types mentioned above are summarized in table 2.5.

2.6.2 Noise Mechanisms in HBT

Being constructed from two PN–junctions, and modeled as a bundle of passive and active elements, the HBT has its own noise behavior. Hence, PN–junction noise is discussed firstly, followed by a the noise model the represents the noise mechanisms in a HBT.

Noise in PN–Junction

The main noise mechanisms that exist in a PN–junction are shot and flicker noise, which can be represented by a single shunt current source. In actual diode the serial resistances of the contacts, r_c, add thermal noise source to the diode (this noise is represented by a serial voltage source). These two sources form the equivalent circuit shown in figure 2.51, where the values of the noise sources are given by [1][17]

$$\overline{v^2} = 4KTr_c\Delta f \quad (2.74)$$

$$\overline{i^2} = 2qI_D\Delta f + K_f I_D^a \frac{\Delta f}{f} \quad (2.75)$$

Noise Type	Origin	Expression	DC Current Dependence	Frequency Dependence	Amplitude Distribution
Shot	PN–junction	$\overline{i_{sn}^2} = 2qI_D\Delta f$	I_D	none	Gaussian
Flicker (1/f)	semiconductor Imperfections	$\overline{i_{fn}^2} = K_f I_D^a \frac{\Delta f}{f}$	I_D^a	$\frac{1}{f}$	non-Gaussian
Johnson (Thermal)	resistor	$\overline{i_{tn}^2} = 4KT\frac{1}{R}\Delta f$	none	none	Gaussian
Burst	unknown	$\overline{i_{bn}^2} = K_b I_D^b \cdot \frac{\Delta f}{1+\left(\frac{f}{f_b}\right)^2}$	I_D^b	$\frac{1}{1+\left(\frac{f}{f_b}\right)^2}$	non-Gaussian
Avalanche	PN–junction avalanche	(unpredictable)	strong dependence	(random spikes)	non-Gaussian

Table 2.5: Summary of noise types

Note that r_d doesn't contribute any noise since it's a synthetic resistor for modeling purposes only.

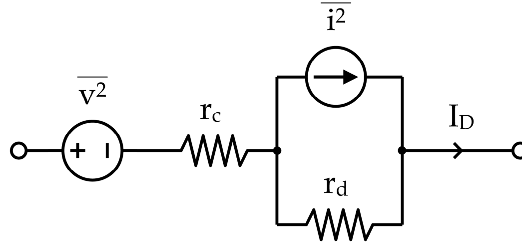


Figure 2.51: Small signal equivalent circuit of a PN–junction with noise sources. $r_d = \frac{KT}{qI_D}$ is the differential resistance of the diode

Noise in HBT

The intrinsic noise sources of a HBT are as follows [1][17]:

1. Shot noise due to the flow of current through BC junction
2. Shot noise due to the flow of current through BE junction
3. Flicker noise due to recombination in the transistor — mostly in the emitter [18]
4. Burst noise due to traps of a certain type [1]
5. Thermal noise due to contact resistors — mainly r_b

Shot noise in BC junction (No.1) can be represented by a current source placed between the collector and emitter terminals. Its value is given by [1]

$$\overline{i_c^2} = 2qI_C\Delta f \tag{2.76}$$

Base noise current consists of BE junction shot noise (No.2), flicker noise (No.3), and burst noise (No.4). They all can be lumped in a single current source of [1]

$$\overline{i_b^2} = 2qI_B\Delta f + K_f I_B^a \frac{\Delta f}{f} + K_b I_B^b \cdot \frac{\Delta f}{1 + \left(\frac{f}{f_b}\right)^2} \tag{2.77}$$

connected between the base and emitter terminals. Finally, the noise originated by the base resistance r_b (No.5) is modeled by a voltage source that equals to the thermal noise generated by this resistance [1]:

$$\overline{v_b^2} = 4KTr_b\Delta f \tag{2.78}$$

The resistors r_e and r_{out} do not generate any noise because they are synthetic resistors intended for modeling.

The noise mechanisms described here form the model depicted in figure 2.52.

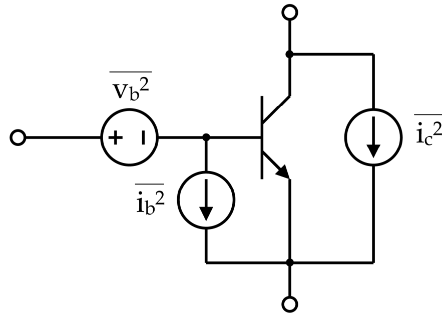


Figure 2.52: Noise sources in HBT

The model of noise presented above is also known as *Fukui model*. Its main advantages over other models are the absence of correlation between the different noise sources and the ability to adapt the model to more complex versions of transistor model [19][20].

References for Chapter 2

[1] Paul R. Gray and Robert G. Meyer, *Analysis and Design of Analog Integrated Circuits*, John Wiley & Sons, third edition, 1993.

[2] A. S. Grove, *Physics and Technology of Semiconductor Devices*, John Wiley & Sons, 1967.

[3] Steven E. Laux and Wai Lee, "Collector signal delay in the presence of velocity overshoot", *IEEE Electron Device Letters*, vol. 11, no. 4, pp. 174–176, April 1990.

[4] Herbert Kroemer, "Heterostructure bipolar transistors and integrated circuits", *Proceedings of the IEEE*, vol. 70, no. 1, pp. 13–25, January 1982.

[5] Mathias Kahn, *Transistor Bipolaire À Hététrojonction GaInAs/InP Pour Circuits Ultra-Rapides : Structure, Fabrication et Caractérisation*, PhD thesis, Universite Paris XI Orsay, June 2004.

[6] Dieter K. Schroder, *Semiconductor Material and Device Characterization*, John Wiley & Sons, 1990.

[7] Adir Bar Lev, *Semiconductors and Electronic Devices*, Prentice–Hall, second edition, 1984.

[8] Benny Sheinman, Edward Wasige, Matthias Rudolph, Ralf Doerner, Victor Sidorov, Shimon Cohen, and Dan Ritter, "A peeling algorithm for extraction of the HBT small-signal equivalent circuit", *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 12, pp. 2804–2810, December 2002.

[9] Solon Jose Spiegel, Dan Ritter, R. A. Hamm, A. Feygenson, and P. R. Smith, "Extraction of the InP/GaInAs heterojunction bipolar transistor small-signal equivalent circuit", *IEEE Transactions on Electron Devices*, vol. 42, no. 6, pp. 1059–1064, June 1995.

[10] S. J. Mason, "Power gain in feedback amplifier", *Trans. IRE Professional Group on Circuit Theory*, vol. CT-1, no. 2, pp. 20–25, June 1954.

[11] Madhu S. Gupta, "Power gain in feedback amplifiers, a classic revisited", *IEEE Transactions on Microwave Theory and Techniques*, vol. 40, no. 5, pp. 864–879, May 1992.

[12] Paolo Antognetti and Giuseppe Massobrio, *Semiconductor Device Modeling with SPICE*, McGraw–Hill, 1988.

[13] Fujiang Lin, Tianshu Zhou, Bo Chen, Ban Leong Ooi, and Pang Shyan Kooi, "Extraction of VBIC model for SiGe HBTs made easy by going through gummel–poon model", *Proceedings of SPIE*, vol. 4228, pp. 249–258, October 2000.

[14] Colin C. McAndrew, Jerold A. Seitchik, Derek F. Bowers, Mark Dunn, Mark Foisy, Ian Getreu, Marc McSwain, Shahriar Moinian, James Parker, David J. Roulston, Michael Shröter, Paul Van Wijnen, and Lawrence F. Wagner, "VBIC95, the vertical bipolar inter-company model", *IEEE Journal of Solid-State Circuits*, vol. 31, no. 10, pp. 1476–1483, October 1996.

[15] Günther Palfinger, *Modelling the Heterojunction Bipolar Transistor with VBIC*, PhD thesis, Technical University Graz, January 2000.

[16] Agilent Technologies, *Nonlinear Devices*, Advanced Design System 2003A Documentation, May 2003.

[17] David A. Johns and Ken Martin, *Analog Integrated Circuit Design*, John Wiley & Sons, 1997.

[18] Emanuel Cohen, "Design of wide-band distributed amplifier based on InP/GaInAs hetrojunction bipolar transistors", Master's thesis, Technion – IIT, 2002.

[19] M. Rudolph, R. Doerner, L. Klapproth, and P. Heymann, "An HBT noise model valid up to transit frequency", *IEEE Electron Device Letters*, vol. 20, no. 1, pp. 24–26, January 1999.

[20] Laurent Escotte, Jean-Phillippe Roux, Robert Plana, Jacques Graffeuil, and Andreas Gruhle, "Noise modeling of microwave heterojunction bipolar transistors", *IEEE Transactions on Electron Devices*, vol. 42, no. 5, pp. 883–889, May 1995.

Chapter 3

Photodiode

Optical communication receivers comprise an element that converts an optical signal to electrical current or voltage. The device used for this intention in this work is the photodiode (*PD*) — which is the subject of this chapter. The chapter begins with an overview on the physics of PD. Then comes a full description of the PD fabrication process and the measurements and characterization concerned with PD, including the extraction of a small signal model. Noise mechanisms that reside in a PD are the last topic discussed in the chapter.

3.1 Photodiode Physical Characteristics

The physics of PD has great importance to the design of the diode itself and the entire circuit neighboring it. This section overviews the basic principles and limitations of the PD. Both DC and AC characteristics are discussed to cover all aspects of operation. In the forthcoming chapters these principles will be referred to in the context of optoelectronic integration.

3.1.1 DC Characteristics

DC characteristics include the physics of detection and the conversion of an optical signal to electrical one, and definitions of parameters that characterize a detector. These subjects are detailed hereinafter.

Physics of Normal Operation

Photodiode is a reverse biased PN–junction. Photons generate electron–hole pairs in its depletion region. The generated charge carriers drift due to the electric field in the depletion region to the neutral regions of the diode. Since holes and electrons are of opposite charge they drift towards opposite sides of the depletion region. In terms of electrical currents, their currents sum up to a total photo current that can be measured at the diode's nodes [1][2]. As light is absorbed also in the neutral regions, generation takes place also there. Electron–hole pairs generated in the neutral regions, far away from the junction, recombine before arriving the junction and don't contribute to the photo current. However, charge carriers generated within a diffusion length from the junction, diffuse to the junction and add up to the total photo current [1]. The entire process of absorbing an incident photon, the following generation of electron–hole pair, and their drift to the neutral regions — is illustrated in figure 3.1.

As will be explained below, the wider the depletion region the more efficient the photodiode is. For this reason photodiodes are usually implemented by PIN–diodes with wide depletion region and narrow neutral regions [1]. This improves their responsivity and frequency performance, to be elucidated hereinafter.

Quantum Efficiency and Responsivity

Incident optical power, P_{opt}, consists of photon flux. The incident photon flux density (per unit area) is given by [2]

$$\phi_{opt} = \frac{P_{opt}}{A \cdot \frac{hc}{\lambda}} \tag{3.1}$$

where A is the detection area of the photodiode, h is Planck's constant, c is the speed of light, and λ is the wavelength of the light. Let us assume that the light is incident from

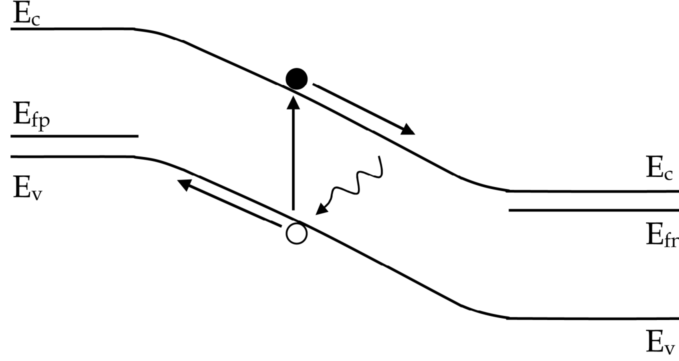


Figure 3.1: Process of photon detection by a photodiode. Photon is marked by wavy arrow, electron by black circle, and hole by white circle.

the top side of the photodiode. The photon flux density that actually penetrates the PD and absorbed in it is

$$\phi_0 = \phi_{opt} \cdot (1 - R_{top}) \tag{3.2}$$

where R_{top} is the reflectivity of the top surface of the PD. Light is absorbed in an exponential profile decaying from the top surface downwards with a coefficient denoted by α. Hence, the electron–hole pair generation rate in the PD is [1][2]

$$G(x) = \phi_0 \cdot \alpha e^{-\alpha x} \tag{3.3}$$

where x is the offset from the top surface into the PD bulk.

It can be concluded from (3.3) that for high PD efficiency, i.e. for exploiting as much photons for detection as possible, the PD should be thick enough to include most of the exponential absorption profile. For this reason PIN–diode structure is commonly used, allowing one to engineer the thickness of the depletion region. A cross-section that illustrates the structure of a top-illuminated PD is shown in figure 3.2.

The external quantum efficiency of a photodiode measures the ratio between the number of incident photons and the number of photons that actually generate current that is measurable at the PD nodes. It is given by [2]

$$\eta = \frac{hc}{\lambda} \cdot \frac{I_{PD}}{qP_{opt}} \tag{3.4}$$

where I_{PD} is the lighting current measured at the PD nodes. Usually the *responsivity* of a PD is of interest. The responsivity is defined as

$$\Re = \frac{I_{PD}}{P_{opt}} \tag{3.5}$$

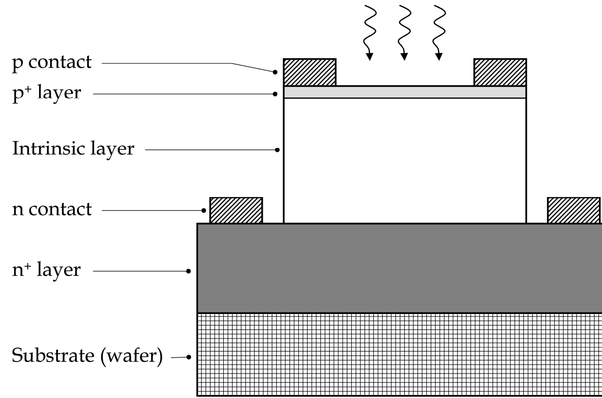


Figure 3.2: Cross-section of a top-illuminated photodiode

Combining (3.4) and (3.5) gives a useful equation:

$$\Re = \eta \cdot q \frac{\lambda}{hc} \tag{3.6}$$

Equation (3.6) enables one to derive the quantum efficiency, η, by means of simple measurement of $\Re$.

3.1.2 AC Characteristics

Photodiodes are prevalently used for the detection of rapidly changing signals such as digital communications. Such signals pose a demand on the PD to be able to cope with high frequency signals, i.e. to detect them with no attenuation. The PD junction capacitance, together with the load resistance, set one limitation to the PD frequency performance. Another restriction is originated by the transit time of charge carriers in the junction. These two phenomena, together with the generation mechanism of the PD, construct a small signal model of a PD.

Diode Capacitance and Resistances

As mentioned above, a PD is always held under reverse bias conditions to keep the detection current the largest current that flows in the PD. As a consequence, the capacitance of a PD is its junction capacitance. In the PD depicted in figure 3.2, the area that contributes to the capacitance is determined by the smaller of the p^+ and the intrinsic layers. Evidently, the junction width, which is equal to the intrinsic layer thickness, also determines the capacitance, obtaining

$$C_j = \epsilon_0 \epsilon_r \cdot \frac{A_j}{W_{PD}} \tag{3.7}$$

where A_j is the junction area, and W_{PD} is its width. In addition, leakage in the junction may change the current flowing in the PD. This can be modeled by means of a resistor connected in parallel to the photo generation element, denoted by r_{out}. Finally, the resistance of the diode contacts are modeled by a serial resistor, denoted by $r_{contact}$.

Transit Time

The time it takes for generated charge carriers to arrive their target contact is an intrinsic limit for the bandwidth of the PD. Once generated in the depletion region, electron and holes drift by the electric field to the contacts. In most cases, the electrons' velocity is significantly greater than the holes' velocity. Therefore, it is preferable to illuminate the PD from the p-layer side, so that most of the generated holes are close to their target contact, and have shorter distance to travel [3]. In addition, the following should be noted: The diffusion mechanism is significantly slower than drift in the depletion region. Consequently, only generation in the depletion region is desired, which is achieved by the wide PIN structure. Large bandgap n contact avoid generation in the neutral regions.

To calculate the frequency response due to transit time one assumes that the optical signal applied to the PD is of the form $\phi_0\left(1+e^{j\omega t}\right)$. Solving the continuity and the current density equations with the above assumptions yields a current density of [2]

$$\begin{aligned} J(\omega) = {} & q\phi_0\alpha W_{PD}\left[\frac{e^{-\alpha W_{PD}}-1}{\alpha W_{PD}\left(\alpha W_{PD}-j\omega t_{tr}^{h}\right)}+\frac{e^{-\alpha W_{PD}}\left(e^{j\omega t_{tr}^{h}}-1\right)}{j\omega t_{tr}^{h}\left(\alpha W_{PD}-j\omega t_{tr}^{h}\right)}\right] \\ & - q\phi_0\alpha W_{PD}\left[\frac{e^{-\alpha W_{PD}}-1}{\alpha W_{PD}\left(\alpha W_{PD}+j\omega t_{tr}^{e}\right)}+\frac{e^{j\omega t_{tr}^{e}}-1}{j\omega t_{tr}^{e}\left(\alpha W_{PD}+j\omega t_{tr}^{e}\right)}\right] \end{aligned} \tag{3.8}$$

where t_{tr}^{h} and t_{tr}^{e} are the hole and electron transit time, respectively. The transit times are calculated by simply dividing W_{PD} by the saturation velocities. The pole caused by transit time can be found by plotting $|J(\omega)|$ and extracting the frequency at which the current density degrades by 3 dB. This frequency is usually referred to as *–3dB bandwidth*.

The –3dB bandwidth due to transit time can be approximated by a more simple expression [3]:

$$f_{-3dB} \approx \frac{K}{2\pi}\cdot\frac{v_{sat}^{h}}{W_{PD}} \tag{3.9}$$

where v_{sat}^{h} is the hole saturation velocity and K is an empirical constant. Since the PDs of this work are made of GaInAs, the calculations of the –3dB bandwidth due to transit time are based on the parameters detailed in table 3.1. It should be noted that [3] assumes that the electrons' velocity is at least 2 times greater than the holes' velocity, and thus concludes that $K = 5.2$. In GaInAs the velocities are relatively close to each other, and a curve fitting is required to determine K.

The results of these calculations show that a 8000 Å junction has transit time of 4 psec, and 6000 Å yield 2.9 psec. The –3dB bandwidth versus junction width is illustrated in figure 3.3, and as can be observed, (3.8) and (3.9) yield very close results.

Parameter	Description	Value	Reference
v_{sat}	Electron saturation velocity	$7 \cdot 10^6 \left[\frac{\text{cm}}{\text{sec}}\right]$	[4], [5]
v^h_{sat}	Hole saturation velocity	$4.9 \cdot 10^6 \left[\frac{\text{cm}}{\text{sec}}\right]$	[6], [5]
α	Absorption coefficient	$0.68 \left[\mu\text{m}^{-1}\right]$	[7]
K	Constant of (3.9)	4.14	(curve fitting)

Table 3.1: Physical properties of GaInAs and parameters used for calculating the –3dB bandwidth of a photodiode

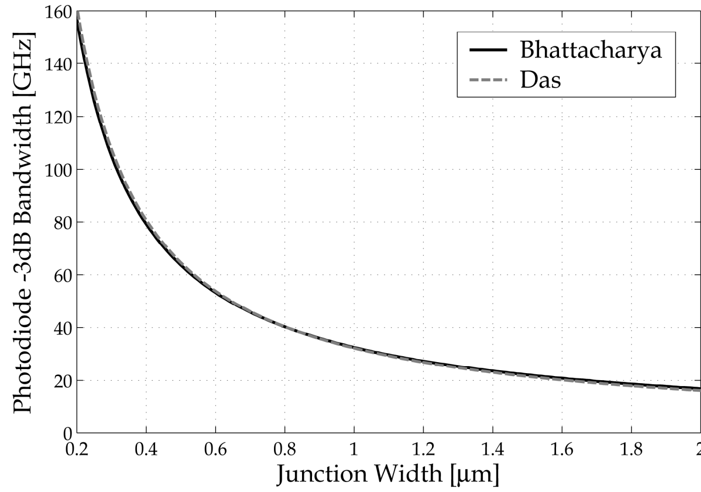


Figure 3.3: GaInAs Photodiode –3dB frequency due to transit time versus junction width. Calculated both according to Bhattacharya [2] and Das [3], i.e. (3.8) and (3.9), respectively

Small Signal Model

The phenomena and elements mentioned above sum up to the small signal model shown in figure 3.4. The generation current is represented by a current source, i_{PD}, and the charge carrier transit time is modeled by the frequency dependence of i_{PD}, namely

$$i_{PD}(\omega) = i_{PD}(P_{opt}, \Re) \cdot \frac{1}{1 + j\omega\tau_{PD}} \tag{3.10}$$

where τ_{PD} is the effective transit time in the PD derived from the –3dB frequency, and $i_{PD}(P_{opt}, \Re)$ is the small signal current generated in the depletion region, according to the incident optical power and the diode's responsivity.

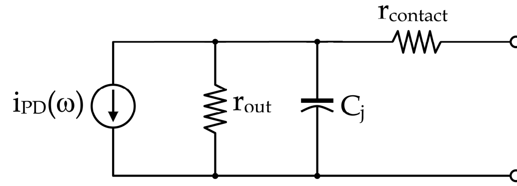


Figure 3.4: Small signal model of a photodiode

3.2 Photodiode Fabrication and Structure

Due to its bandgap of 0.75 eV the $Ga_{0.47}In_{0.53}As$ is the material of choice for photodiodes intended for 1.55 μm wavelength [8]. In this work the GaInAs layers were grown on semi-insulating InP wafers and processed later on. This section details the fabrication process and the final structure of the PDs.

3.2.1 Fabrication Process

Photodiodes can be fabricated either separately on a dedicated wafer or monolithically with electronic circuits. Both methods were used in this work. In the monolithic case the diodes may share the base, collector and subcollector layers and their corresponding contacts. The requirement for a thick photodiode slows the HBTs as their transit time is increased. Alternatively, the photodiode layers can be grown separately, prior to the HBT layers, enabling the optimization of each device individually. However, fabrication of the latter type comprises more masks, and the high topography limits both PD processing precision and the minimal distance between adjacent HBTs. When fabricated separately (not monolithically) another mask set is used to allow fast and efficient fabrication process. This procedure is described below.

Wafer Layer Structure

The layers are similar to the HBT layers described in section 2.2.1, with the following exceptions: (a) The emitter layers are omitted, (b) there is no composition grading in the p^+ layer (the counterpart of the HBT's base layer), and (c) the n^+ layer is made of InP (contrary to the HBT's GaInAs subcollector) with a thin GaInAs contact layer above it. The layers and their properties are summarized in table 3.2. As the layers are grown by the same MOMBE system used for HBT layer growth, their properties are very similar. The cap layer seals and protects the wafer till the beginning of the fabrication process in the clean rooms. The layer structure after growth is illustrated in figure 3.5.

Figure 3.5: Layer structure on wafer as grown by MOMBE system

Junction Process

In the clean rooms the cap layer is removed and PR is flattened on the wafer and shaped by means of photolithography. The mask that defines the shapes is called "diode mesa". Now the p^+ and the intrinsic layers are wet etched to define the mesa of the junction, and then the PR is removed. The results of this process are shown in figure 3.6. Note that the etching progress is stopped by the InP etch-stop layer, and then the etch-stop layer is etched by a different selective solution.

Contact Process

Next step is creating the p and n contacts. Photo resist is applied and exposed through "contact metal" mask. Metal is deposited on the wafer (the composition of the metal is similar to that of the HBT, see section 2.2.1 on page 36) and molded by liftoff process, leaving metal areas on the p^+ and n^+ layers for electrical contacts. The p contact is

Layer	Thickness [Å]	Composition	Type	Dopant	Doping Concentration $[cm^{-3}]$
Cap Layer	500	InP	intrinsic	—	—
p^+	280	$Ga_{0.47}In_{0.53}As$	P^+	C	$\sim 4 \cdot 10^{19}$
Intrinsic	8000	$Ga_{0.47}In_{0.53}As$	N^-	Si	$\sim 7 \cdot 10^{15}$
Etch Stop	200	InP	intrinsic	—	—
n Contact	200	GaInAs	N^+	Si	$\sim 4 \cdot 10^{19}$
n^+	4000	InP	N^+	Si	$\sim 4 \cdot 10^{19}$
Substrate (Wafer)	500$[\mu m]$	InP	semi–insulator	Fe	—

Table 3.2: Layer growth properties

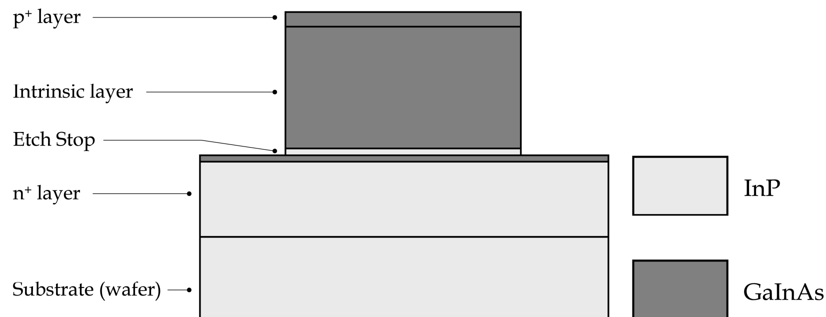


Figure 3.6: Wafer layers after diode mesa etching

designed as a ring so that it doesn't block the incident light and prevent its penetration into the diode. The layers at this level are shown in figure 3.7.

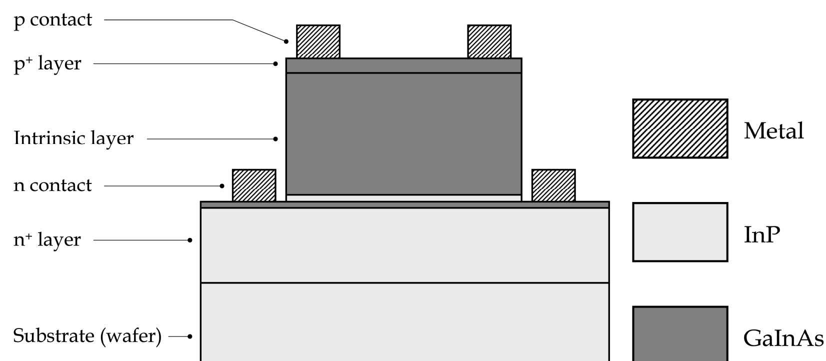


Figure 3.7: Wafer layers after contact implementation

When contacts are done the PD is electrically isolated from the rest of the wafer by means of etching the n^+ layer around the PD and so exposing the insulator wafer. This is done by exploiting the mask called "isolation". The layers of the photodiode at this level are illustrated in figure 3.8.

Passivation and Interconnect

Polyimide is spun on the wafer and dry etched by using the "Polyimide cover" mask. This mask opens a round opening in the Polyimide above the p^+ layer in order to expose the PD to incident light. Also areas on the metal contact are exposed to enable connecting the PD the the interconnect circuitry. A cross section of a Polyimide-coated PD is depicted in figure 3.9. Just like in HBT, the Polyimide acts both as passivation and

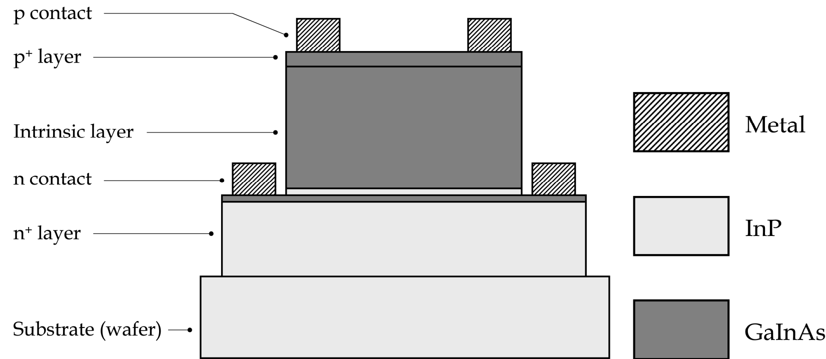


Figure 3.8: Wafer layers after "isolation" mask

protection/isolation layer.

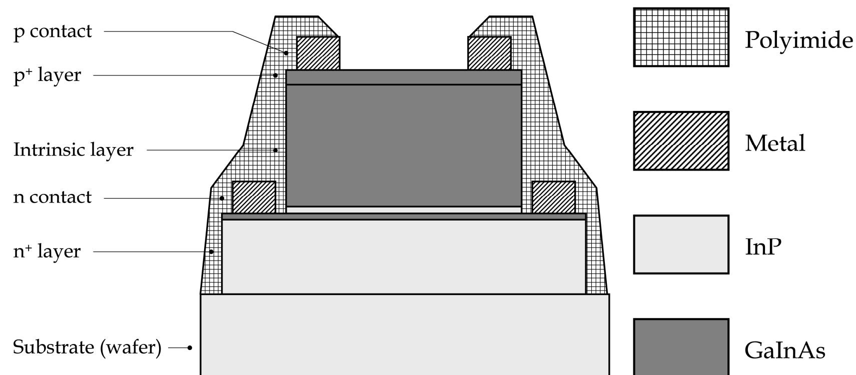


Figure 3.9: Wafer layers after "Polyimide cover" mask

Photodiodes are connected to the external world via transmission lines that transfer the electrical signals from the PD output nodes to connection pads. These pads might be connected to external circuits or probing and measurement systems. The interconnect is implemented by means of lithography of "metal 1" mask, metal deposition, and liftoff.

Anti-Reflective Coating

The last level in PD fabrication is anti-reflective coating (*ARC*). The ARC implemented in this work was a first-order single layer SiN coating. The thickness of the layer was chosen to be equal to $\frac{1}{4}\lambda \cdot \frac{1}{n_{ARC}}$, where $\lambda = 1.55$ μm is the wavelength of the incident light in air, and n_{ARC} is the diffraction coefficient of the ARC. The diffraction coefficient should be $n_{ARC} = \sqrt{n_{GaInAs} \cdot n_{air}}$ (where $n_{GaInAs} = 3.4$ and $n_{air} = 1$ are the diffraction

coefficients of GaInAs and air, respectively). The material chosen for this purpose was SiN, which when deposited in high temperature sports a diffraction coefficient of 2 and its thickness can be set to 2000 Å.

Due to the high temperature involved in the SiN deposition process no photo resist can be applied prior to the deposition, and as a result liftoff process cannot be implemented. Hence, the SiN layer is dry etched by means of plasma. The main drawback of using plasma etching is that isolating substrate surfaces become conductive when hit by plasma. To solve this problem only metal pads are exposed, leaving all the rest of the wafer protected by photo resist during the dry etch, and so coated by the ARC. The mask used for this lithography is "anti-reflective", and the final structure of a PD at this level is illustrated in figure 3.10.

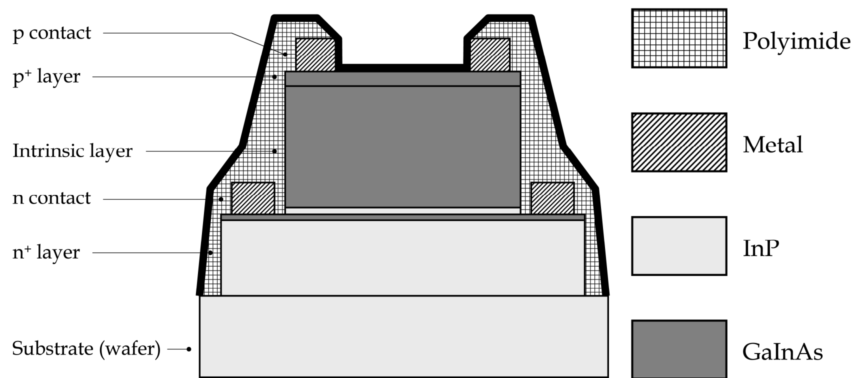


Figure 3.10: Cross section of a completed photodiode with anti-reflective coating

3.2.2 Photodiode Geometric Structure

The geometric structure of a photodiode is designed to minimize the junction capacitance, C_j, and the serial resistance, $r_{contact}$, of the diode. The thickness of the intrinsic layer, as well as the existence of an ARC, improve the responsivity, $\Re$. Before we proceed to PD structure, the process is summarized step by step in table 3.3 to facilitate referring to the different steps.

As mentioned before, the p contact is ring-shaped to allow passage of incident light through the contact opening. In addition, a $10 \times 5.5 \mu m^2$ rectangular pad is attached to the ring enabling the connection with metal 1 later on. Shown in figure 3.11a is a p contact of a small PD (10 μm opening). The ring isn't closed (and thus has a pincers shape) due to processing considerations (steps D and E in table 3.3). Note the rectangular pad evident in the upper center of this picture.

Junction area is basically determined by the p contact. Since wet etching of GaInAs placed under metal surface is considerably faster than naked GaInAs, and the junction

Step Index	Mask Number	Mask/Process Name	Description	Sketch Figure	SEM Figure
A		Cap Layer Removal	Cap layer removal by wet etch		
B	1	*Diode Mesa*	Lithography for diode mesa etch		
C		Diode Etch	Diode wet etch	3.6	
D	2	*Contact Metal*	Lithography for p and n contacts		
E		Metal Deposition	Contact metal deposition and liftoff	3.7	3.11a
F	3	*Isolation*	Lithography and n^+ layer wet etch	3.8	
G		Polyimide Spread	Polyimide spread by spinning for PD protection and isolation		
H	4	*Polyimide Cover*	Lithography and Polyimide dry etch for photodiode protection and isolation; contacts and optical window are opened	3.9	3.11b 3.11c
I	5	*Metal 1*	Lithography for metal 1		
J		Metal Deposition	Metal 1 deposition and liftoff		3.11d
K		SiN Deposition	SiN deposition for ARC		
L	6	*Anti-Reflective*	Lithography and SiN dry etch	3.10	

Table 3.3: PD fabrication process steps

is relatively thick, a margin of 2 μm is added to diode mesa area further than the p contact (steps B and C). At this level all of the small signal elements are known: junction capacitance, C_j, is determined by the diode mesa area, and contact resistance, $r_{contact}$, is dominated by the internal perimeter of the p contact (contribution of the n contact to $r_{contact}$ is negligible). Apparently, the transit time, τ_{PD}, is set by the layer structure, and has nothing with processing and etching.

The etching involved in the isolation process (step F) and the Polyimide covering (steps G and H) include both the diode mesa and the n contact, and so form the final PD footprint on the wafer. In the case of 25 μm PDs discussed in this section, the footprint sums up to a $62 \times 52 \mu m^2$ rectangle. Vias to the contacts and an optical window are left open during the Polyimide etch (step H), as shown in figure 3.11b and 3.11c. Metal 1 connects to these vias (steps I and J), as illustrated in figure 3.11d.

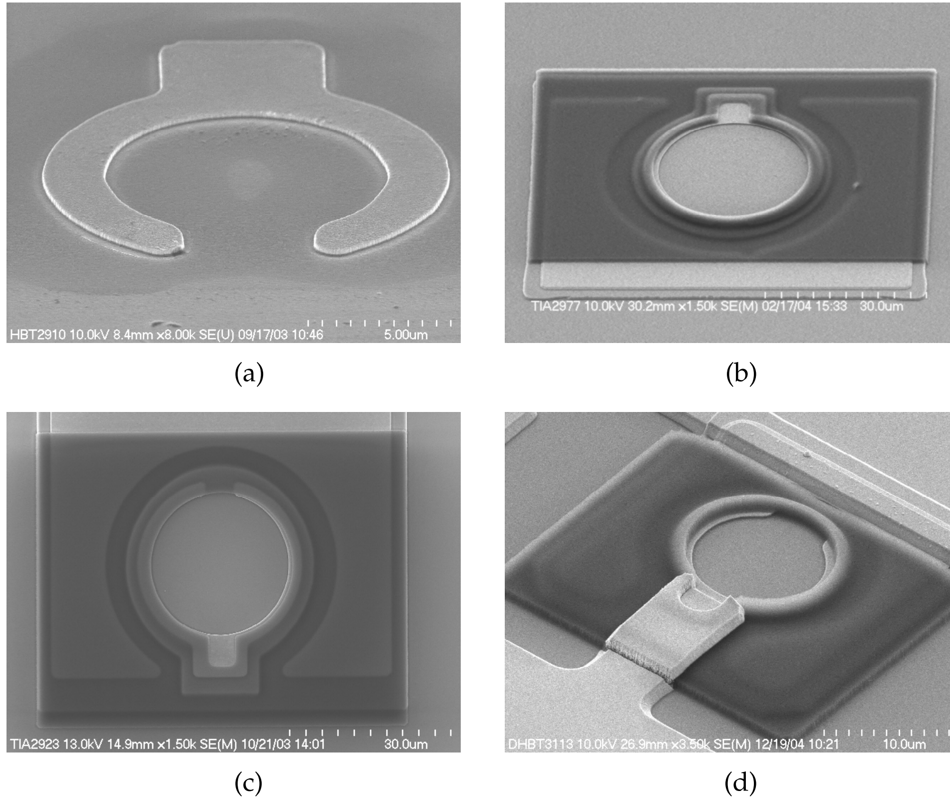


Figure 3.11: SEM images of a photodiode at various fabrication process steps: (a) contact metal deposition and liftoff (b) polyimide cover — bird-eye view (c) polyimide cover — top view (d) metal 1 deposition and liftoff

3.3 Photodiode Measurements and Characterization

Like HBTs, photodiodes require characterization to pave the way towards a simulation model. Despite of the simplicity of the PD model, its measurement procedures include sophisticated optoelectronic measurements. In order to extract the PD's characters and model measurements of four categories are carried out: DC, voltage–capacitance, RF small signal, and optoelectronic small signal.

Some of the measurements are similar to those taken form a HBT: I–V curve of the junction is measured to characterize leakage currents under reverse bias conditions (see section 2.3.1); C–V curve is used for choosing the optimal voltage to be used for biasing the PD, and for estimating the capacitance of a PD (refer to section 2.3.3); and a one-port RF measurement is taken from a small area PD in order to find C_j and $r_{contact}$. Nonetheless, the optoelectronic measurements require a measurement system of their own.

3.3.1 Measurement Equipment and Setup

The method used for optoelectronic measurement is the optical heterodyne detection [9][10]. The measurement system consists of two tunable laser sources, 50/50 optical coupler, electrical spectrum analyzer, and appropriate probes and cables. The laser sources (Anritsu MG9638A and HP 8168C) generate beams slightly shifted in wavelength. Both beams are inserted to a 50/50 coupler (EmiTek), so that the power at the coupler's output is modulated with the frequency that equals to the frequency difference between the two laser sources. The modulated optical signal is then applied to the PD under test. Finally, the electrical output of the PD is measured by a spectrum analyzer (Agilent E4446A). A schematic diagram of the system mentioned above is depicted in figure 3.12.

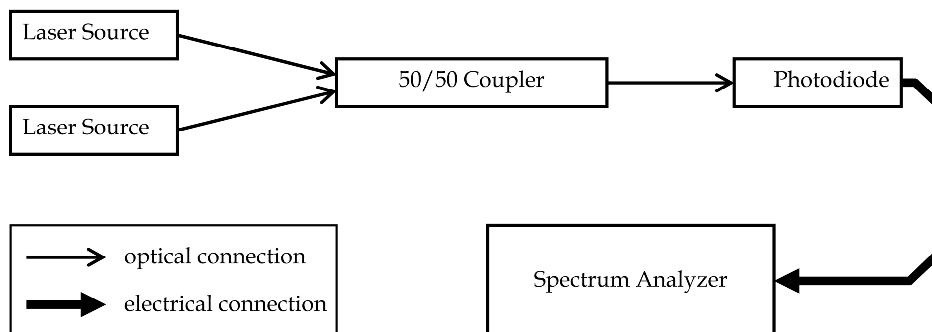
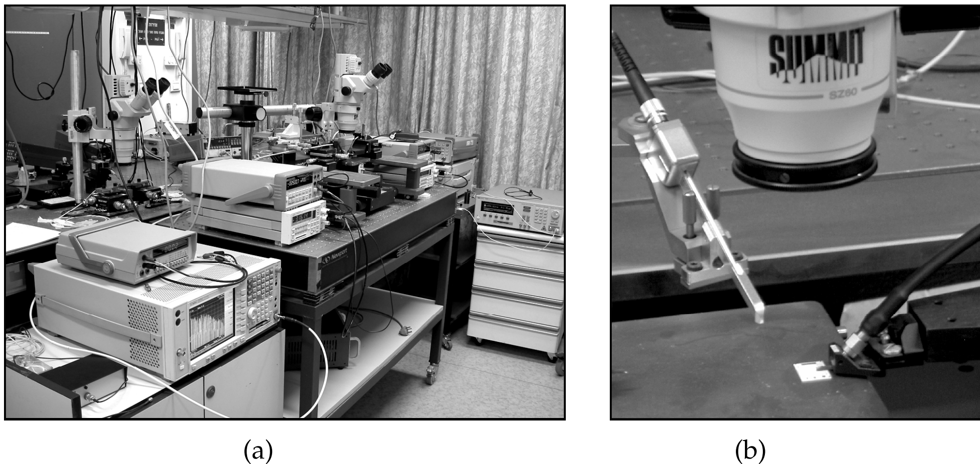


Figure 3.12: Diagram of the optoelectronic measurement system

Biasing is applied to the PD through an external bias tee (not illustrated in figure 3.12). Pictures of the entire system and the probes used for probing a PD are shown in figure 3.13.



(a) (b)

Figure 3.13: Pictures of the optoelectronic measurement system: (a) general view (b) optical and electrical probes for PD probing. Devices in (a) are (from left): spectrum analyzer, power supplies, probe station, power supplies, and tunable lasers.

In order to facilitate the measurement one of the tunable lasers is programmed to sweep over several wavelengths so that the modulation varies from 500 MHz to 40 GHz. In addition, the spectrum analyzer is configured to trace the maximum value at each frequency. Extraction of the measured data points from the curve is apparent.

3.3.2 Responsivity Measurement

Besides its own value, measurement of the responsivity of a PD is a helpful manner to locate the laser beam above the center of the PD. The optical DC power (P_{opt}) is measured by an optical power meter, and the DC current flowing in the PD (I_{PD}) is measured. Once maximum photo current is observed proper position of the optical probe is assured, and derivation of the responsivity and external quantum efficiency (from (3.5) and (3.6)) is evident. Results of responsivity measurements of PDs fabricated for this work are listed in table 3.4.

3.3.3 Measurement and Processing

Measurement procedure and data processing are described hereinafter.

Internal Diameter [μm]	Intrinsic Layer Thickness [Å]	ARC	Responsivity $\Re\left[\frac{A}{W}\right]$	External Quantum Efficiency, η
25	8000	Yes	0.586	0.469
20	8000	No	0.344	0.275
15	8000	No	0.344	0.275
10	8000	No	0.344	0.275
20	6000	No	0.180	0.144
15	6000	No	0.180	0.144
10	6000	No	0.180	0.144

Table 3.4: Measured responsivity and external quantum efficiency of various PDs

Measurement Procedure

An optoelectronic measurement is taken from the PD under test, using the system illustrated in figure 3.12. The biasing voltage of the PD is derived from the C–V measurement, choosing the minimum voltage (in its absolute value) that achieves widest depletion region. A one-port S–parameter measurement is taken from the PD (by means of network analyzer) to extract its resistance and capacitance without lighting.

Small Signal Model Extraction

The measured data are intended for the construction of a small signal model, to be utilized later on for circuit simulations. The parameters are used as follows:

- Data from spectrum analyzer indicate the –3dB bandwidth of the PD when loaded by a 50 Ω load. Gain is normalized to its low frequency value.
- One-port S–parameters (S_{22}, for instance) are used for extraction of C_j and $r_{contact}$.

Firstly, S–parameters are converted to Y–parameters. The capacitance of the PD is given by

$$C_{port2} = C_j + C_{pad} = \frac{\operatorname{Im}(Y_{22})}{\omega} \tag{3.11}$$

Since pad capacitance, C_{pad}, is known, derivation of C_j is obvious. Shown in figure 3.14 is an extraction of C_{port2}. Like in HBT model extraction, the capacitance value is valid only in the intermediate frequencies, since in high frequencies the impedance increases due to inductance of the metal line connecting the PD to its pad.

Extraction of $r_{contact}$ makes use of TLM measurement results (see section 2.3.2). Once Z, L_T, and R_c are known from a TLM measurement, $r_{contact}$ is estimated by

$$r_{contact} \approx r_c \cdot \frac{P_{in}}{Z} \tag{3.12}$$

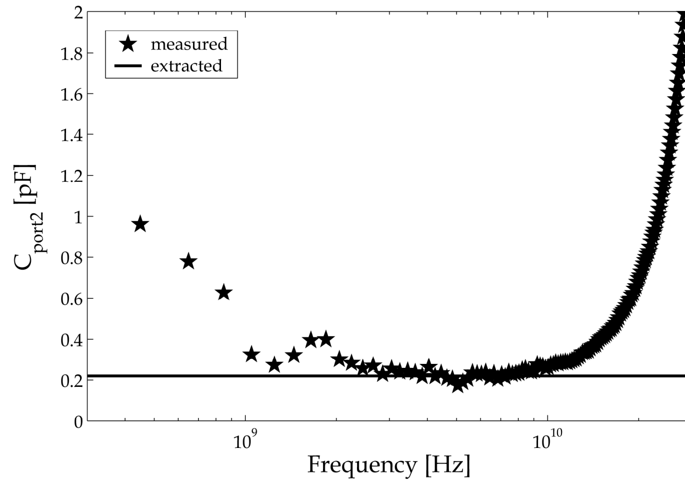


Figure 3.14: Extraction of C_j

where P_{in} is the internal perimeter of the p contact, and provided that the width of the ring-shaped contact is greater than L_T.

The product $(r_{contact} + 50\Omega)\, C_{port2}$ sets a pole in the PD frequency response. The frequency at which this pole occurs can confirm the extraction of C_j and $r_{contact}$. In addition, another pole exists at $\omega = \frac{1}{\tau_{PD}}$. This pole is extracted from the optoelectronic measurement and compared to (3.9).

Actual Extracted Parameters

The parameters extracted from the PDs used in this work are detailed in table 3.5. As ev-

Internal Diameter [μm]	Intrinsic Layer Thickness [Å]	C_j [fF]	$R_{contact}$ [Ω]	τ_{PD} [psec]
25	8000	180	7	—
20	8000	162	8	35
15	8000	127	12	—
10	8000	98	16	20
20	6000	190	8	28
15	6000	153	12	25
10	6000	115	16	19

Table 3.5: Extracted small signal parameters of various PDs

ident in table 3.5, the transit times are considerably longer than expected (see figure 3.3). We have expected $\tau_{PD} = 4$ psec for the 8000 Å PDs, and 3 psec for the 6000 Å diodes. In addition, the measured junction capacitances are higher than expected. This is in accordance with C–V measurements performed on the wafers, showing (in figure 3.15) that the 8000 Å collector wafer exhibits only 6000 Å depleted, and only 4500 Å is depleted in the 6000 Å collector wafer. The problem is caused by unwanted contaminations involved in the layer growth process, that increase the doping level in the intrinsic layer. The above reduces the bandwidth of the PDs in two ways: (a) the thinner junction increases C_j, and (b) charge carriers generated in the neutral region transport by means of diffusion, that is much slower than drift in a depletion region. As a consequence, τ_{PD} increases. Using (2.16) with $D_h = 4\ \frac{\text{cm}^2}{\text{sec}}$ and $W = 2000$ Å yields transit time of 50 psec, which explains the significant slowing of the PDs.

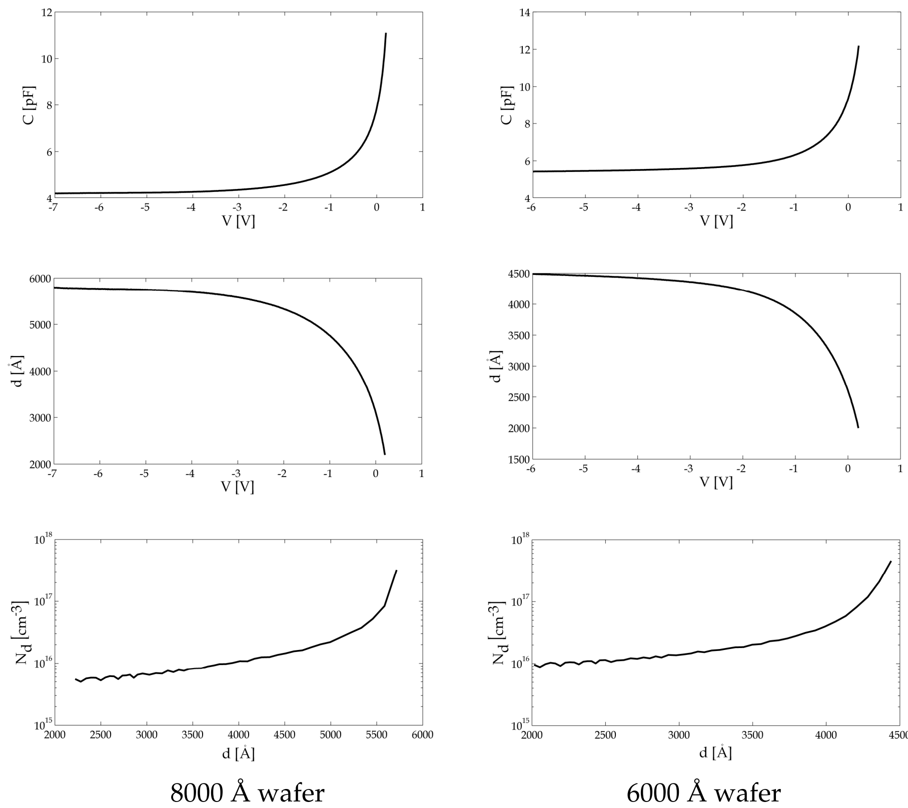


Figure 3.15: C–V measurements and processed data of both wafers.
Top: C–V curve; middle: depletion region width vs. applied voltage; bottom: doping concentration vs. distance from the metallurgic junction. Measured on a 210 x 100 μm^2 large area diode.

3.4 Noise in Photodiode

This section deals with sources of noise in a PD and their impact on PD noise performance. Also discussed here is the effect of noise exhibited by a PD on the performance of the entire optoelectronic system. The following chapters will make use of the equations derived here to calculate the noise performance of the optoelectronic integrated circuits discussed there.

3.4.1 Noise Mechanisms in Photodiode

As a PN–junction the noise of a PD consists of shot noise — generated within the junction, and Johnson noise — generated by parasitic and load resistances. Since in ultra-fast optical communications only high frequency behavior is of interest, flicker noise is neglected.

The total DC current flowing in the PD and contributing to its noise is given by [2]

$$I_n = I_{PD} + I_{dark} + I_{bg} \tag{3.13}$$

where I_{PD} is the DC photo current, I_{dark} is the dark current, a result of thermal generation in the depletion region, and I_{bg} is the photo current generated by background undesirable radiation (usually ambient light). From (2.66) this current introduces a shot noise current of

$$\overline{i_{sn}^2} = 2qI_n\Delta f \tag{3.14}$$

On top of that, the serial resistance of the PD, $r_{contact}$, and the load resistance attached to the PD (denoted by R_L), generate thermal noise current expressed by (according to (2.71))

$$\overline{i_{tn}^2} = \frac{4KT\Delta f}{(r_{contact} + R_L) \parallel r_{out}} \tag{3.15}$$

In (3.15), r_{out} refers only to the output resistance originated by leakages with ohmic behavior. The total noise current can therefore be represented by a single current source, as shown in figure 3.16. As the noise sources has no correlation, the total noise current is simply

$$\overline{i_n^2} = \overline{i_{sn}^2} + \overline{i_{tn}^2} \tag{3.16}$$

The relations between noise current and signal detection are described in the next section.

3.4.2 Relations Between Noise and Detected Signal

In communication system design signal-to-noise ratio (*SNR*), in terms of power, is of major importance. In order to calculate SNR, we denote the RMS input power and current

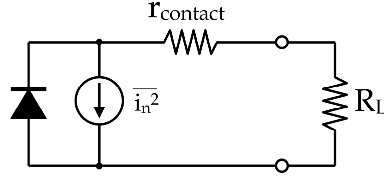


Figure 3.16: Noise source of a photodiode

by p_{opt} and i_{PD}, respectively. According to (3.4), RMS current is given by

$$i_{PD} = \frac{\lambda}{hc} \cdot q\eta p_{opt} \tag{3.17}$$

and SNR can be written as follows [2]:

$$SNR = \frac{i_{PD}^2 \left(r_{contact} + R_L\right)}{\overline{i_n^2} \left(r_{contact} + R_L\right)} = \frac{\left(\frac{\lambda}{hc} \cdot q\eta p_{opt}\right)^2}{2qI_n\Delta f + \frac{4KT\Delta f}{(r_{contact}+R_L)\|r_{out}}} \tag{3.18}$$

The smallest signal that can be detected by a PD is the power that generates a photo current as large as the noise current. This limit is called "noise-equivalent power" (*NEP*) and derived from the input power, p_{opt}, at which SNR is unity. Substituting unity in (3.18) and normalizing to frequency yields

$$NEP = \frac{hc}{\lambda} \cdot \frac{1}{q\eta} \sqrt{2qI_n\Delta f + \frac{4KT}{\left(r_{contact} + R_L\right) \parallel r_{out}} \Delta f} \tag{3.19}$$

It can be concluded from (3.19) that in order to minimize NEP (i.e. reduce the noise) η and R_L should be as large as possible, and I_{dark} and I_{bg} should be as small as possible. Nonetheless, the higher the R_L, the lower the –3dB bandwidth of the PD is.

References for Chapter 3

[1] David J. Roulston, *Bipolar Semiconductor Devices*, McGraw–Hill, 1990.

[2] Pallab Bhattacharya, *Semiconductor Optoelectronic Devices*, Prentice–Hall, second edition, 1997.

[3] Mukunda B. Das, "Optoelectronic detectors and receivers: Speed and sensitivity limits", *Proceedings of the 1998 Conference on Optoelectronic and Microelectronic Materials Devices*, pp. 15–22, December 1988.

[4] T. H. Windhorn, L. W. Cook, and G. E. Stillman, "The electron velocity-field characteristic for n-InGaAs at 300k", *IEEE Electron Device Letters*, vol. 3, no. 1, pp. 18–20, January 1982.

[5] D. Wake, R. H. Walling, I. D. Henning, and D. G. Parker, "Planar-junction, top-illuminated GaInAs/InP pin photodiode with bandwidth of 25 GHz", *Electronics Letters*, vol. 25, no. 15, pp. 967–969, July 1989.

[6] P. Hill, J. Schlafer, W. Powazinik, M. Urban, E. Eichen, and R. Olshansky, "Measurement of hole velocity in n-type InGaAs", *Applied Physics Letters*, vol. 50, no. 18, pp. 1260–1262, May 1987.

[7] D. A. Humphreys, R. J. King, D. Jenkins, and A. J. Moseley, "Measurement of absorption coefficients of $Ga_{0.47}In_{0.53}As$ over the wavelength range of 1.0–1.7 μm", *Electronics Letters*, vol. 21, no. 25/26, pp. 1187–1189, December 1985.

[8] Hong Wang and Geok Ing Ng, "Electrical properties and transport mechanisms of InP/InGaAs HBTs operated at low temperature", *IEEE Transactions on Electron Devices*, vol. 48, no. 8, pp. 1492–1497, August 2001.

[9] Satoki Kawanishi, Atsushi Takada, and Masatoshi Saruwatari, "Wide-band frequency-response measurement of optical receivers using optical heterodyne detection", *IEEE Journal of Lightwave Technology*, vol. 7, no. 1, pp. 92–98, January 1989.

[10] D. A. Humphreys and C. A. Park, "High accuracy frequency response measurements of mm-wave photodiodes using a DFB heterodyne system with a novel detection scheme", *IEE Colloquium on Optical Detectors*, pp. 6/1–6/5, January 1990.

Chapter 4

Optoelectronic Lumped Circuit

Design of a lumped preamplifier for optical communications is the subject of this chapter. Lumped amplifier is one approach of implementing a preamplifier, that can employ feedback and simple biasing topologies and exploit their advantages. The chapter begins with a short introduction on the building blocks of a bipolar amplifier. The following sections introduce and justify the topologies chosen for this work — both from electrical and optoelectronic viewpoints. The third and last part of the chapter summarizes the measurements, characterization, and performance of the circuits fabricated for this work.

4.1 Amplifier Building Blocks

The bipolar transistor has three basic configurations used in analog circuits — common emitter, common base, and common collector — each exhibits different behavior in terms of gain, bandwidth, node impedances, and noise performance. Consider the configurations shown in figure 4.1. In common emitter (*CE*) configuration the emitter is connected

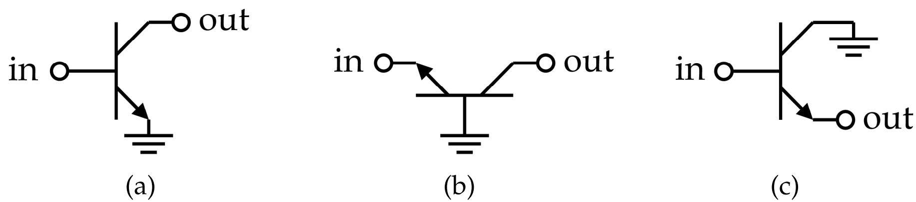


Figure 4.1: Bipolar transistor configurations: (a) common emitter (b) common base (c) common collector. Note that ground symbol refers to AC ground.

to an AC ground, the base is used as the network input, and the collector is the output. In common base (*CB*) configuration the base is grounded, the emitter acts as the input, and the collector is the output. Finally, in common collector (*CC*) configuration the base and the emitter act as input and output, respectively, and the collector is grounded. The analysis of all configurations can be carried out with the T model (figure 2.38), but CE and CC analysis becomes easier with the hybrid-π model — illustrated in figure 4.2. This model necessitates new nomenclature for the model elements. The customary notations that differ from the T model are detailed in table 4.1.

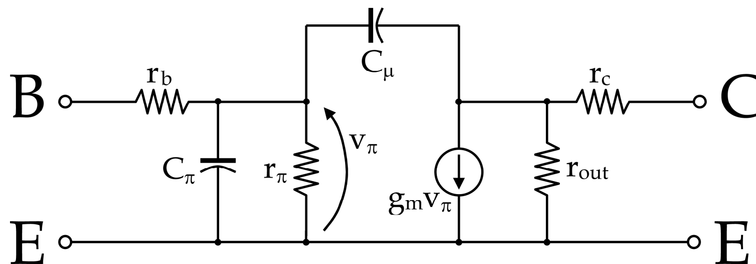


Figure 4.2: Small signal hybrid-π model of a bipolar transistor, neglecting r_{ee}

Element	Description	Expressed by T model terms
r_π	BE resistance	$r_\pi = r_e(\beta + 1)$
C_π	BE capacitance	$C_\pi = C_{BE}$
C_μ	BC capacitance	$C_\mu = C_{BC}$
g_m	Transconductance	$g_m = \frac{qI_C}{KT}$

Table 4.1: Hybrid-π terms expressed by T model terms

4.1.1 Bipolar Single-Stage Amplifier

The amplifiers presented in this work consist of several stages, each incorporates a single HBT to provide the required gain or buffering. Derivation of the characteristics of the available single-stage amplifiers are outlined in this section. In the below calculations r_c is neglected, and r_{out} is usually considered as infinite. However, due to the relatively low current gain (β) of HBTs the common approximations made in most textbooks were reconsidered in the calculations here.

Common Emitter

When utilizing CE configuration as an amplification stage a resistor, R_C, is externally connected between the emitter and the supply voltage, V_{CC}. This topology is depicted in figure 4.3a. Analyzing the small signal model of this circuit, shown in figure 4.3b, yields voltage gain of [1]

$$a_v = -g_m\left(R_C \parallel r_{out}\right) \cdot \frac{r_\pi}{r_\pi + r_b} \approx -g_m R_C \tag{4.1}$$

assuming that r_b is negligible in comparison to r_π. The sign of a_v is negative, that is, CE stage is a phase inverting stage. As expected, the current gain of the circuit, when loaded by short circuit, is

$$a_i = \beta \tag{4.2}$$

The input resistance is given by

$$R_{in} = r_b + r_\pi \approx r_\pi \tag{4.3}$$

whilst the total input impedance is given by

$$Z_{in} = r_b + \left(r_\pi \parallel \frac{1}{sC_\pi} \parallel \frac{1}{sC_\mu a_v}\right) \tag{4.4}$$

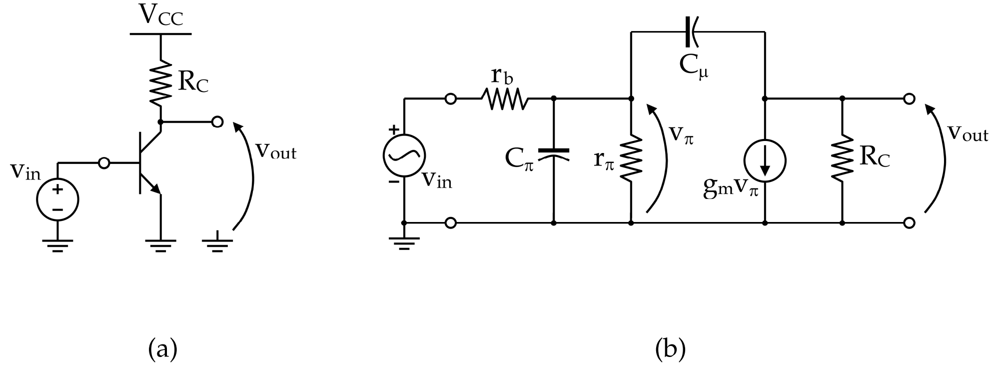


Figure 4.3: Common emitter stage: (a) circuit topology (b) small signal representation

The multiplication of C_μ by a_v is a result of Miller effect, as C_μ connects the stage's input with its output (refer to appendix B). The output resistance is expressed by

$$R_{out} = r_{out} \parallel R_C \approx R_C \tag{4.5}$$

and the total output impedance is

$$Z_{out} = r_{out} \parallel R_C \parallel \frac{1}{sC_\mu} \tag{4.6}$$

When examining the bandwidth of a_v the source and load resistances, denoted by R_S and R_L respectively, should be taken into account. Base resistance, r_b, can be lumped in R_S in serial connection, and R_C should be lumped in R_L in parallel connection. By deriving the frequency dependency of a_v one obtains

$$a_v = -g_m R_L \cdot \frac{r_\pi}{r_\pi + R_S} \cdot \frac{1 - \frac{s}{\omega_z}}{\left(1 - \frac{s}{\omega_{p1}}\right)\left(1 - \frac{s}{\omega_{p2}}\right)} \tag{4.7}$$

where

$$\omega_z = \frac{g_m}{C_\mu} \tag{4.8}$$

$$\omega_{p1} = -\frac{r_\pi + R_S}{r_\pi R_S} \cdot \frac{1}{C_\pi + C_\mu \left(1 + g_m R_L + \frac{R_L}{r_\pi \| R_S}\right)} \tag{4.9}$$

$$\omega_{p2} = -\left(\frac{1}{C_\mu R_L} + \frac{1}{C_\pi R_L} + \frac{1}{C_\pi \left(r_\pi \parallel R_S\right)} + \frac{g_m}{C_\pi}\right) \tag{4.10}$$

are the angular frequencies of the zero and the poles of (4.7). The zero's frequency, ω_z, is extremely high, and thus its effect is ignored.

Common Emitter with Degeneration Resistor

In the CE circuit mentioned above I_C is exponential with the DC voltage of the input. Introducing a resistor between the emitter and ground "degenerates" this exponential relation and makes it approximately linear. Such a resistor, denoted by R_E, also modifies the AC behavior of the stage. The topology and small signal representation of a CE stage with degeneration resistor are shown in figure 4.4.

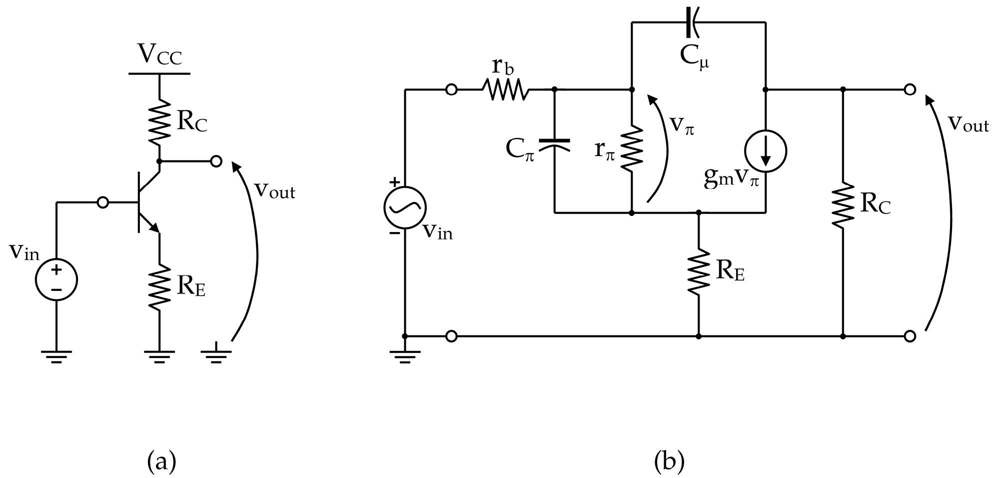


Figure 4.4: Common emitter stage with degeneration resistor: (a) circuit topology (b) small signal representation

The serial resistance of the emitter, r_{ee}, can be lumped in with R_E rather than neglected. The transconductance of a degenerated CE stage is [1]

$$G_m = \frac{g_m}{1 + g_m R_E} \tag{4.11}$$

Equation (4.11) implies that G_m is smaller than g_m by a factor $1 + g_m R_E$. The voltage gain of the stage is given by

$$a_v = -G_m \left[R_C \parallel r_{out}\left(1 + g_m R_E\right)\right] \cdot \frac{r_\pi}{r_\pi + r_b} \approx -G_m R_C \tag{4.12}$$

If R_E is significantly greater than $\frac{1}{g_m}$ then a_v reduces to

$$a_v \approx -\frac{R_C}{R_E} \tag{4.13}$$

The degeneration resistor doesn't change current gain, namely

$$a_i = \beta \tag{4.14}$$

The input resistance is given by

$$R_{in} = r_b + r_\pi + R_E(\beta + 1) \approx r_\pi + R_E(\beta + 1) \tag{4.15}$$

and the total input impedance is given by

$$Z_{in} = r_b + \left(r_\pi \parallel \frac{1}{sC_\pi} \parallel \frac{1}{sC_\mu a_v} \right) + R_E(\beta + 1) \tag{4.16}$$

The output resistance is expressed by

$$R_{out} = r_{out}\left[1 + g_m\left(r_\pi \parallel R_E\right)\right] \parallel R_C \approx R_C \tag{4.17}$$

and the total output impedance is

$$Z_{out} = r_{out}\left[1 + g_m\left(r_\pi \parallel R_E\right)\right] \parallel R_C \parallel \frac{1}{sC_\mu} \tag{4.18}$$

Since the degeneration resistor, R_E, acts as a feedback network [1] the bandwidth of such a stage is greater than CE stage without degeneration.

Common Base

Shown in figure 4.5a is CB topology, accompanied by a small signal representation, based on the T model, in figure 4.5b. Base resistance, r_b, is neglected. Examination of the circuit yields transconductance of

$$G_m = \frac{\alpha}{r_{ee} + r_e + \frac{r_b}{\beta+1}} \approx g_m \frac{r_\pi}{r_\pi + r_b} \approx \frac{\alpha}{r_{ee} + r_e} \tag{4.19}$$

thus the voltage gain equals to

$$a_v = G_m\left(R_C \parallel r_{out}\right) \approx G_m R_C \approx \frac{R_C}{r_{ee} + r_e} \tag{4.20}$$

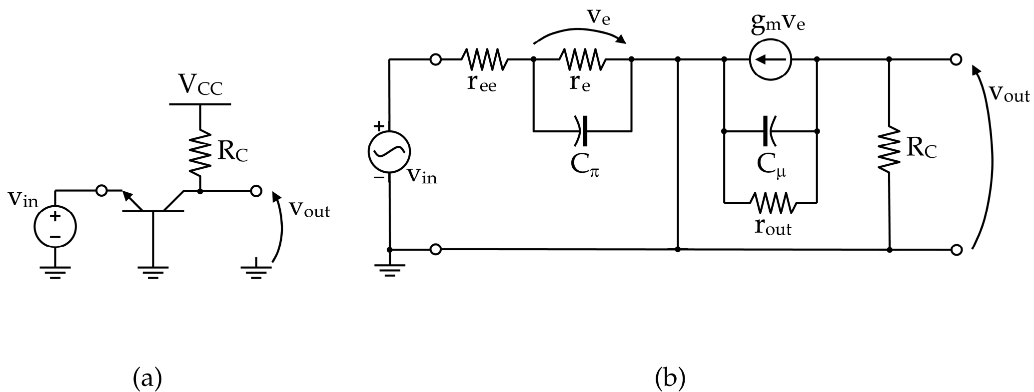


Figure 4.5: Common base stage: (a) circuit topology (b) small signal representation

and the current gain of the circuit, when loaded by short circuit, is

$$a_i = \alpha \tag{4.21}$$

The input resistance is given by

$$R_{in} = r_{ee} + r_e + \frac{r_b}{\beta + 1} \approx r_{ee} + r_e \tag{4.22}$$

The total input impedance is given by

$$Z_{in} = r_{ee} + \left(r_e \parallel \frac{1}{sC_\pi} \right) + \frac{r_b}{\beta + 1} \approx r_{ee} + \left(r_e \parallel \frac{1}{sC_\pi} \right) \tag{4.23}$$

The output resistance of the stage is expressed by

$$R_{out} = (r_{out} + r_b) \parallel R_C \approx R_C \tag{4.24}$$

and the total output impedance is

$$Z_{out} = \left[\left(r_{out} \parallel \frac{1}{sC_\mu} \right) + r_b \right] \parallel R_C \approx R_C \parallel \frac{1}{sC_\mu} \tag{4.25}$$

Derivation of the bandwidth of a_i yields frequency dependency as follows:

$$a_i = \frac{\alpha}{1 - \frac{s}{\omega_p}} \tag{4.26}$$

where

$$\omega_p \approx -\frac{g_m}{C_\pi} \approx -\omega_T \tag{4.27}$$

and ω_T is the current gain cutoff angular frequency of the transistor ($\omega_T = 2\pi f_T$).

Common Collector (Emitter Follower)

The common collector stage, also known as "emitter follower", consists of the topology depicted in figure 4.6a. Its small signal equivalent circuit is illustrated in figure 4.6b, based on the hybrid-π model. In the following equations, r_{ee} and r_{out} are neglected, and r_b is lumped together with the source resistance as R_S. Load resistance, if existent, can be included in R_E as they are connected in parallel to each other. The voltage gain of an emitter follower is

$$a_v = \frac{1}{1 + \frac{r_\pi + R_S}{R_E(\beta+1)}} \approx 1 \tag{4.28}$$

and the current gain of the circuit, when loaded by short circuit, is

$$a_i = -(\beta + 1) \tag{4.29}$$

The input resistance of a CC stage is given by

$$R_{in} = r_b + r_\pi + (\beta + 1) R_E \tag{4.30}$$

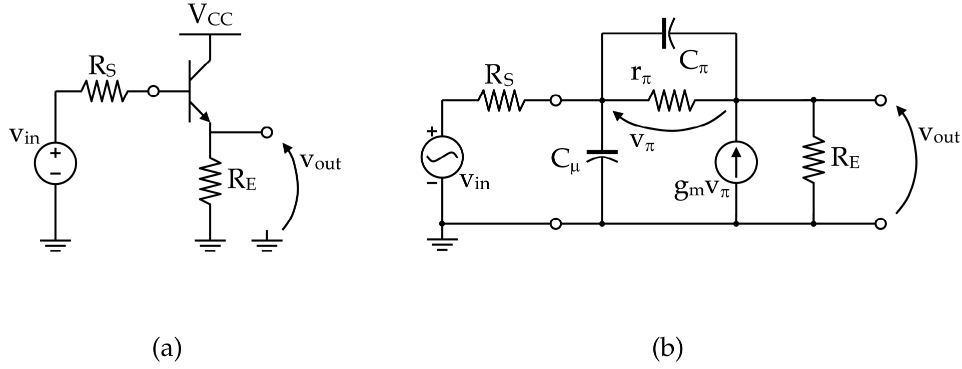


Figure 4.6: Common collector stage: (a) circuit topology (b) small signal representation

and the total input impedance is given by

$$Z_{in} = r_b + \frac{1}{sC_\mu} \parallel \left[\left(r_\pi \parallel \frac{1}{sC_\pi} \right) + (\beta + 1) R_E \right] \tag{4.31}$$

The output resistance of the stage is

$$R_{out} = r_{ee} + r_e + \frac{R_S}{\beta + 1} \tag{4.32}$$

and the total output impedance is

$$Z_{out} = r_{ee} + \left(r_e \parallel \frac{1}{\beta + 1} \cdot \frac{1}{sC_\pi} \right) + \frac{R_S}{\beta + 1} \tag{4.33}$$

Frequency dependency of a_v is of the form

$$a_v = \frac{1}{1 + \frac{r_\pi + R_S}{R_E(\beta+1)}} \cdot \frac{1 - \frac{s}{\omega_z}}{1 - \frac{s}{\omega_p}} \tag{4.34}$$

where

$$\omega_z \approx -\frac{g_m}{C_\pi} \approx -\omega_T \tag{4.35}$$

$$\omega_p = -\frac{1}{C_\pi \left(r_\pi \parallel \frac{R_S + R_E}{1 + g_m R_E} \right)} \tag{4.36}$$

are the angular frequencies of the zero and the pole of (4.34). Inclusion of C_μ in the calculation of the frequency response introduces another pole to (4.34). It should be noted that loading a CC stage with capacitive load can move the poles and the zero, and result in instability.

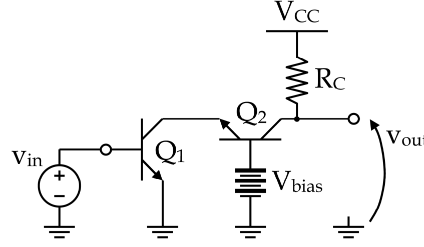


Figure 4.7: Circuit topology of a cascode stage

Cascode

Cascode is a two-transistor stage that consists of a CE stage followed by a CB stage, as shown in figure 4.7. The main idea that underlies the cascode topology is to load the CE stage (Q1) with low resistance, namely, the low input resistance of the CB stage (Q2). This reduces the voltage gain of Q1, a_{v1}, and thus eliminates the Miller effect on $C_{\mu 1}$ (i.e. C_μ of Q1) and expands the bandwidth of Q1 (and the entire cascode stage as well). The total voltage gain of the cascode is still determined by g_{m1} and R_C, or, when using degeneration resistance, R_E and R_C, thus

$$a_v = -g_{m1} \left(R_C \parallel \beta \cdot r_{out2} \right) \cdot \frac{r_{\pi 1}}{r_{\pi 1} + r_{b1}} \approx -g_{m1} R_C \tag{4.37}$$

similarly to (4.1). In the degenerated case the gain is similar to (4.12). The current gain is given by

$$a_i = \beta \cdot \alpha \approx \beta \tag{4.38}$$

just like a CE stage. The input resistance equals to R_{in} of a CE stage (see (4.3) and (4.15)), but the total input impedance ((4.4) or (4.16)) lacks the term of C_μ because a_{v1} is very small. The output resistance of a cascode stage is given by

$$R_{out} \approx \beta \cdot r_{out2} \parallel R_C \approx R_C \tag{4.39}$$

Bandwidth of the cascode is calculated by the zero-value time constant method [1]. The time constant obtained by this fashion is

$$\tau_{-3dB} \approx C_{\pi 1} R_{\pi 1} + C_{\mu 1} R_{\mu 1} + C_{\pi 2} R_{\pi 2} + C_{\mu 2} R_{\mu 2} \tag{4.40}$$

where

$$R_{\pi 1} = r_{\pi 1} \parallel \frac{R_S + R_E}{1 + g_{m1} R_E} \tag{4.41}$$

$$R_{\mu 1} = R_L + (1 + G_m R_L) \left[(r_{\pi 1} + (\beta + 1) R_E) \parallel R_S \right] \tag{4.42}$$

$$R_{\pi 2} = r_{\pi 1} \parallel r_{e2} \tag{4.43}$$

$$R_{\mu 2} = r_{b2} + R_L \tag{4.44}$$

Like in CE, r_b is included in the source resistance, R_S, and R_C is connected in parallel to the load resistance so that both form R_L. R_E is a degeneration resistor, if existent, and G_m is evaluated from (4.11). The bandwidth (i.e. –3dB frequency) is therefore

$$f_{-3dB} = \frac{1}{2\pi\tau_{-3dB}} \tag{4.45}$$

which is significantly higher than the bandwidth of a "plain" CE stage.

4.1.2 Feedback

Being a lumped circuit, the circuit discussed in this chapter utilizes the benefits of feedback network. This section concisely introduces the basic concepts of feedback and their effect on circuit design.

Feedback Effect on Gain

Consider the negative-feedback network shown in figure 4.8. The transmission function of the amplifier is $A(\omega)$, and F is the function of the feedback network. The closed-loop

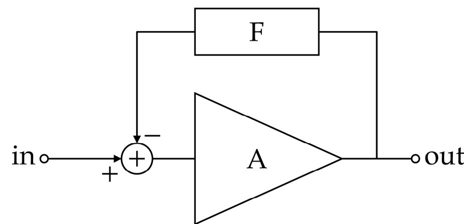


Figure 4.8: Negative-feedback network scheme

gain is given by [1][2]

$$G_{CL} = \frac{A}{1 + AF} \approx \frac{1}{F} \tag{4.46}$$

The approximation in (4.46) assumes that $AF \gg 1$, which is the case for most feedback circuits. In addition, it is evident from (4.46) that the amplifier's gain, A, is divided by a factor of $(1 + AF)$ when the loop is closed.

Feedback Effect on Bandwidth

It's assumed that the amplifier of figure 4.8 has a pole at ω_0, that is,

$$A(\omega) = \frac{A}{1 - \frac{s}{\omega_0}} \tag{4.47}$$

Substituting (4.47) in (4.46) yields closed-loop gain of

$$G_{CL}(\omega) = \frac{A}{1 + AF} \cdot \frac{1}{1 - \frac{s}{(1+AF)\omega_0}} \tag{4.48}$$

Equation (4.48) implies that the bandwidth of the amplifier, ω_0, is multiplied by a factor of $(1 + AF)$ when the loop is closed.

Types of Feedback

A feedback network can be applied to an amplifier in four types of connections:

1. Serial-In/Parallel-Out (*SIPO*) — the feedback samples voltage at the amplifier's output and introduces voltage to its input. The closed-loop circuit behaves like a voltage amplifier.
2. Parallel-In/Serial-Out (*PISO*) — the feedback samples current at the amplifier's output and introduces current to its input. The closed-loop circuit behaves like a current amplifier.
3. Parallel-In/Parallel-Out (*PIPO*) — the feedback samples voltage at the amplifier's output and introduces current to its input. The closed-loop circuit behaves like a transimpedance amplifier (*TIA*).
4. Serial-In/Serial-Out (*SISO*) — the feedback samples current at the amplifier's output and introduces voltage to its input. The closed-loop circuit behaves like a transconductance amplifier (*TCA*).

The amplifiers designed for this work are TIAs (refer to section 4.3). Hence, they incorporate PIPO-type feedback networks.

PIPO Feedback and Input/Output Impedance

An operational amplifier (*opamp*) with PIPO feedback resistor is shown in figure 4.9. Being an opamp it has very high input impedance (denoted by $z_{in.a}$) and very low output impedance ($z_{out.a}$). The input impedance of the feedback TIA is therefore [1]

$$Z_{in} = \frac{R_f \parallel z_{in.a}}{1 + AF} \approx \frac{R_f}{a_v} \tag{4.49}$$

where R_f is the feedback resistor, a_v is the voltage gain of the opamp, and AF is the loop gain. If the circuit is loaded by R_L then the loop gain can be written as

$$AF = a_v \cdot \frac{z_{in.a}}{z_{in.a} + R_f} \cdot \frac{R_f R_L}{R_f R_L + z_{out.a} R_f + z_{out.a} R_L} \tag{4.50}$$

For $z_{in.a} \longrightarrow \infty$ and $z_{out.a} \longrightarrow 0$ the right hand of (4.50) reduces to a_v, and thus the approximation in (4.49) is justified. The output impedance of the feedback TIA is given

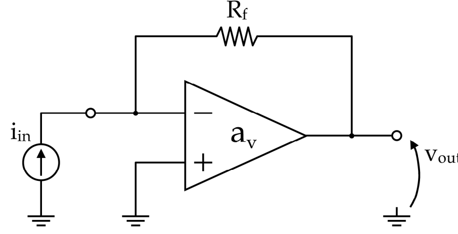


Figure 4.9: Scheme of opamp with PIPO feedback

by

$$Z_{out} = \frac{R_f \parallel z_{out.a}}{1 + AF} \approx \frac{z_{out.a}}{a_v} \tag{4.51}$$

Equations (4.49) and (4.51) show that PIPO feedback exhibits low impedance both at the input and the output of the circuit — as expected from a TIA.

4.1.3 Noise in Single- and Multiple-Stage Amplifiers

The noise mechanisms and model presented in section 2.6 are now applied to the different amplifying stages and to feedback. Since the circuits constructed for this work are intended for high frequencies, flicker and burst noises are neglected in the following discussion.

Noise of a circuit limits the minimum signal that can be detected by the circuit. As a consequence, noise generation mechanisms are converted to *input-equivalent noise generators* — all appear at the input of the circuit. Also here, recalculations have been required due to the unique nature of HBTs.

Noise in Common Emitter Stage

Using the noise model illustrated in figure 2.52, the noise of a CE stage can be represented by a voltage source introduced at the input of the stage:

$$\frac{\overline{v_n^2}}{\Delta f} = 4KTR_S + 2qI_B R_S^2 + \frac{1}{g_m^2} \cdot \frac{|Z_\pi + R_S|^2}{|Z_\pi|^2} \left(\frac{4KT}{R_L} + 2qI_C \right) \tag{4.52}$$

where

$$Z_\pi = r_\pi \parallel \frac{1}{sC_\pi} \tag{4.53}$$

is the base–emitter impedance, I_B and I_C are the DC currents of the base and the collector, respectively, R_S includes r_b, and R_L includes R_C in parallel. Alternatively, a current source can replace the voltage source at the input to represent the noise, namely

$$\frac{\overline{i_n^2}}{\Delta f} = \frac{\overline{v_n^2}}{\Delta f} \cdot \frac{|Z_\pi + R_S|^2}{|Z_\pi|^2 R_S^2} \tag{4.54}$$

which results in

$$\frac{\overline{i_n^2}}{\Delta f} = \frac{|Z_\pi + R_S|^2}{|Z_\pi|^2}\left[\frac{4KT}{R_S} + 2qI_B + \frac{1}{g_m^2 R_S^2}\cdot\frac{|Z_\pi + R_S|^2}{|Z_\pi|^2}\left(\frac{4KT}{R_L} + 2qI_C\right)\right] \tag{4.55}$$

Equations (4.52) and (4.55) imply that the noise generated by the collector and the load is attenuated when referred to the input, due to the gain of the stage. If the CE stage is degenerated be a resistor, R_E, then Z_π becomes

$$Z_\pi = \left(r_\pi \parallel \frac{1}{sC_\pi}\right) + R_E(\beta + 1) \tag{4.56}$$

the input noise voltage source is

$$\frac{\overline{v_n^2}}{\Delta f} = 4KTR_S + 2qI_B R_S^2 + \frac{1}{G_m^2}\cdot\frac{|Z_\pi + R_S|^2}{|Z_\pi|^2}\left(\frac{4KT}{R_E} + \frac{4KT}{R_L} + 2qI_C\right) \tag{4.57}$$

thus (4.55) becomes

$$\frac{\overline{i_n^2}}{\Delta f} = \frac{|Z_\pi + R_S|^2}{|Z_\pi|^2}\left[\frac{4KT}{R_S} + 2qI_B + \frac{1}{G_m^2 R_S^2}\cdot\frac{|Z_\pi + R_S|^2}{|Z_\pi|^2}\left(\frac{4KT}{R_E} + \frac{4KT}{R_L} + 2qI_C\right)\right] \tag{4.58}$$

As evident, the relatively low gain of a degenerated stage (accentuated by G_m, which is lower than g_m) increases the input referred noise. In addition, R_E contributes its own noise term.

Noise in Common Base Stage

Although the starting point for noise calculation is similar to CE stage, CB stage has no current gain (α, to be precise), and as a result, every noise current source introduced at the output of the stage is referred "as is" to the input. Therefore, the noise at the input can be represented as a current source as follows:

$$\frac{\overline{i_n^2}}{\Delta f} = 2qI_E + 4KT\left(\frac{1}{R_L} + \frac{1}{R_S}\right) \tag{4.59}$$

Here R_S includes r_{ee} and R_L includes r_c.

Noise in Common Collector Stage

Recall that CC stage acts as a voltage follower, with approximately unity voltage gain. Should a noise voltage source be applied to the output of the stage, it would be referred back to the input with no attenuation. The noise voltage at the input is therefore given by:

$$\frac{\overline{v_n^2}}{\Delta f} = 2qI_C\,|Z_e \parallel R_E \parallel Z_L|^2 + 2qI_B\left|\frac{1}{sC_\mu} \parallel Z_\pi\right|^2 + 4KT\,(R_L + R_S) \tag{4.60}$$

where Z_π is defined by (4.56), Z_L and R_L are the load impedance and resistance, respectively, R_S includes r_b, and

$$Z_e = r_{ee} + \left(r_e \parallel \frac{1}{\beta + 1} \cdot \frac{1}{sC_\pi} \right) \tag{4.61}$$

is the emitter impedance from the stage output viewpoint.

Noise in Cascode Stage and Multiple-Stage Amplifier

When more than one stage are connected in cascade, total noise performance should be calculated from the last (output) stage back to the first one. Each stage is loaded by the following stage and fed by the previous stage. Only the last stage is loaded by a "real" load, and the first stage is fed by a "real" source. Hence, attention should be paid not to include twice elements that act both as source and load. When referring back a noise source through a CE stage it is attenuated, as described above. As a rule of thumb in multiple-stage design, the higher the gain of the stage, the closer to the input it should be placed.

Applying the above method to a degenerated cascode stage (figure 4.7) yields an input noise current of

$$\frac{\overline{i_n^2}}{\Delta f} = D_{\pi s} \left[\frac{4KT}{R_S} + 2qI_{B1} + \frac{1}{G_{m1}^2 R_S^2} \cdot D_{\pi s} \left(\frac{4KT}{R_E} + 2qI_{C1} + 2qI_{E2} + \frac{4KT}{R_C} \right) \right] \tag{4.62}$$

where

$$D_{\pi s} = \frac{|Z_\pi + R_S|^2}{|Z_\pi|^2} \tag{4.63}$$

and assuming that the stage is loaded by an infinite resistance, so that $R_L \approx R_C$.

Feedback Effect on Noise

It can be shown the the existence of a feedback network doesn't change the input noise source of the amplifier [1][2][3]. However, noise generated by the feedback network is directly referred to the input of the circuit. For example, a PIPO feedback network implemented by means of a single resistor, R_f, contributes the following term to the total input current noise:

$$\frac{\overline{i_n^2}}{\Delta f} = \frac{4KT}{R_f} \tag{4.64}$$

In all of the calculations in this section the bandwidth, Δf, that should be taken is $\frac{\pi}{2}$ times the –3dB bandwidth of the examined circuit [3].

4.1.4 Summary

The following amplifying stages and network have been discussed: common emitter (with and without degeneration), common base, common collector, cascode, and feedback network. The details are generalized and summarized below.

- *Common Emitter* stage has voltage gain of $-g_m R_C$ (it's a phase inverting stage), current gain of β, and high input and output resistances. It has low bandwidth due to Miller effect that boosts C_μ.
- *Degenerated Common Emitter* stage has voltage gain of $-\frac{R_C}{r_e+R_E}$, current gain of β, and high input and output resistances. Its bandwidth is higher than a non-degenerated CE stage, but its gain is lower.
- *Common Base* stage acts as a current follower. It has voltage gain of $\frac{R_C}{r_e+r_{ee}}$ (it's a phase preserving stage), unity current gain, low input resistance, and high output resistance. Its bandwidth is as high as f_T of the transistor.
- *Common Collector (Emitter Follower)* stage acts as a voltage follower. It has unity voltage gain (it's a phase preserving stage), current gain of $-(\beta+1)$, extremely high input resistance, low output resistance, and bandwidth of approximately f_T. There are another pole and zero in the stage's transfer function that may cause instability, if the stage is loaded by a capacitive load.
- *Cascode Stage* is actually a CE stage followed by a CB stage. It has the same voltage gain, current gain, and input resistance like a CE stage, but extremely high output resistance. The bandwidth is significantly higher than a CE stage due to the elimination of Miller effect impact.
- *Feedback Network* multiplies the bandwidth by a factor of $(1+AF)$, but divides the gain by the very same factor. It also improves the linearity of the circuit, and sets the gain to approximately $\frac{1}{F}$.

Noise performance and characteristics of the above have also been detailed and completely derived.

4.2 Transimpedance Architecture

This section reviews the design of two feedback TIAs. It starts from the design of opamp, feedback network and stability, discusses the implementation of an input stage and biasing, and analyzes the noise performance and power supply rejection of the circuits. Finally, the layout of the circuits is discussed as it has great effect on circuit behavior and performance.

4.2.1 Operational Amplifier Design

Design of a high-bandwidth opamp consists of the following considerations:

1. The opamp should be phase inverting since it is subject to connection to a feedback network
2. The gain-bandwidth product should be maximized
3. The input resistance should be as high as possible
4. The output resistance should be as low as possible

Due to the above the opamp incorporates a cascode stage [4][5][6], which inverts the phase and has high gain-bandwidth product. Additionally, a degeneration resistor is applied to simplify the biasing of the stage and increase its input resistance. Nonetheless, a cascode stage exhibits high output resistance. Due to this the cascode if followed by a CC stage [4][6][7][8][9]. The high input resistance of the CC stage doesn't load the cascode output, and the low output resistance fits an opamp design. The topology is depicted in figure 4.10. As evident, the CC stage requires a current sink to drain the emitter current of Q4 into. This connection is provided by the feedback network, as will be explained in the next section. Finally, a capacitor (C_{E2}) is connected in parallel to the degeneration resistor (R_{E2}) for bandwidth enhancement purposes [10].

Biasing is applied to the base of Q3 by means of three serial diodes, which exhibit an almost-constant voltage drop of about 900 mV. This approximation can be represented in small signal terms by a voltage divider, in which the voltage is divided between R_{B3} and the differential resistance of the three diodes. Each diode is implemented by shorting BC–junction of a transistor and using BE–junction as a diode [11].

4.2.2 Feedback and Stability

A shunt-shunt (PIPO) feedback resistor, R_f, is connected between the output and the input of the opamp. As a consequence, the emitter DC current of Q4 is drained to the input node via the feedback resistor. However, should Q4 be loaded by a low resistance load, the opamp output would be divided between the feedback and the load. To avoid

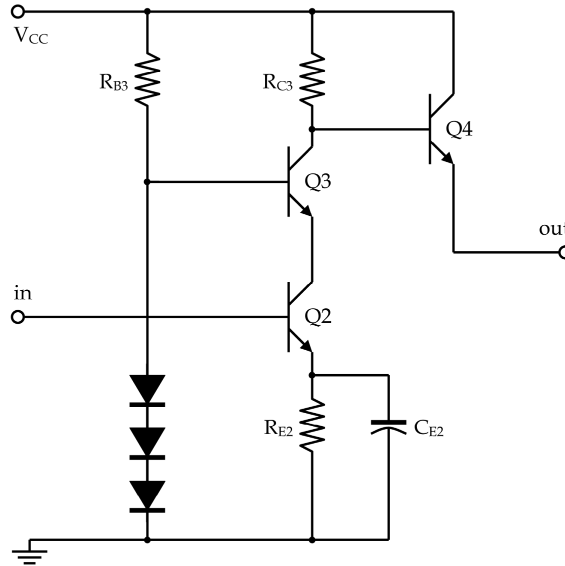


Figure 4.10: Topology of the opamp

this a CC output stage, with high input resistance, is added to the circuit outside the feedback loop [7][9]. The entire closed-loop circuit is shown in figure 4.11.

An important note should be indicated. The virtual ground approximation assumes that input signal of an opamp is significantly smaller than its output signal. As a result, the input node can be referred to as AC ground when looking from the output node. According to this, the emitter of Q4 is connected to a virtual ground by R_f, i.e. the CC stage of Q4 is loaded by R_f, that plays the role of R_E.

Phase margin of slightly less than 60° (namely, 57°) was chosen, since it exhibits flat closed-loop response within the bandwidth and as steep descent as $-40\ \frac{\text{dB}}{\text{dec}}$ above the bandwidth [1]. The reduction to 57° is intended for adding some peaking to the curve to compensate for the photodiode attenuation at high frequencies [12].

Derivation of the open-loop gain (assuming that the voltage gain of the CC stages is approximately unity, and all transistors are identical) yields opamp low frequency transimpedance gain of

$$A = \alpha\beta R_{C3} \approx \beta R_{C3} \tag{4.65}$$

and feedback transmission of

$$F = \frac{1}{R_f} \tag{4.66}$$

thus the loop gain is given by

$$AF \approx \beta \frac{R_{C3}}{R_f} \tag{4.67}$$

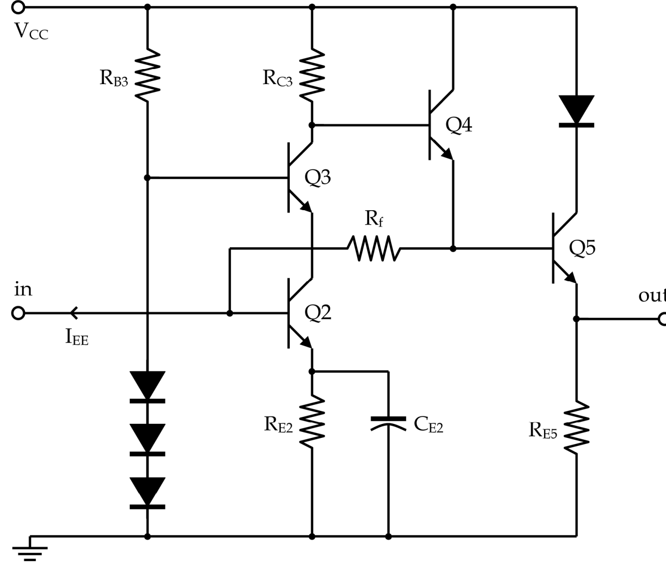


Figure 4.11: Closed-loop TIA topology (*Version B*)

Inserting (4.65) and (4.67) in (4.46) yields closed loop transimpedance gain of

$$G_{CL} \approx R_f \tag{4.68}$$

Anyway, the voltage gain of the opamp is

$$a_v \approx -\frac{R_{C3}}{r_{e2} + r_{ee} + R_{E2}} \tag{4.69}$$

and since the feedback connects in parallel to the input there is no voltage loop, so this is the closed-loop voltage gain as well. As far as frequency response is considered, it is dominated by the cascode as CC stages sport wide-band response [13]. Due to the smaller-than-unity voltage gain of the CE stage in the cascode, $C_{\mu 2}$ can be neglected. Full derivation of the transimpedance gain yields

$$A = \frac{\beta R_{C3}}{\left(1 - \frac{s}{\omega_{p1}}\right)\left(1 - \frac{s}{\omega_{p2}}\right)\left(1 - \frac{s}{\omega_{p3}}\right)} \tag{4.70}$$

where

$$\omega_{p1} = -\frac{1}{C_{\pi 2}[r_{\pi 2} \parallel (\beta + 1)(R_{E2} + r_{ee})]} \tag{4.71}$$

$$\omega_{p2} = -\frac{1}{(C_{\mu 3} + C_{\mu 4})R_{C3}} \tag{4.72}$$

$$\omega_{p3} = -\frac{1}{C_{\pi 3} r_{e3}} \tag{4.73}$$

are the first (dominant), second, and third poles, respectively. Note that the third pole occurs at a very high frequency, and has no effect on stability. Equations (4.70)–(4.72) imply that in order to acquire a specified phase margin, the first pole can be adjusted by R_{E2} and by Q2's emitter current (that changes $r_{\pi 2}$), and the second pole is adjustable through R_{C3}. Changes in R_{C3} also affect the low frequency gain, though. The above also suggests that the higher the β, the lower ω_{p1} is, but ω_{p2} and phase margin are unaffected by β.

4.2.3 Input Stage

Derivation of (4.70)–(4.73) assumed that the input current source is ideal, and so has no capacitance. Should it has capacitance, C_S, a pole would occur at

$$\omega_p \approx -\frac{1}{C_S \frac{R_f}{a_v}} \tag{4.74}$$

If the source is a photodiode its capacitance is relatively high, which may result in bandwidth reduction [13]. To cope with this a current buffer, in the form of CB stage, may be introduced at the input [5][7][9][10][14][15]. The input resistance of a CB stage is very low, to ensure that the pole due to C_S occurs at a high frequency. Moreover, the capacitor seen from the opamp input node is C_μ of the CB stage. This comes up to the circuit illustrated in figure 4.12.

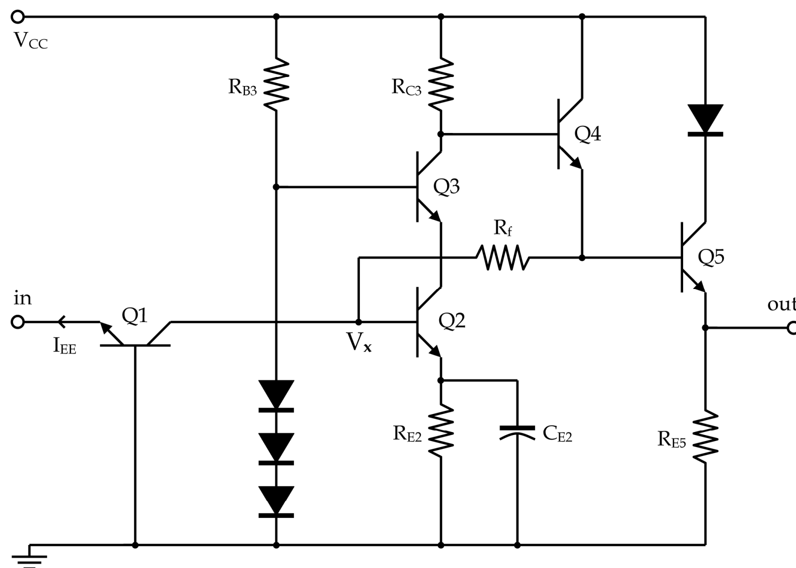


Figure 4.12: TIA with common base input stage (*Version A*)

Henceforth, for the sake of convenience, the circuit topology with CB input stage (shown in figure 4.12) is nicknamed *version A*. The topology without input stage (shown in figure 4.11) is called *version B*. The actual values used in both circuits are detailed in table 4.2.

Element	Value in Version A	Value in Version B
R_f	500 [Ω]	500 [Ω]
R_{C3}	115 [Ω]	150 [Ω]
R_{E2}	10 [Ω]	13 [Ω]
C_{E2}	2 [pF]	3 [pF]
R_{B3}	1.5 [kΩ]	1.5 [kΩ]
R_{E5}	400 [Ω]	400 [Ω]

Table 4.2: Actual values used in TIA designs

4.2.4 Biasing

Examination of the circuits shows that the DC current of Q4 is drained out of the circuit through the input node. Let this biasing current be denoted by I_{EE}. In addition, the supply voltage is denoted by V_{CC}. It is evident that the voltage at the base of Q3 is $3V_{on}$ (where V_{on} is the voltage drop on a diode under forward bias), that is, it is supply independent.

DC current flowing in the cascode branch is easily set in circuit version B, by setting DC voltage at the input node and exploiting the fact that Q2 is degenerated. However, circuit version A lacks this control, and therefore the voltage at the base of Q2 (denoted by V_x in figure 4.12) should be studied.

Inspection of circuit topology of version A reveals that V_x is stabilized by the negative feedback. To calculate its value we assume that all transistors are identical and in the forward active mode, and the output resistance, r_{out}, is infinite. The path along which V_x is calculated is illustrated in figure 4.13. According to the above assumptions one obtains

$$V_x = \frac{\beta R_f + \alpha R_{C3}}{\beta R_G R_f + \alpha R_{C3}(R_G + R_{E2})} \cdot \left[R_{E2} \left(V_{CC} + \alpha R_{C3} \frac{V_{CC} - V_{on}}{\beta R_f + \alpha R_{C3}} - \alpha R_f I_{EE} - 2V_{on} \right) + R_G V_{on} \right] \tag{4.75}$$

where

$$R_G = R_{E2} + \alpha^2 R_{C3} - \frac{\alpha}{\beta} R_f + \frac{\alpha^3 R_{C3}^2}{\beta R_f + \alpha R_{C3}} \tag{4.76}$$

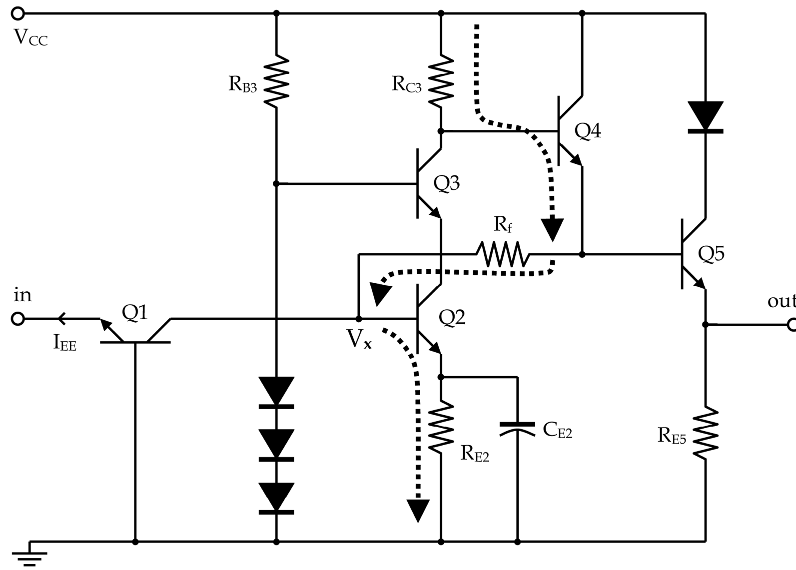


Figure 4.13: TIA version A and the path along which V_x is calculated (dashed arrows)

For large values of β (4.75) reduces to

$$V_x = \frac{R_{E2}}{R_{E2} + R_{C3}} \cdot (V_{CC} - I_{EE} R_f - 2V_{on}) + V_{on} \tag{4.77}$$

Due to the low β of contemporary HBTs, (4.75) has great importance for determining the DC current flowing in the cascode stage and the overall circuit performance. Small currents increase r_e (namely, decrease g_m), and large currents evoke Kirk effect, in which C_μ (C_{BC}) rises. Shown in figure 4.14 are the effects of β, R_{E2}, and R_{C3} on V_x, using $V_{on} = 0.9$ V, $V_{CC} = 5$ V, and $I_{EE} = 2$ mA. Values for R_{E2} and R_f are taken from table 4.2. As evident from the figure, low values of β result in decrease in the cascode DC current, enhancing the complexity of the circuit biasing.

The currents of all stages were designed to set the transistors at their optimal biasing point in terms of bandwidth, Kirk effect, and avalanche breakdown. DC simulation results are summarized in table 4.3 (V_{EE} is the DC voltage at the input node, to which I_{EE} is applied).

4.2.5 Noise Performance

Noise performance of both circuit versions are derived hereinafter. Firstly, noise performance of Q5 is calculated, and then substituted as the load noise term in the noise calculation of Q4. The result is the load noise of the cascode, and so the input-referred noise source of the opamp is obtained. At this level feedback noise is also calculated,

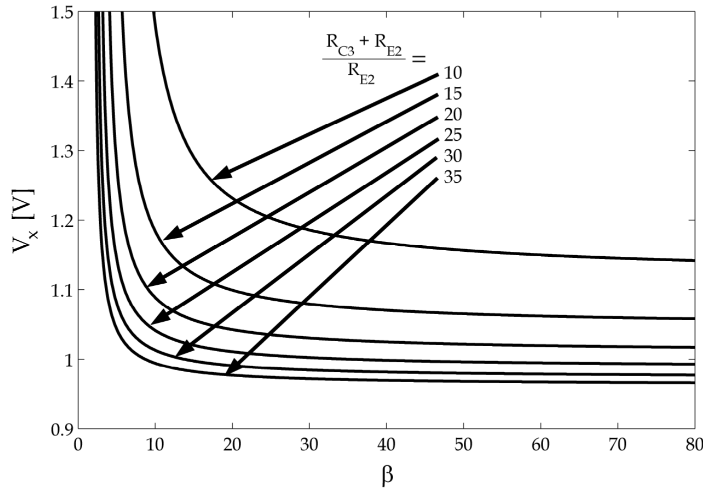


Figure 4.14: V_x versus β and R_{C3}. Values used for calculations: $V_{on} = 0.9$ V, $V_{CC} = 5$ V, $I_{EE} = 2$ mA, $R_{E2} = 10\ \Omega$, and $R_f = 500\ \Omega$

and we are done with version B. As of version A, the result so far is the load noise of Q1, and the input-referred noise source of the entire circuit is obtained. Since the circuits fabricated in this work are connected to 50 Ω measurement pads and systems, a serial

	Version A 8000 Å Collector	Version A 6000 Å Collector	Version B 8000 Å Collector	Version B 6000 Å Collector
V_{CC}	4.0 [V]	3.7 [V]	4.5 [V]	3.2 [V]
I_{EE}	2.5 [mA]	1.0 [mA]	3.5 [mA]	1.5 [mA]
V_{EE}/V_x	–0.893/1.02 [V]	–0.846/1.05 [V]	999 [mV]	951 [mV]
Q1	$V_{CE} = 1.91$ [V] $I_C = 2.42$ [mA]	$V_{CE} = 1.90$ [V] $I_C = 0.97$ [mA]	—	—
Q2	$V_{CE} = 1.51$ [V] $I_C = 6.76$ [mA]	$V_{CE} = 1.42$ [V] $I_C = 10.2$ [mA]	$V_{CE} = 1.57$ [V] $I_C = 5.06$ [mA]	$V_{CE} = 1.49$ [V] $I_C = 3.87$ [mA]
Q3	$V_{CE} = 1.66$ [V] $I_C = 6.54$ [mA]	$V_{CE} = 1.05$ [V] $I_C = 9.83$ [mA]	$V_{CE} = 2.11$ [V] $I_C = 4.90$ [mA]	$V_{CE} = 1.09$ [V] $I_C = 3.74$ [mA]
Q4	$V_{CE} = 1.66$ [V] $I_C = 2.67$ [mA]	$V_{CE} = 1.99$ [V] $I_C = 1.33$ [mA]	$V_{CE} = 1.67$ [V] $I_C = 3.70$ [mA]	$V_{CE} = 1.43$ [V] $I_C = 1.65$ [mA]
Q5	$V_{CE} = 1.66$ [V] $I_C = 3.46$ [mA]	$V_{CE} = 2.00$ [V] $I_C = 2.02$ [mA]	$V_{CE} = 1.67$ [V] $I_C = 4.61$ [mA]	$V_{CE} = 1.44$ [V] $I_C = 2.15$ [mA]

Table 4.3: Simulated biasing points of all TIA transistors

matching resistor of 42 Ω is connected to the circuit outputs. If a circuit would have been embedded in a complete receiver system, the matching resistor might have had been removed. Hence, the noise generated by this resistor is not taken into account in the following calculations, which assume an infinite voltage-sampling load resistance.

The noise referred to the input of Q5 is given by

$$\frac{\overline{v_{n5}^2}}{\Delta f} = 2qI_{C5} \left| Z_{e5} \parallel R_{E5} \right|^2 + 2qI_{B5} \left| \frac{1}{sC_{\mu5}} \parallel Z_{\pi5} \right|^2 + 4KT(r_b + R_{E5}) \tag{4.78}$$

where Z_{e5} and $Z_{\pi5}$ are defined by (4.61) and (4.56), respectively. This noise source is referred to as the load noise source of Q4. In the case of Q4, R_f acts as an emitter resistor, and the load resistance (i.e. the input resistance of Q5) is considered as infinite. Now,

$$\frac{\overline{v_{n4}^2}}{\Delta f} = 2qI_{C4} \left| Z_{e4} \parallel R_f \right|^2 + 2qI_{B4} \left| \frac{1}{sC_{\mu4}} \parallel Z_{\pi4} \right|^2 + 4KT(r_b + R_f) + \frac{\overline{v_{n5}^2}}{\Delta f} \tag{4.79}$$

Here the last term demonstrates the fact that in CC stages the load noise voltage is referred back to the input "as is". Now we come to the cascode. The requested input-referred noise source is current source as the input of a TIA samples current. Hence, we can assume that the source resistance, R_S, is infinite, so that the term $\frac{Z_\pi R_S}{Z_\pi + R_S}$ in (4.62) reduces to Z_π solely. $D_{\pi s}$ models the voltage drop on the source resistance, r_b in this case, and thus can be approximated by unity. Also, r_b is negligible compared to R_S of a current source, and therefore can be omitted from the source's thermal noise term. This yields

$$\frac{\overline{i_{n2}^2}}{\Delta f} = 2qI_{B2} + \frac{1}{G_{m2}^2 \left| Z_{\pi2} \right|^2} \cdot \left[\frac{4KT}{R_{E2}} + 2qI_{C2} + 2qI_{E3} + \frac{4KT}{R_{C3}} + \frac{\overline{v_{n4}^2}}{\Delta f} \frac{1}{R_{C3}^2} \right] \tag{4.80}$$

Adding the noise of R_f gives the total input-referred noise current of version B:

$$\frac{\overline{i_{nB}^2}}{\Delta f} = \frac{\overline{i_{n2}^2}}{\Delta f} + \frac{4KT}{R_f} \tag{4.81}$$

In version A Q1 introduces it own noise current, therefore the circuit's total input-referred noise current is given by

$$\frac{\overline{i_{nA}^2}}{\Delta f} = \frac{\overline{i_{nB}^2}}{\Delta f} + 2qI_{E1} \tag{4.82}$$

According to the above, the input-referred noise current and current spectral density of the circuits are detailed in table 4.4. Differences between wafers evoke from the different biasing required for each circuit to achieve maximum bandwidth, which directly affects DC currents and noise.

The dominant noise sources are those originated by I_{B1}, R_f, and I_{B2}, as they aren't attenuated when referred back to the input. As a consequence, noise exhibited by version A is as twice as high as version B, due to the noise introduced by the common base input stage. Noise current of less than 25 pA/$\sqrt{\text{Hz}}$ is sufficient for a fast optical communication front-end receiver [16]. Noise simulations show a similar picture, as shown in figure 4.15.

	Noise Current Spectral Density [pA/$\sqrt{\text{Hz}}$]	Total Noise Current [μA]	Noise Bandwidth [GHz]
Version A, 8000 Å	30.0	5.5	33.5
Version A, 6000 Å	21.4	3.9	32.8
Version B, 8000 Å	9.2	1.9	40.5
Version B, 6000 Å	8.6	1.6	33.9

Table 4.4: TIA Noise performance

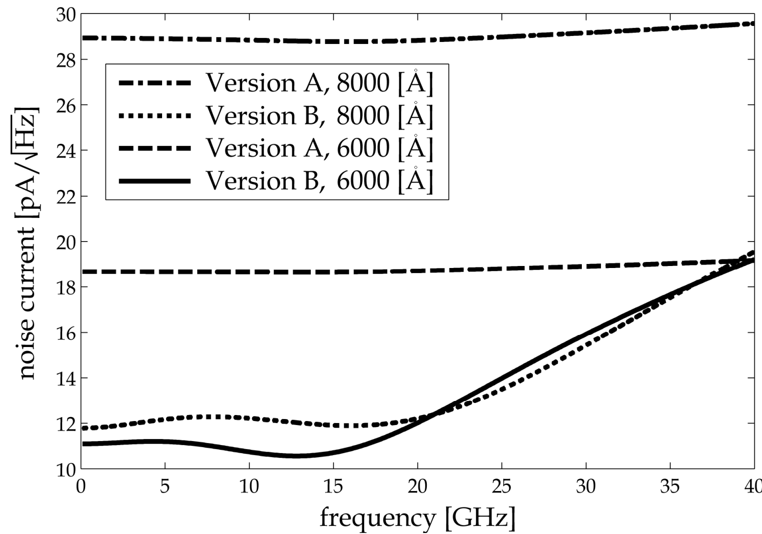


Figure 4.15: Simulated noise current

4.2.6 Power Supply Rejection

A receiver on chip includes the light detector, TIA, automatic gain control, clock and data recovery, decision circuit, and other digital circuits [14]. The TIA on the receiver chip will thus be vulnerable to power supply noise generated by all other circuits on the chip. Since optoelectronic TIAs are most commonly single-ended amplifiers [4][14] the power supply noise is almost directly transferred to their output. Thus, the power supply rejection ratio (PSRR) is an important consideration in circuit design of high speed TIAs. In addition, techniques for enhancing PSRR, such as folded cascode topology, are infeasible in the bipolar NPN technology in which use was made for this work.

In the below discussion PSRR of version A is derived firstly, and then version B is

referred. As mentioned above (section 4.2.4), the circuit is biased by a voltage source, V_{CC}, and current source, I_{EE}. Variations of V_{CC} are transferred to the output via three paths:

A. Variations of V_x, which are transferred to the output.

B. Variations of the voltage at the base of Q3.

C. The voltage at the output of the cascode stage follows V_{CC}.

Denoting the supply voltage noise by v_{nCC}, the amplified power supply noise at the output is given by:

$$v_{oCC} = v_{oA} + v_{oB} + v_{oC} \tag{4.83}$$

where v_{oA}, v_{oB}, and v_{oC} are the output noise due to paths A, B, and C, respectively. Now,

$$v_{oA} = -v_{nCC}\frac{\partial V_x}{\partial V_{CC}} \cdot \frac{R_{C3}}{r_{e2} + R_{E2}} \tag{4.84}$$

$$v_{oB} = -v_{nCC}\frac{R_D}{R_D + R_{B3}} \cdot \frac{R_{C3}}{r_{e3} + r_{out2}} \approx 0 \tag{4.85}$$

$$v_{oC} = v_{nCC} \tag{4.86}$$

where R_D is the differential resistance of the three-diode branch at the base of Q3. We have assumed that the closed-loop transimpedance gain is R_f, and that the gain of the emitter followers is unity. Inserting (4.77) in (4.84) yields that for $\beta \gg 1$

$$v_{oA} = -v_{nCC}\frac{R_{E2}}{R_{E2} + R_{C3}} \cdot \frac{R_{C3}}{r_{e2} + R_{E2}} \tag{4.87}$$

obtaining the supply gain

$$A_{CC} = \frac{v_{oCC}}{v_{nCC}} = -\frac{R_{E2}}{R_{E2} + R_{C3}} \cdot \frac{R_{C3}}{r_{e2} + R_{E2}} + 1 \tag{4.88}$$

Note that A_{CC} consists of negative and positive terms that may cancel each other if designed properly. Whilst paths A and B include phase-inverting stages, path C keeps the phase unchanged. The V_{CC}–PSRR is:

$$PSRR(V_{CC}) = \left|\frac{Z_T}{A_{CC}}\right| = \left|\frac{R_f}{\frac{R_{E2}}{R_{E2}+R_{C3}} \cdot \frac{R_{C3}}{r_{e2}+R_{E2}} - 1}\right| \tag{4.89}$$

The second supply, I_{EE}, is implemented by a voltage source, V_{EE}, in series with a resistor. The resistor is connected in parallel to the photodiode. Denoting the biasing resistor by R_{EE}, the PSRR of V_{EE} is given by:

$$PSRR(V_{EE}) = \left|\frac{Z_T}{A_{EE}}\right| = R_{EE} \tag{4.90}$$

Equation (4.90) implies that the PSRR of V_{EE} is easily controlled by determining the value of R_{EE}. Since R_{EE} is from the order of 1 kΩ, the supply noise at V_{EE} consumes very small portion of the circuit's noise budget. Hence, only the PSRR of V_{CC} is discussed below.

In order to enhance the PSRR given in (4.89) the first term in the denominator should be close to unity. Namely, the resistors R_{E2} and R_{C3} should be set to minimize the denominator of (4.89). This task is not straight forward because these resistors also set the open loop gain and poles. In addition, the self-cancellation term depends on the frequency response of the cascode stage, which results in a reduction in PSRR as frequency increases.

The PSRR of circuit version B, without the CB input stage, is given by

$$PSRR = \left| \frac{R_f}{\frac{R_D}{R_D+R_{B3}} \cdot \frac{R_{C3}}{r_{e3}+r_{out2}} - 1} \right| \tag{4.91}$$

Since the first term in the denominator is much smaller than unity, no self-cancellation effect takes place in (4.91), and the expected PSRR of version B is much lower than version A at low frequencies. However, this term increases with frequency thanks to $C_{\mu 2}$, that shorts Q3's emitter to the AC ground introduced at the input node. Simulated PSRRs of version A and B are shown in figure 4.16. Calculation of the PSRR according to (4.89) and (4.91), with the parasitic resistances of the transistors lumped into R_{E2} and R_{C3}, yields a low frequency rejection ratio of 65 dBΩ for version A, and 54 dBΩ for version B. These results are in accordance with the simulated PSRRs. As evident in figure 4.16, version A exhibits significantly better PSRR than version B at low frequencies. At higher frequencies version B improves and version A deteriorates.

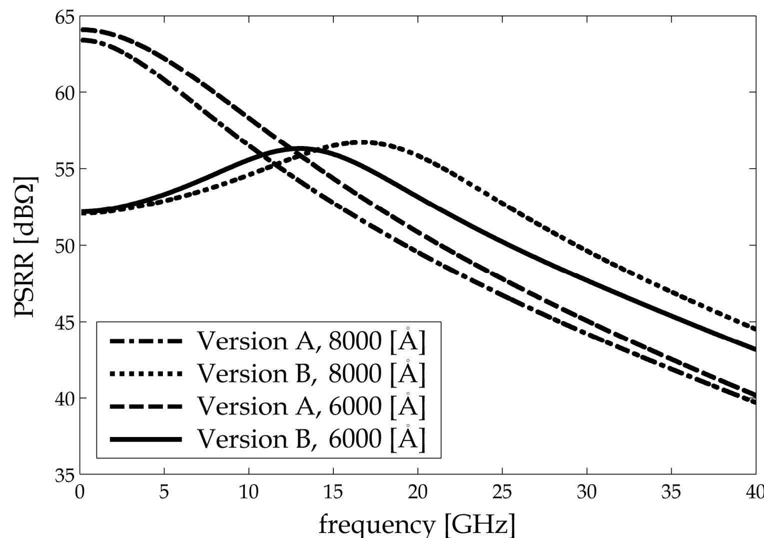


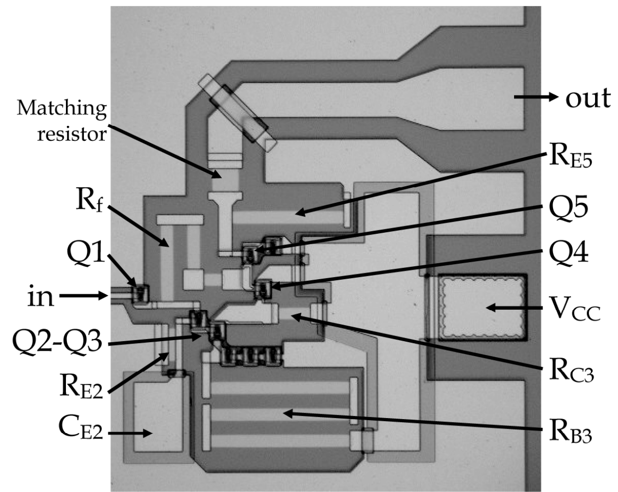
Figure 4.16: Simulated PSRR

4.2.7 Layout

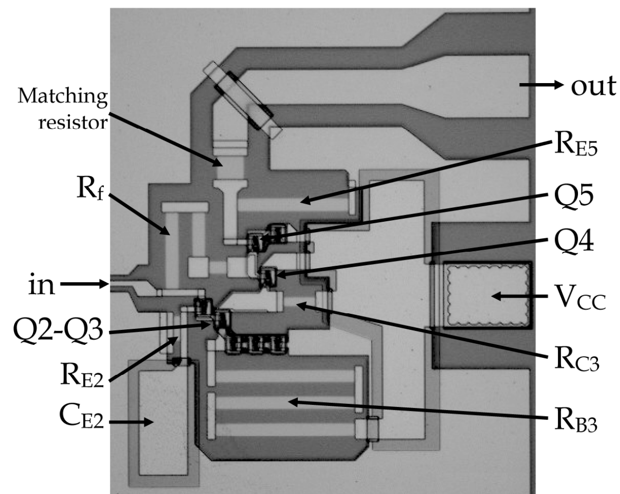
Design of layout for high frequencies, such as millimeter-waves, requires extra care to parasitics. Hence, layout of the circuits was designed according to the following guidelines:

- Elements should be as close to each other as possible to keep the circuit lumped.
- Metal inter-connect lines should be as thick as possible to reduce their inductance. This can be implemented both by widening the lines and by heaping metal 2 on metal 1.
- Resistors should be kept away from any AC ground to eliminate parasitic capacitances.
- Resistors should be as wide as possible for two reasons: (1) to reduce the effect of lithography imprecision, and (2) to prevent self heating of the thin NiCr film.
- Voltage supplies must be stabilized by means of capacitance to ground.

The layouts resulted from the guidelines mentioned above and circuit topologies are depicted in figure 6.1. Total die size is 826x808 μm^2. As can be seen, circuits are intended for measurement in a 50 Ω system, that is, they comprise matching resistors and 50 Ω output transmission lines. V_{CC} can be applied to its corresponding pad by either using a DC needle probe or a dual RF probe with G-S-G-S-G configuration. Finally, all pads are designed for 150 μm pitch G-S-G probes.



(a)



(b)

Figure 4.17: Layout of TIAs: (a) version A (b) version B. Total die size is 826x808 μm^2

4.3 Optoelectronic Integration

Integration of a photodiode with a TIA is the last stage in photoreceiver's front-end design. The goal of such an integration is to achieve maximum responsivity, gain, and bandwidth measured from the optical input down to the amplified electrical output, as well as minimum noise. For this sake PD's dimensions, responsivity, bandwidth, and noise performance, together with TIA's gain, bandwidth, and noise performance, should be considered.

4.3.1 Photodiode Dimensions and Optoelectronic Bandwidth

The diameter of a light beam coming from an optical fiber with lens-shaped edge is 5 μm. However, positioning of the fiber above a PD is somewhat difficult, especially when a movable optical probe is used. Increasing the diameter of the PD can help with this, but this increases the capacitance of the diode (see table 3.5). In addition, the fact that the lighting generation occurs in the center of the diode, far away from the p contact, results in high contact resistance, because current has to travel a long way through the p layer to the contact.

A photodiode introduces two poles to the optoelectronic frequency response — one emerges from the capacitance of the diode, C_j, and the resistance by which the PD is loaded, namely, $R_{contact} + Z_{in}$ (= PD's contact resistance + TIA's input impedance); another pole is originated by the PD's transit time, τ_{PD}. While the latter has nothing with the diameter of the PD, the former is a direct outcome of the dimensions.

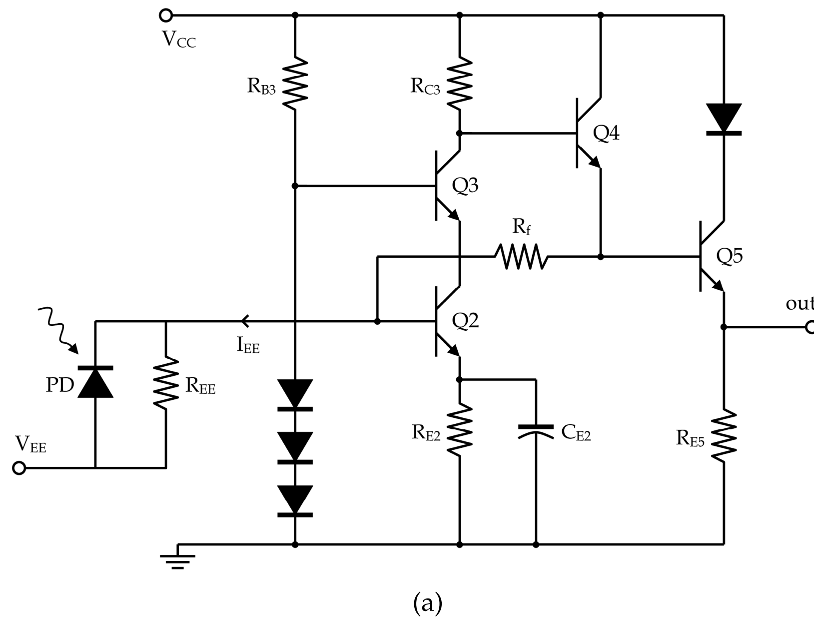
In the layout each of the circuit versions has been joined with 10, 15, and 20 μm PDs. Evidently, the smallest PD is expected to achieve the best bandwidth.

4.3.2 Layout Integration

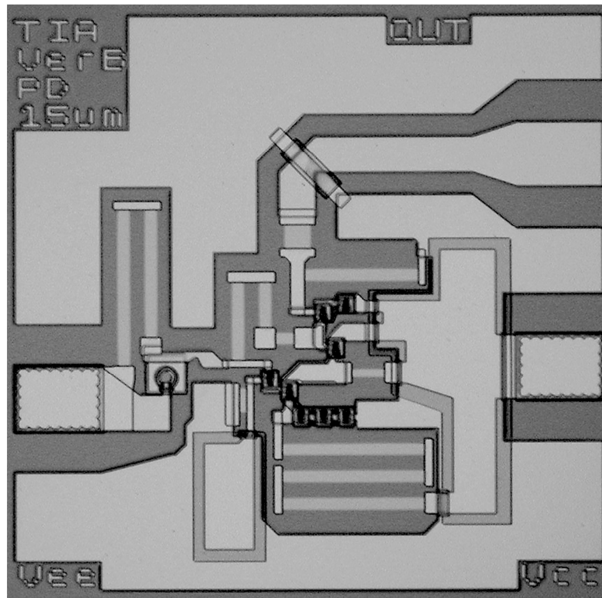
Photodiodes connect directly to TIAs' input nodes. Due to biasing considerations, n contact of the PD is connected to the TIA's input node, and p contact is connected to a biasing pad, that acts as an AC ground. As mentioned before, biasing is applied by means of a resistor connected in parallel to the PD, as shown in figure 4.18a. Applying V_{EE} sets both the reverse bias of the PD as well as I_{EE}, where R_{EE} converts the voltage source of V_{EE} to current source. An optoelectronic circuit is shown in figure 4.18b. The pad of V_{EE} is designed for a 150 μm pitch G-S-G probe. The value of R_{EE} is 1 kΩ for all circuit and PD versions.

4.3.3 Optoelectronic Noise Performance

The noise equivalent power of a monolithic circuit can be calculated from (3.19) and is given by



(a)



(b)

Figure 4.18: TIA version B with 15 μm photodiode and biasing resistor: (a) scheme (b) layout

$$NEP = \frac{hc}{\lambda} \cdot \frac{1}{q\eta} \sqrt{2qI_n\Delta f + \frac{4KT}{R_{EE}}\Delta f + \overline{i^2_{nTIA}}} \quad (4.92)$$

where I_n is the DC current flowing in the PD, and $\overline{i^2_{nTIA}}$ is the input-referred noise current of the TIA. Results of NEP calculations of each circuit are given in table 4.5, based on simulations and $I_n = 5\ \mu$A.

Photodiode Diameter [μm]	Collector Width [Å]	Bandwidth with version A [GHz]	Bandwidth with version B [GHz]	NEP with version A [μW]	NEP with version B [μW]
10	8000	19.2	19.6	14.9	7.0
15	8000	19.0	18.8	14.8	6.9
20	8000	18.8	17.9	14.7	6.7
10	6000	18.1	16.7	18.2	11.5
15	6000	17.6	15.8	18.0	11.1
20	6000	17.2	15.0	17.0	10.9

Table 4.5: NEP of optoelectronic circuits

4.4 Circuit Measurements and Characterization

Both electrical and optoelectronic measurements are reviewed in this section. The electrical measurements are used for characterization of the TIA in terms of transimpedance gain. The optoelectronic measurement provides information on the bandwidth of the entire monolithic front-end photoreceiver. Measurement procedures and configurations are detailed hereinafter.

4.4.1 TIA Measurements and Characterization

S–parameter measurements were carried out on each circuit to find its optimal biasing and characterize its electrical performance.

Measurement Equipment and Setup

Setup of the S–parameter measurement is quite similar to the HBT setup described in section 2.3.4. I_{EE} is applied to port 1 of the network analyzer through the internal bias tee of the analyzer, and port 2 is left unbiased (floating). V_{CC} is applied to is corresponding pad either by means of a DC needle or by a dual RF probe.

Transimpedance Measurement

S–parameters measured on the TIA under test are converted to transimpedance gain using the following equity [17]:

$$Z_T = \left| Z_0 \cdot \frac{S_{21}}{1 - S_{11}} \right| \tag{4.93}$$

where Z_0 is the characteristic impedance of the S–parameter measurement system, 50 Ω in most cases. Presenting the data in dBΩ format (i.e. $20 \log_{10} |Z_T|$) enables extraction of the –3dB frequency of the TIA.

4.4.2 OEIC Measurements and Characterization

Measurement of optoelectronic integrated circuits (*OEICs*) includes small signal frequency response characterization solely, described below.

Measurement Equipment and Setup

Measurement equipment and setup is similar to section 3.3.1, except that the output of the TIA is DC-floating, eliminating the need for bias tee in the output probe. However, V_{CC} is applied by another RF probe (with a bias tee, terminated by 50 Ω in its RF node), and a DC needle is used for applying V_{EE}. An optical probe is used for the optical input signal. Shown in figure 4.19 is the probing setup mentioned above.

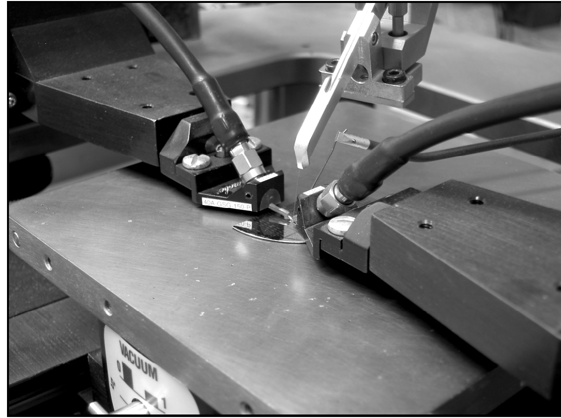


Figure 4.19: Probing of a circuit

Frequency Response Measurement

This measurement is similar to PD frequency response measurement (refer to section 3.3.1). Data acquisition and processing is also alike. However, the curve here includes more than one complex pole, all occur at the same frequency, and so more complicated fitting is applied. A curve fitted to measured data is depicted in figure 4.20.

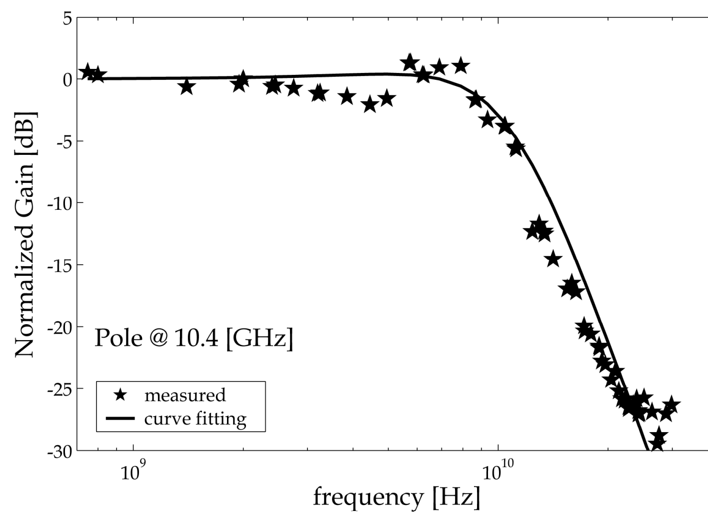


Figure 4.20: A curve fitted to OEIC measured data. Measured on circuit version B with 10 μm photodiode

4.5 Circuit Performance

This section summarizes the performance of all lumped circuits of this work, as well as compares measured results to simulated ones. Both electrical and optoelectronic results are discussed.

4.5.1 TIA Performance

All data are summarized in table 4.6 and depicted in figure 4.21. It should be noted that Z_T includes the effect of a matching resistor and 50 Ω load at the output.

	Version A, 8000 Å Collector	Version B, 8000 Å Collector	Version A, 6000 Å Collector	Version B, 6000 Å Collector
V_{CC}	4.0 [V]	4.5 [V]	3.7 [V]	3.2 [V]
I_{EE}	2.5 [mA]	3.5 [mA]	1.0 [mA]	1.5 [mA]
DC Power Consumption	58 [mW]	71 [mW]	57 [mW]	30 [mW]
Z_T — simulated	44.3 [dBΩ]	45.2 [dBΩ]	44.1 [dBΩ]	44.6 [dBΩ]
Z_T — measured	39.7 [dBΩ]	45.5 [dBΩ]	43.5 [dBΩ]	44.2 [dBΩ]
f_{-3dB} — simulated	21.3 [GHz]	25.8 [GHz]	20.9 [GHz]	21.6 [GHz]
f_{-3dB} — measured	22.0 [GHz]	21.6 [GHz]	21.5 [GHz]	20.8 [GHz]

Table 4.6: Simulated and measured performance of TIAs

The results imply that best performances, in terms of flatness of response, are achieved by the 8000 Å collector wafer. This is due to the fact that about 2000 Å (out of 8000) of the collector isn't depleted. In the 6000 Å collector wafer only 4500 Å is depleted. As a result, the latter suffers from small phase margin and consequently excessive peaking in the closed-loop response.

4.5.2 OEIC Performance

Results of the optoelectronic measurements are summarized in table 4.7. As expected, widest bandwidth is achieved with 10 μm photodiodes (section 4.3.1). The measured bandwidths are considerably smaller than the simulated ones. This is due to the non-depleted regions in the collector layer, in which generated charge carriers move by diffusion. Since diffusion is significantly slower than drifting in a depletion region, τ_{PD} is increased, and bandwidth is reduced.

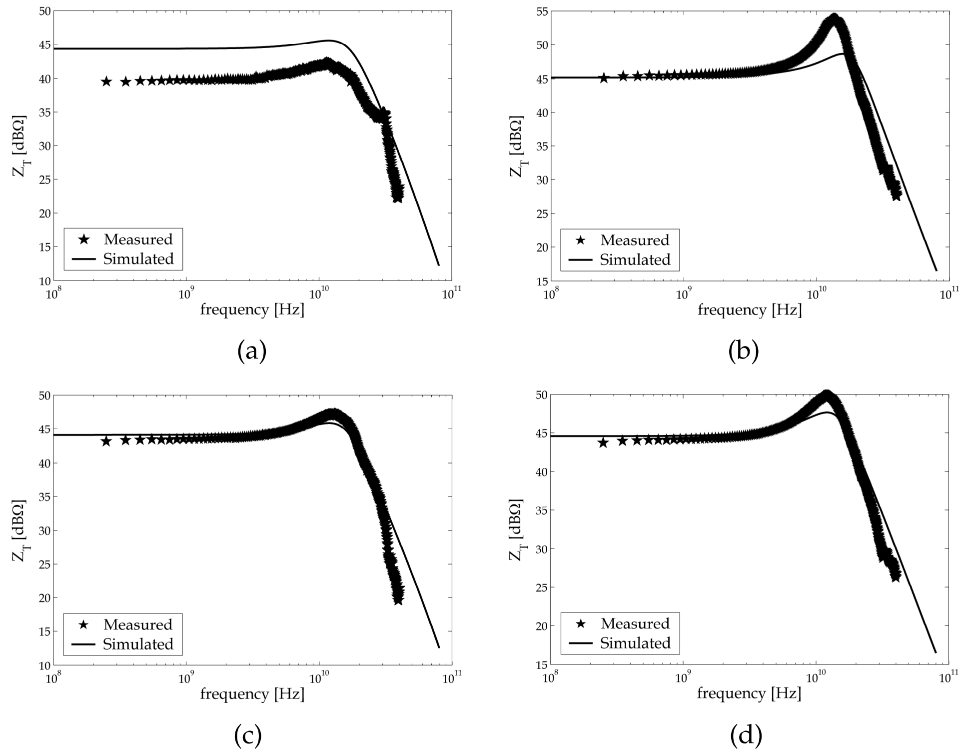


Figure 4.21: Simulated and measured performance of various TIAs: (a) version A, 8000 Å collector (b) version B, 8000 Å collector (c) version A, 6000 Å collector (d) version B, 6000 Å collector

	Version A, 8000 Å Collector	Version B, 8000 Å Collector	Version A, 6000 Å Collector	Version B, 6000 Å Collector
10 μm PD	14.2 (19.2) [GHz]	11.2 (19.6) [GHz]	11.2 (18.1) [GHz]	10.4 (16.7) [GHz]
15 μm PD	13.0 (19.0) [GHz]	10.4 (18.8) [GHz]	7.9 (17.6) [GHz]	8.7 (15.8) [GHz]
20 μm PD	8.7 (18.8) [GHz]	9.3 (17.9) [GHz]	6.2 (17.2) [GHz]	10.4 (15.0) [GHz]

Table 4.7: Optoelectronic bandwidth of TIAs with various photodiodes — measured (simulated)

The optoelectronic gain is given by

$$G_{oe} = \Re \cdot Z_T \tag{4.94}$$

Substituting measured results in (4.94) and omitting the effect of the output matching

resistor yields gain of 61.1 and 119.2 V/W for circuits A and B with 8000 Å collector, while the 6000 Å wafer exhibits 49.6 and 53.7 V/W for circuits A and B, respectively. The 8000 Å diodes have higher responsivity, and thus their corresponding amplifiers have higher optoelectronic gain.

References for Chapter 4

[1] Paul R. Gray and Robert G. Meyer, *Analysis and Design of Analog Integrated Circuits*, John Wiley & Sons, third edition, 1993.

[2] David A. Johns and Ken Martin, *Analog Integrated Circuit Design*, John Wiley & Sons, 1997.

[3] Behzad Razavi, *Design of Analog CMOS Integrated Circuits*, McGraw-Hill, 2001.

[4] M. Govindarajan and S. R. Forrest, "Design considerations for wide-band p-i-n/HBT monolithic transimpedance optical receivers", *IEEE Journal of Lightwave Technology*, vol. 2, pp. 367–378, February 1993.

[5] Zhongyuan Chang and Willy M. C. Sansen, "Stability and noise performance of constant transimpedance amplifier with inductive source", *IEEE Transactions on Circuits and Systems*, vol. 36, pp. 264–271, February 1989.

[6] Eugene John and Mukunda B. Das, "Design and performance analysis of InP-based high-speed and high-sensitivity optoelectronic integrated receivers", *IEEE Transactions on Electron Devices*, vol. 41, pp. 162–172, February 1994.

[7] Tongtod Vanisri and Chris Toumazou, "Integrated high frequency low-noise current-mode optical transimpedance preamplifiers: Theory and practice", *IEEE Journal of Solid-State Circuits*, vol. 30, pp. 677–685, June 1995.

[8] Ye Lu and Mourad N. El-Gamal, "A 2.3V low noise, low power, 10 GHz bandwidth Si-bipoar transimpedance preamplifier for optical receiver front-ends", *The 2001 IEEE International Symposium on Circuits and Systems*, vol. 4, pp. 834–837, May 2001.

[9] D. Huber, M. Bitter, M. Dülk, S. Fischer, E. Gini, A. Neiger, R. Schreieck, C. Bergamaschi, and H. Jäckel, "A 53 GHz monolithically integrated InP/InGaAs PIN/HBT receiver OEIC with an electrical bandwidth of 63 GHz", *Proceedings of the 12th Conference on Indium Phosphide and Related Materials*, pp. 325–328, May 2000.

[10] B. Wilson and I. Darwazeh, "Low input resistance transimpedance optical preamplifier for fiber optic local area networks", *IEEE International Symposium on Circuits and Systems*, vol. 3, pp. 2531–2534, June 1988.

[11] Paolo Antognetti and Giuseppe Massobrio, *Semiconductor Device Modeling with SPICE*, McGraw–Hill, 1988.

[12] J. Bellon and M. J. N. Sibley, "Frequency response compensation of transit time limited pin photodiode", *Electronics Letters*, vol. 36, no. 14, pp. 1222–1223, July 2000.

[13] Robert G. Meyer and Robert Alan Blauschild, "A wide-band low-noise monolithic transimpedance amplifier", *Journal of Solid-State Circuits*, vol. SC-21, pp. 530–533, August 1986.

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

[15] Benny Sheinman, *Modeling and Performance of Ultrafast InP Based Heterojunction Bipolar Transistors and Circuits*, PhD thesis, Technion – IIT, October 2004.

[16] Kevin W. Kobayashi, "State-of-the-art 60 GHz, 3.6 K-Ohm transimpedance amplifier for 40 Gb/S and beyond", *IEEE RFIC Symposium*, pp. 55–58, June 2003.

[17] Andreas Leven, Ralf Reuter, and Yves Baeyens, "Unified analytical expressions for transimpedance and equivalent input noise current of optical receivers", *IEEE Transactions on Microwave Theory and Techniques*, vol. 48, no. 10, pp. 1701–1706, October 2000.

Chapter 5

Optoelectronic Distributed Circuit

The distributed amplifier, also referred to as traveling wave amplifier (*TWA*), is an efficient topology that gets the most from a transistor. This chapter deals with the design of a TWA that acts as a preamplifier for optical communications. The design is based on a previous work of Emanuel Cohen [1] that proved the feasibility of broadband TWA using the technology of the Technion. Cohen's amplifier was designed as a power amplifier, and thus redesign was required to reduce the input impedance. The first topic is the fundamentals of distributed amplifiers; a discussion on the architecture chosen for this work is the following subject. The third section deals with the optoelectronic integration, that is, the considerations in connecting a photodiode to a TWA. Finally, measurement, characterization, and performance of the circuit are discussed.

5.1 Distributed Amplifier Fundamentals

Distributed amplifiers consist of two transmission lines and one or more amplification stages. The basic elements and principles of a TWA are briefly discussed below.

5.1.1 Transmission Lines

Transmission line is an electrical medium used to transmit power, in the form of electromagnetic wave, from one point to another. Its characteristics and behavior are summarized in this section from an electrical viewpoint.

Basic Concepts

Shown in figure 5.1 is the typical connection of a transmission line. A source with output resistance, R_s, or impedance, Z_s, introduces an AC signal at the input node of the line. On the output side the line is loaded by an impedance denoted by Z_L, and the characteristic impedance of the line itself is Z_0. The voltage reflection coefficient at the input is given by [2][3]

$$\Gamma_{in} = \frac{Z_0 - Z_s}{Z_0 + Z_s} \tag{5.1}$$

and at the output it is

$$\Gamma_L = \frac{Z_L - Z_0}{Z_L + Z_0} \tag{5.2}$$

S–parameters are related to reflection coefficients in the following way: S_{11} is actually Γ_{in} when $\Gamma_L = 0$, that is, when $Z_L = Z_0$. Similarly, S_{22} equals Γ_L when $\Gamma_{in} = 0$ ($Z_s = Z_0$).

The electrical field of the wave traveling along the line is of the form [3]

$$\overrightarrow{E} = E_0 e^{-\gamma x} \hat{x} = E_0 e^{-\alpha x} e^{-j\beta x} \hat{x} \tag{5.3}$$

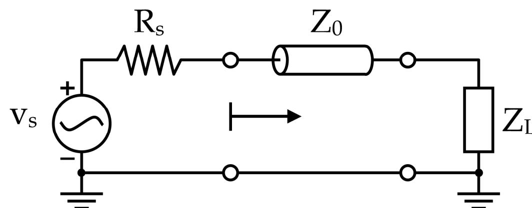


Figure 5.1: Transmission line electrical connection. The arrow indicates the direction of power propagation

where E_0 is the amplitude, x is position along the line, and γ is the complex propagation constant. γ can be expressed as a function of propagation constant, β, and attenuation constant, α, as follows:

$$\gamma = \alpha + j\beta \tag{5.4}$$

For a lossless line $\alpha = 0$, which means that the signal isn't attenuated when propagating along the line. Another quantity related to electrical field is the electrical length of a line, defined as

$$E = \beta l \tag{5.5}$$

where l is the length of the transmission line. Lastly, the group velocity of the wave is given by [3]

$$v_g = \left(\frac{\partial\gamma}{\partial\omega}\right)^{-1} \tag{5.6}$$

and the phase velocity by

$$v_p = \frac{\omega}{\gamma} \tag{5.7}$$

Transmission Line Model

Transmission lines can be modeled in many cases by infinitesimal lumped elements such as inductors, capacitors, and resistors. A lossless line can be modeled as a chain or L–C links (figure 5.2a), whilst the model of a lossy line includes series and parallel resistances (figure 5.2b) [2][3]. Note that in lossless line $R \longrightarrow 0$ and $G \longrightarrow 0$. All of the parameters in the model (L, C, R, and G) are per unit length. The quantities of a transmission line can be expressed by its model parameters, making the design of a transmission line more handy.

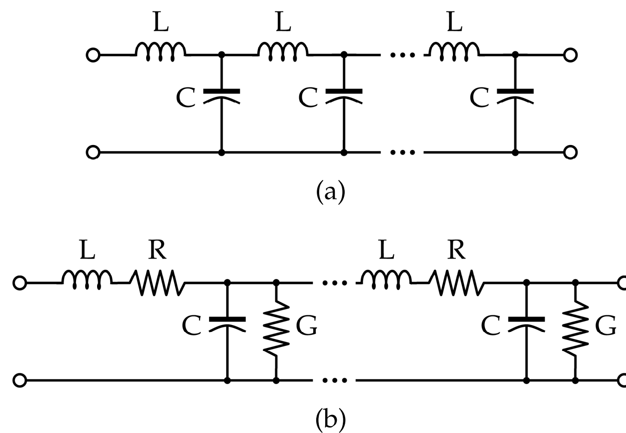


Figure 5.2: Models of (a) lossless and (b) lossy transmission lines

The characteristic impedance of a lossy transmission line can be expressed by its model parameters as

$$Z_0 = \sqrt{\frac{R + j\omega L}{G + j\omega C}} \tag{5.8}$$

while for lossless line (5.8) reduces to

$$Z_0 = \sqrt{\frac{L}{C}} \tag{5.9}$$

Table 5.1 summarizes the different line quantities as a function of the model parameters. The approximations for lossy lines are valid if $RG \ll LC, RC, LG$ which is the case

Quantity	Description	Units (M.K.S)	Value in lossless line	Value in lossy line
Z_0	Line characteristic impedance	$[\Omega]$	$Z_0 = \sqrt{\frac{L}{C}}$	$Z_0 = \sqrt{\frac{R+j\omega L}{G+j\omega C}}$
α	Attenuation constant	$\left[\frac{rad}{m}\right]$	$\alpha = 0$	$\alpha \approx \frac{R}{2Z_0} + \frac{GZ_0}{2}$
β	Propagation constant	$\left[\frac{rad}{m}\right]$	$\beta = \omega\sqrt{LC}$	$\beta \approx \omega\sqrt{LC}$
v_g	Group velocity	$\left[\frac{m}{sec}\right]$	$v_g = \frac{1}{\sqrt{LC}}$	$v_g \approx \frac{1}{\sqrt{LC}}$
v_p	Phase velocity	$\left[\frac{m}{sec}\right]$	$v_p = \frac{1}{\sqrt{LC}}$	$v_p \approx \frac{\omega}{\frac{R}{2Z_0}+\frac{GZ_0}{2}+j\omega\sqrt{LC}}$

Table 5.1: Transmission line quantities expressed by its model parameters

in most transmission lines. In addition, group velocity is assumed independent of frequency; this is perfectly true only if $RC = GL$ [3].

Manipulating the equations of table 5.1 reveals that for a lossless transmission line L and C are given by [2]

$$L = Z_0 \cdot \frac{l}{v_p} \tag{5.10}$$

and

$$C = \frac{1}{Z_0} \cdot \frac{l}{v_p} \tag{5.11}$$

Equations (5.10) and (5.11) are useful in practical circuit design since electromagnetic simulation tools prevalently calculate Z_0 and v_p, not L and C.

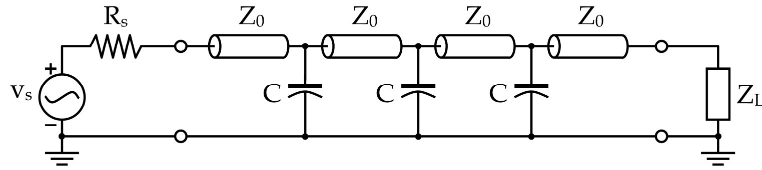


Figure 5.3: Transmission line with periodical capacitive discontinuities

Bragg Frequency
Uniform transmission lines are very rare in practice. Actual lines commonly suffer from discontinuities such as ground connection bridges, junctions, and parasitic capacitances. Many times the distracting capacitances are periodically positioned, and the distance between two adjacent capacitances is much smaller than the wavelength (as illustrated in figure 5.3). In this case the additional capacitances can be referred to as a distributed capacitance (per unit length) and be added to the capacitance of the transmission line, C. However, due to the periodical nature of the discontinuities the line has limited bandwidth caused by Bragg scattering. The bandwidth, so-called "Bragg frequency", is given by [4]

$$f_{Bragg} = \frac{1}{\pi \cdot l\sqrt{LC}} \tag{5.12}$$

where l is the total length of the line. As a rule of thumb, when designing a circuit that comprises transmission lines, f_{Bragg} should be at least 2 to 3 times the circuit's bandwidth [1].

Coplanar Waveguide

There are several types of transmission lines used in integrated circuits, each has its own advantages and drawbacks. As the InP process in which the circuits are fabricated doesn't offer more than one metal level (see figure 2.18), the coplanar waveguide (*CPW*) is a natural choice for transmission lines [1]. The structure of CPW, shown in figure 5.4, consists of a signal line sandwiched between two ground lines — all rest on the same planar substrate (the semi-insulating InP wafer, in our case). The parameters that determine the characteristics of the line are the thickness of the metal, t, the width of the signal line, W, the distance between signal and ground lines, G, the length of the line, l, the conductivity of the metal, σ, the dielectric coefficient of the substrate, ϵ_r, and the dielectric loss factor of the substrate, $\tan\delta$ [4]. The values taken for metal 1 CPWs, made of gold metal lines over InP substrate, are listed in table 5.2. Using these values with reasonable values of W and G yield Z_0 of 20 to 60 Ω, as illustrated in figure 5.5.

Generally speaking, the greater G, the smaller C, and thus the higher Z_0. Similarly, the greater W, the smaller L, and thus the lower Z_0. However, the higher the values of C

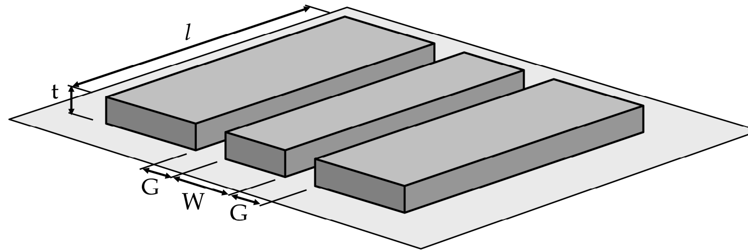


Figure 5.4: Structure of coplanar waveguide

Parameter	Description	Value
ϵ_r	Dielectric coefficient of InP	12.4
$\tan\delta$	Dielectric loss factor of InP	$5 \cdot 10^{-4}$
σ	Conductivity of Au	$48.8 \cdot 10^6 \frac{1}{\Omega\text{m}}$
t	metal thickness	0.8 μm

Table 5.2: Values used for calculation of CPW's characteristics

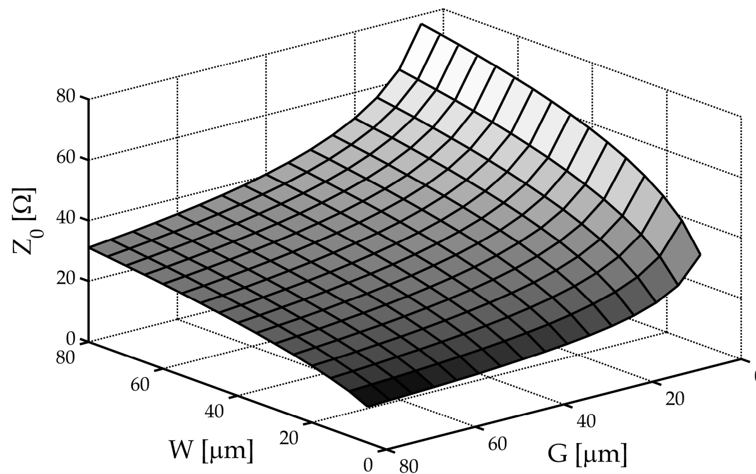


Figure 5.5: Z_0 as a function of W and G

and L, the slower the group and phase velocities (v_g and v_p, respectively).

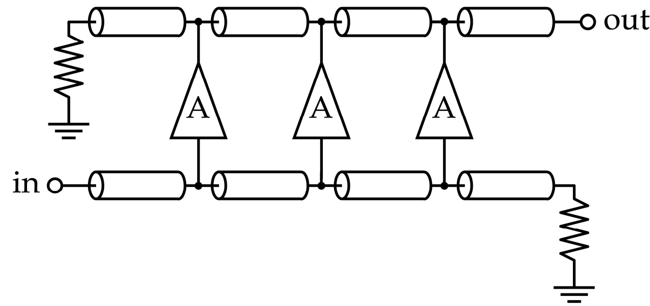


Figure 5.6: Structure of a typical TWA

5.1.2 Distributed Amplifier

Distributed amplifiers (TWAs) consist of input transmission line, amplification stage or stages, and output transmission line. Shown in figure 5.6 is the structure of a typical TWA. The amplifier works as follows. A signal is introduced at the input node and propagating along the input transmission line (the lower line in figure 5.6). Every time the signal passes a junction a portion of it enters the amplification stage connected to the junction (marked by 'A' in the figure). The amplified signals at the outputs of the amplification stages are then collected by the output (upper) transmission line, which forms an amplified signal propagating towards the output node. The details concerned with how a TWA works are discussed hereinafter.

Transistor capacitances as Matching Networks

A matching network is a network the matches the impedance of a load to the impedance of the source, or vice versa. Lumped elements such as resistors, capacitors, and inductors, can be used as matching networks [2][3]. The idea behind TWA is to use the input capacitance of the amplification stage as a matching network between the input line and the stage, and to use the output capacitance as a matching to the output line [5]. As a result, these capacitances share their energy with the inductance and capacitance of the transmission lines, which in turn enhances the bandwidth of the stage.

The input impedance when the input line is loaded by amplification stages is given by [6]

$$Z_{0in} = \sqrt{\frac{R_{in} + j\omega L_{in}}{G_{in} + j\omega C_{in} + \frac{Y_{st,in}}{l_{in}}}} \tag{5.13}$$

where R_{in}, L_{in}, C_{in}, and G_{in} are the parameters of the transmission line, and $Y_{st,in}$ is the input admittance of the amplification stage. Here l_{in} is the distance between two adjacent

stages. In a HBT-based TWA (contrary to the FET case of [6]) $Y_{st,in}$ can be expressed as

$$Y_{st,in} = j\omega C_{st,in} + \frac{1}{R_{st,in}} \tag{5.14}$$

The complex propagation constant is thus given by

$$\gamma_{in} = \alpha_{in} + j\beta_{in} \approx \frac{Z_{in}}{2R_{st,in}l_{in}} + j\omega\sqrt{L_{in}(C_{in} + \frac{C_{st,in}}{l_{in}})} \tag{5.15}$$

Similarly, the output impedance is no other than

$$Z_{0out} = \sqrt{\frac{R_{out} + j\omega L_{out}}{G_{out} + j\omega C_{out} + \frac{Y_{st,out}}{l_{out}}}} \tag{5.16}$$

The complex propagation constant is thus given by

$$\gamma_{out} = \alpha_{out} + j\beta_{out} \approx \frac{Z_{out}}{2R_{st,out}l_{out}} + j\omega\sqrt{L_{out}(C_{out} + \frac{C_{st,out}}{l_{out}})} \tag{5.17}$$

using equivalent nomenclature.

Amplification Stages

The outputs of the stages are added at the output transmission line, rather than multiplied. The outcome of this topology is that stages can be added to the TWA with minimal impact on the bandwidth. There are two considerations limiting the number of stages: signal attenuation and phase shift. Firstly, since the signal is attenuated when traveling along the input line, the line has limited length. Adding more stages and extending the line over its limit will not contribute any signal to the output, only noise.

The limit to number of stages due to attenuation is determined by the following phenomenon: from some number of stages, N, any additional stage will add a very small signal to the output — so small, that it is smaller than the attenuation in the output signal due to the additional stage. Taking into account only the attenuation, the power gain of a TWA is given by [3][1][7]

$$G_p = \frac{g_m^2 Z_L Z_S}{4} \cdot \left| \frac{e^{-N\gamma_{in}l_{in}} - e^{-N\gamma_{out}l_{out}}}{e^{-\gamma_{in}l_{in}} - e^{-\gamma_{out}l_{out}}} \right|^2 \tag{5.18}$$

where g_m is the stage's transconductance, Z_L is the load impedance, and Z_S is the source impedance. According to this, the optimal number of stages is given by

$$N_{opt} = \frac{\ln(\alpha_{in}l_{in}) - \ln(\alpha_{out}l_{out})}{\alpha_{in}l_{in} - \alpha_{out}l_{out}} \tag{5.19}$$

The second limit to the number of stages originates from both gain and noise reasons. Consider the two signal paths depicted in figure 5.7. The delay of the black path equals

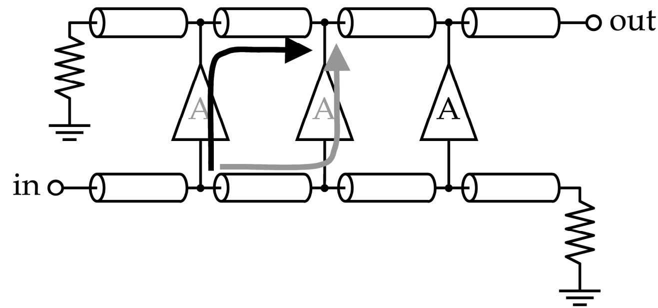


Figure 5.7: Two signal paths with different delays

to the delay of the amplification stage plus the delay of the output line (determined by the phase velocity of the line). The delay of the gray path equals to the stage delay plus the input line delay. If the delays of the input and output lines aren't equal (which is true in most cases), the output signal smears out (disperses). The more stages the TWA employs, the more dispersed the output signal. If phases are shifted so that destructive interference occurs, gain saturates. The more stages in the TWA, the greater shift in phases is accumulated along the lines. Hence, number of stages is limited. Phases are perfectly aligned if, and only if, $\beta_{in} l_{in} = \beta_{out} l_{out}$ and stage delays are equal. Neglecting attenuation issues, the power gain of a TWA is given by [1][8]

$$G_p = \frac{g_m^2 Z_L Z_S}{4} \cdot \frac{1 - \cos(N\omega\Delta T)}{1 - \cos(\omega\Delta T)} \tag{5.20}$$

where ΔT is the delay difference between the paths through two adjacent stages and the corresponding line segments. Hence, the optimal number of stages due to phase shift is

$$N_{opt} = \frac{\pi}{\omega\Delta T} \tag{5.21}$$

where ω equals to the bandwidth of interest.

5.2 Distributed Amplifier Architecture

This section examines the architecture of the TWAs fabricated for this work. It introduces the attenuation compensation architecture,and discusses input impedance required for optoelectronic integration, number of stages, biasing, noise performance, and layout.

5.2.1 Amplification Stage

The architecture chosen for the TWAs is the attenuation compensation architecture [9]. This architecture exhibits the best figures of merit like high bandwidth/f_T ratio and low power consumption [10]. The concept underlying attenuation compensation is better understood by examining the circuit shown in figure 5.8. The amplification stage consists of a CC stage (Q1), followed by a cascode stage (Q2 and Q3). The extremely high output resistance of the cascode eliminates the shunt conductance, G, introduces by the amplification stage at the output line, and so reduces the losses along the line. The CC stage is loaded by the input capacitance of the cascode. This capacitive load can be modeled as a negative resistance at the CC's input (refer to section 4.1.1), which compensates for the losses in the input line. Obviously, this compensation should be used carefully to avoid instability of the TWA.

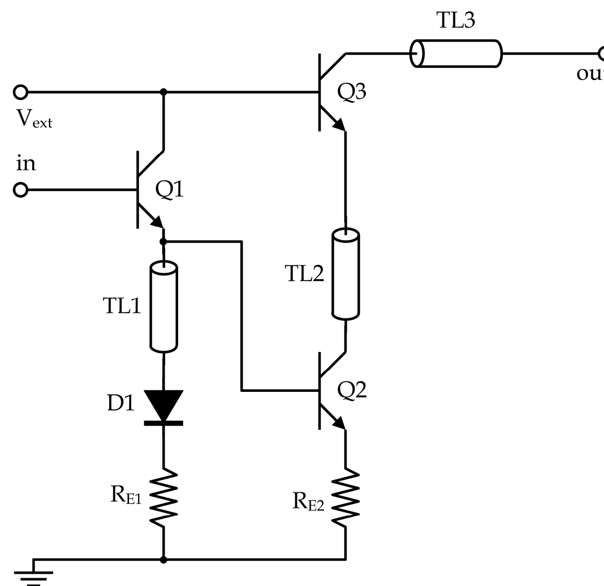


Figure 5.8: Scheme of an amplification stage. Q1 is the CC stage, Q2 and Q3 form the cascode

Three unmatched transmission lines appear in the circuit for inductive peaking that result in bandwidth enhancement [10]. TL1 loads the CC stage, whilst TL2 and TL3 work for the cascode. It should be noted that R_{E1} is relatively low — this relaxes the negative resistance effect caused by Q2, since the small resistance is connected in parallel to Q2's input impedance. In addition, as TL1 behaves like an inductor, it is reflected to Q1's input as a resistor, acting as a matching network between the input line and the amplification stage. The diode marked by D1 is intended for biasing purposes, and despite its existence a small degeneration resistor is added to Q2 (R_{E2}) to further stabilize its biasing point.

5.2.2 Input and Output Characteristic Impedances

The TWA is designed to connect with a 50 Ω system. As a consequence, the output impedance should be 50 Ω. The output line is loaded by the stages' output capacitances (C_μ of Q3), which is relatively small, and by ground air bridges. Hence, the line is designed to a higher impedance than 50 Ω to allow for impedance reduction by the additional capacitances.

As for the input line, its characteristic impedance should be as low as possible to obtain maximum bandwidth. Unfortunately, the attenuation compensation topology reduces the input capacitance of the stage, because the negative resistance is coupled with a negative capacitance [9]. This brings about the need to add capacitances in the form of ground air bridges. In addition, the input line is geometrically designed to have small Z_0. Shown in figure 5.9 is the simulated input impedance of a TWA. The impedance cannot be further reduced because of limitations such as physical dimensions of the transmission line and input capacitance of the stages.

On top of the above considerations, the lines should have the same phase velocity

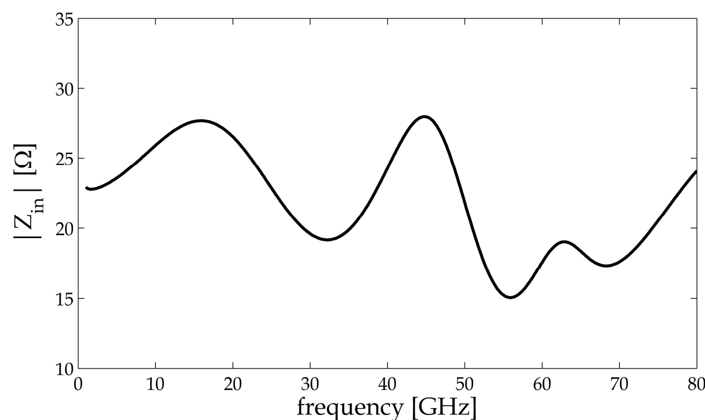


Figure 5.9: Input impedance of a TWA

and a sufficiently high Bragg frequency. The results of an optimization made for the transmission lines are detailed in table 5.3. Note that the above involves rough approximations regarding the stage's input capacitance and other parasitics. Simulations show that the input line Z_0 is approximately 25 Ω and phase velocity of $86 \cdot 10^6 \, \frac{\text{m}}{\text{sec}}$ — nearly matched to v_p of the output line.

Parameter	Value in the input line	Value in the output line
W	20 [μm]	14 [μm]
G	7 [μm]	36 [μm]
l	208 [μm]	208 [μm]
Parallel capacitance	25 [fF]	23 [fF]
Z_0	33 [Ω]	53 [Ω]
v_p	$92 \cdot 10^6 \left[\frac{\text{m}}{\text{sec}}\right]$	$83 \cdot 10^6 \left[\frac{\text{m}}{\text{sec}}\right]$
f_{Bragg}	141 [GHz]	127 [GHz]

Table 5.3: Actual values of TWA's transmission lines

5.2.3 Number of Stages

Number of stages was determined according to simulation results. The transimpedance gain, Z_T, of a 4-stage amplifier is 44.1 dBΩ; 5-stage amplifier exhibits 45.5 dBΩ; and 6 stages yield 46.4 dBΩ. In addition, 6-stage amplifier has poor noise performance and wavy group delay (sections 5.2.5 and 5.2.6 below) compared to 4- and 5-stage amplifiers. According to this, only 4- and 5-stage amplifiers were fabricated.

5.2.4 Biasing

As evident in figure 5.8 each amplification stage has three biasing voltages: the DC voltages of the input and output lines, and an extra biasing voltage, denoted by V_{ext}. In a complete TWA the input and output voltages are applied by bias tees directly to their corresponding lines, but V_{ext} requires a separate biasing line that travels through all of the stages. When a PD is connected to the amplifier's input a 1 kΩ resistor acts as a bias tee for the input line voltage (denoted by V_{PDn}), and another biasing pad (V_{PDp}) is connected directly to the p contact of the PD. The biasing scheme is illustrated in figure 5.10.

5.2.5 Noise Performance

There are 3 types of noise sources in a TWA:

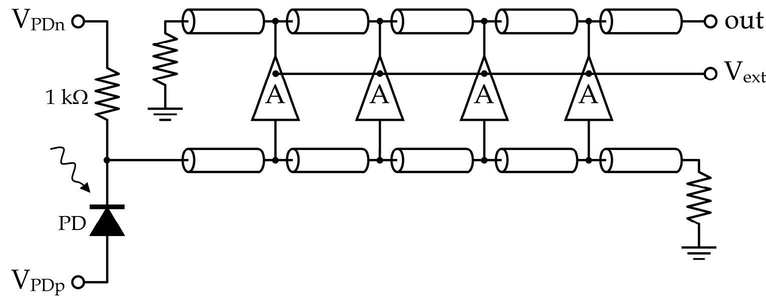


Figure 5.10: Scheme of TWA's biasing

1. Noise originated by the matching resistor of the output transmission line
2. Noise of the amplification stages
3. Noise of the input line matching resistor

The noise generated by the matching resistors are given by (2.70). However, noise of the output matching is attenuated when referred back to the input, whilst noise of the input matching isn't — it is even divided by the attenuation along the input transmission line, that is, amplified when referred to the input. As for the stages' noise, the last stage (the farthermost from the input) contributes more noise than the others due to the attenuation of the input line. Figure 5.11 demonstrates the effect of the last stage on noise performance. Shown in the figure are noise simulation results of an amplifier with 4 stages and a 5-stage amplifier.

The noise of each stage consists of the cascode noise, referred to the stage input through the CC stage. In addition, the CC transistor (Q1) contributes its own noise. Since a CC input stage exhibits poor voltage noise performance, the attenuation compensation architecture suffers from inferior noise performance compared to "standard" CE and cascode stages. Thanks to the low impedance of the input line required for transimpedance amplifiers, this voltage noise is relaxed to some extent when translated to current noise.

5.2.6 Group Delay

Group delay is a major concern in distributed amplifiers, especially when peaking techniques are excessively used. TWA's principle of operation consists mainly on capacitors and inductors, not on R–C networks. Inductive peaking also has the same effect. As a result, gain is very wavy compared to lumped topologies, and so is group delay. This phenomenon is on top of the dependence of group delay on frequency, as far as lossy transmission lines are considered.

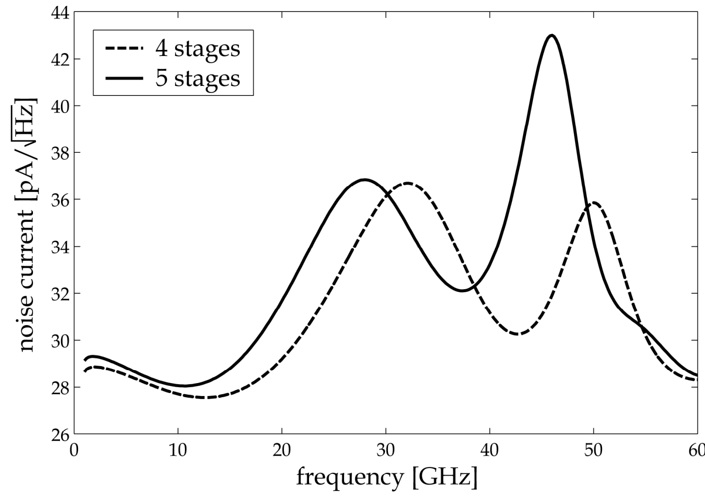
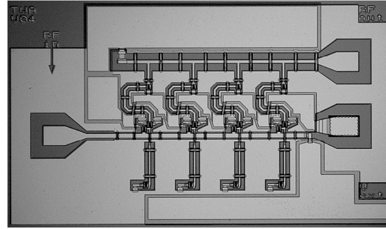


Figure 5.11: Simulated noise current of 4- and 5-stage amplifiers

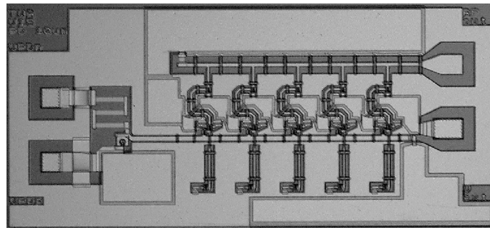
Variation in group delay should be kept below $\frac{1}{4}T$, where T is the time of a single bit transmission. This ensures that the eye (in eye diagram) is at least half open. To achieve this goal most of the peaking is performed by the transmission line between the cascode transistors (TL2), leaving the output peaking line (TL3) minimal. As TL2 is surrounded by two capacitors ($C_{\mu 2}$ and $C_{\pi 3}$), it is somewhat matched and provides a less wavy group delay curve [10].

5.2.7 Layout

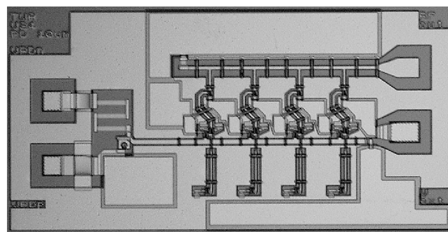
The same guidelines that appear in section 4.2.7 also apply here. Furthermore, due to the nature of TWAs extra attention was paid to parasitic capacitances and inductances. Three versions of circuits were designed, with strong, moderate, and weak peaking lines, which differ from each other by the length of TL2. Each version was fabricated with a 10 μm PD in addition to an electrical version without PD. On top of this, each circuit version was fabricated in 4- and 5-stage versions. Die sizes of electrical circuits are: 1758x1020 μm^2 for the 4-stage versions, and 1966x1020 μm^2 for the 5-stage versions. Die sizes of circuits with a PD are: 2047x1020 μm^2 in the 4 stages case, and 2255x1020 μm^2 for 5 stages. Pictures of 3 different circuits are shown in figure 5.12.



(a)



(b)



(c)

Figure 5.12: Layout of TWAs: (a) strong peaking, 4 stages, electrical version (b) moderate peaking, 5 stage, with PD (c) weak peaking, 4 stages, with PD

5.3 Optoelectronic Integration

Introducing a photodiode to a TWA is the subject of this section. It deals with the size of photodiode, the physical connection between the PD and the amplifier, and the optoelectronic performance.

5.3.1 Photodiode Connection

For a bandwidth exceeding 40 GHz the only photodiode that can be used is the smallest one, namely, the PD with 10 μm wide optical window. This diode exhibits the smallest capacitance (refer to table 3.5) but the same responsivity and transit time as the larger diodes. However, due to the input impedance of the TWA even the small PD's response might decrease before 40 GHz when connected to a TWA. To solve this problem the response of the amplifier is designed with peaking around 40 GHz to compensate for the PD response [11]. In addition, a 200 μm long transmission line was inserted between the PD and the amplifier, acting as a peaking inductor. As shown in figure 5.13, the photodiode has –3dB bandwidth of 37 GHz when connected to a 25 Ω load. The response of the amplifier, with strong peaking around 40 GHz, enhances the bandwidth of the optoelectronic circuit to 63 GHz.

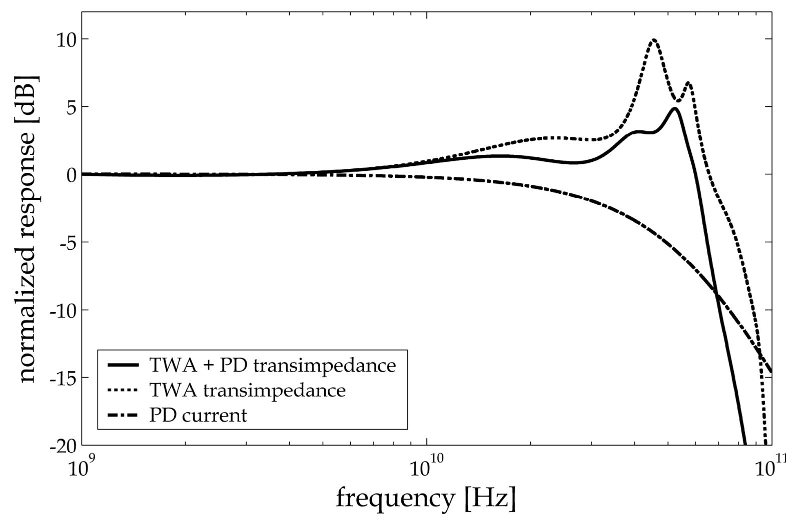


Figure 5.13: Normalized frequency response of a photodiode, TWA, and optoelectronic circuit. The photodiode was simulated with a 25 Ω load.

5.3.2 Optoelectronic Noise Performance

The current noise at the amplifier's input consists of the noise generated by the TWA, the noise of the PD, and the noise of the biasing resistor connected to V_{PDn} (figure 5.10). Once the noise current is known derivation of NEP is apparent, as described in section 4.3.3. Simulated noise and NEP of the various circuits are summarized in table 5.4. Calculations are based upon noise simulations, showing that all circuits exhibit noise current of approximately 29 pA/$\sqrt{\text{Hz}}$ up to 20 GHz. Including the exact noise spectrum is meaningless because of the limitations of the simplified VBIC model used for simulations. The DC current flowing in the PD is assumed 5 μA.

Collector thickness [Å]	Number of stages	Peaking strength	Optical bandwidth [GHz]	Noise spectral density [pA/$\sqrt{\text{Hz}}$]	Total noise [μA]	NEP [μW]
8000	4	Strong	55.5	29	8.6	25.2
8000	4	Moderate	59.2	29	8.8	26.0
8000	4	Weak	63.4	29	9.2	26.0
8000	5	Strong	54.1	29	8.5	24.9
8000	5	Moderate	57.9	29	8.7	25.7
8000	5	Weak	63.3	29	9.1	26.9
6000	4	Strong	54.0	29	8.4	47.4
6000	4	Moderate	57.6	29	8.7	49.0
6000	4	Weak	64.4	29	9.2	51.8
6000	5	Strong	53.2	29	8.4	47.1
6000	5	Moderate	57.0	29	8.7	48.7
6000	5	Weak	63.8	29	9.2	51.6

Table 5.4: NEP of optoelectronic distributed circuits

5.4 Circuit Performance

Distributed amplifiers based upon attenuation compensation architecture are vulnerable to changes in transistor parameters. However, proper adjustment of the biasing can regain the desired performance. All measurement and performance issues are discussed in this section.

5.4.1 Measurements

Electrical measurements of TWAs include only S–parameter measurements. The optoelectronic setup is similar to sections 3.3.1 and 4.4.2. In both measurements DC supplies can be applied either by a DC needle or by a dual G-S-G-S-G probe with appropriate bias tee.

5.4.2 TWA Performance

Simulation and measurement differ because of the following reasons:

1. The transistors exhibit large C_μ and small τ_D for which the circuits have not been designed
2. The VBIC model used in this work does not include modeling of variations in C_μ due to biasing or Kirk effect
3. The model does not include avalanche breakdown and thus biasing points aren't simulated properly

As a consequence, practical circuits allow some adjustment of their biasing points to overcome the changes in C_μ and τ_D. This option lacks in simulations. A comparison between simulation and measurement is demonstrated in figure 5.14. Also the group delay measured results are considerably different from simulated ones for the very same reasons. This is depicted in figure 5.15. Simulated group delay varies 8 psec peak-to-peak, whilst measured delay variations exceed 13 psec (when a centered average line is taken). The wavy group delay is a direct result of the too short τ_D of the transistors.

Output matching (return loss, S_{22}) is affected by C_μ, which is the output capacitance of the amplification stage. Due to the higher-than-expected capacitance return loss is better than –10 dB, instead of –18 dB expected from simulations. This is depicted in figure 5.16. Finally, simulated and measured performances are compared in table 5.5.

5.4.3 OEIC Performance

Measured and simulated results are detailed in table 5.6. The optoelectronic gain is calculated according to (4.94). Like in the TIA case, the 6000 Å circuits exhibit lower opto-

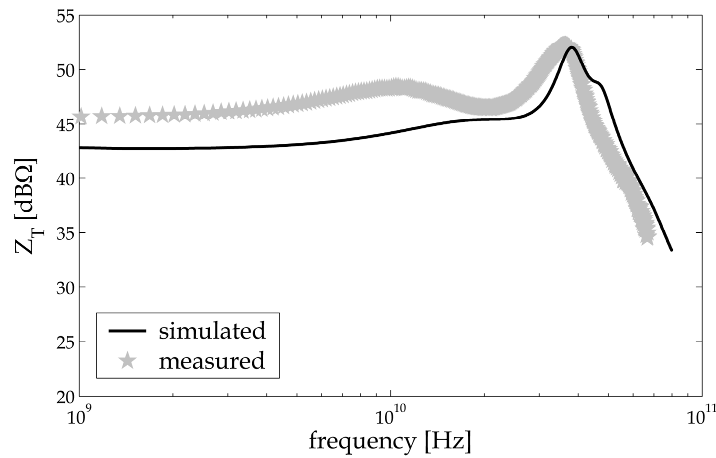


Figure 5.14: Simulated and measured TWA transimpedance gain (5-stage, strong peaking, 8000 Å amplifier)

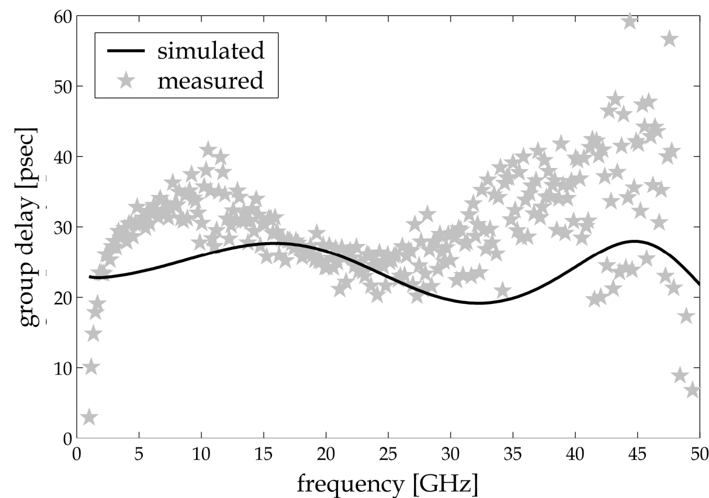


Figure 5.15: Simulated and measured TWA group delay (4-stage, weak peaking, 8000Å amplifier)

electronic gain due to the lower responsivity of the photodiodes. However, these circuits have wider bandwidth in comparison to the 8000 Å circuits. This difference is because the non-depleted region in BC junction is more than 2000 Å thick in the 8000 Å wafer, and only 1500 Å in the 6000 Å wafer. Consequently, the delay originated by diffusion is

longer in the 8000 Å PDs than in the 6000 Å.

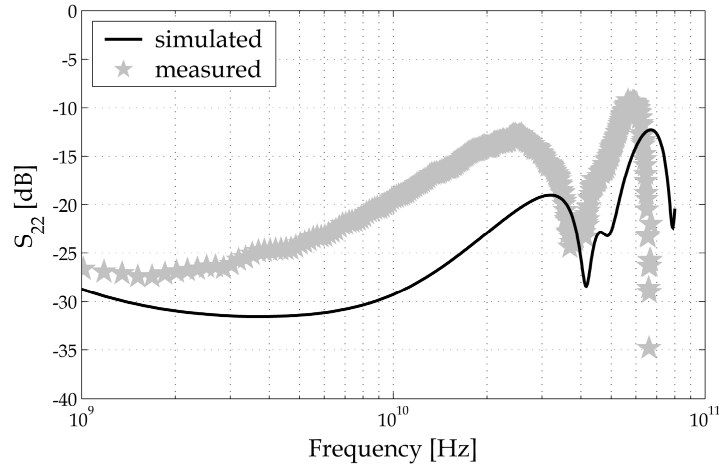


Figure 5.16: Simulated and measured output return loss (4-stage, strong peaking, 8000Å amplifier)

Collector thickness [Å]	Number of stages	Peaking strength	Simulated bandwidth [GHz]	Measured Bandwidth [GHz]	Simulated Gain [dBΩ]	Measured Gain [dBΩ]
8000	4	Strong	65.2	51.9	41.0	44.5
8000	4	Moderate	69.0	$\sim$ 68	41.9	45.2
8000	4	Weak	80.0	55.1	41.3	45.3
8000	5	Strong	62.8	62.0	42.8	42.1
8000	5	Moderate	68.4	49.2	42.5	45.7
8000	5	Weak	76.4	$\sim$ 70	41.9	43.4
6000	4	Strong	60.4	$\sim$ 70	44.3	39.8
6000	4	Moderate	65.2	$\sim$ 80	44.0	38.3
6000	4	Weak	73.8	$\sim$ 80	44.0	43.3
6000	5	Strong	57.8	59.3	45.8	42.2
6000	5	Moderate	62.2	54.1	45.8	44.8
6000	5	Weak	68.8	$\sim$ 70	45.8	44.1

Table 5.5: Simulated and measured TWA electrical performances

Collector thickness [Å]	Number of stages	Peaking strength	Simulated Optical bandwidth [GHz]	Measured Optical Bandwidth [GHz]	Measured Gain [V/W]
8000	4	Strong	55.5	9.8	57.8
8000	4	Moderate	59.2	12.2	62.6
8000	4	Weak	63.4	13.7	63.3
8000	5	Strong	54.1	14.8	43.8
8000	5	Moderate	57.9	13.4	66.3
8000	5	Weak	63.3	10.4	50.9
6000	4	Strong	54.0	17.6	17.6
6000	4	Moderate	57.6	20.2	14.8
6000	4	Weak	64.4	20.2	26.3
6000	5	Strong	53.2	17.5	23.2
6000	5	Moderate	57.0	20.2	31.3
6000	5	Weak	63.8	19.4	28.9

Table 5.6: Simulated and measured TWA optoelectronic performances

References for Chapter 5

[1] Emanuel Cohen, "Design of wide-band distributed amplifier based on InP/GaInAs hetrojunction bipolar transistors", Master's thesis, Technion – IIT, 2002.

[2] Guillermo Gonzalez, *Microwave Transistor Amplifiers – Analysis and Design*, Prentice–Hall, second edition, 1997.

[3] Thomas H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*, Cambridge University Press, 1998.

[4] Rainee N. Simons, *Coplanar Waveguide Circuits, Components, and Systems*, John Wiley & Sons, 2001.

[5] Bal S. Virdee, Avtar S. Virdee, and Ben Y. Banyamin, *Broadband Microwave Amplifiers*, Artech House, 2004.

[6] David M. Pozar, *Microwave Engineering*, John Wiley & Sons, second edition, 1998.

[7] I. D. Robertson and S. Lucyszyn (Editors), *RFIC and MMIC Design and Technology*, The Institution of Electrical Engineers, 2001.

[8] Mark J. W. Rodwell, Scott T. Allen, Ruai Y. Yu, Michael G. Case, Uddalak Bhattacharya, Madhukar Reddy, Eric Carman, Masayuki Kamegawa, Yoshiyuki Konishi, Joe Pusl, and Rajasekhar Pullela, "Active and nonlinear wave propagation devices in ultrafast electronics and optoelectronics", *Proceedings of the IEEE*, vol. 82, no. 7, pp. 1037–1059, July 1994.

[9] Kevin W. Kobayashi, Reza Esfandiari, and Aaron K. Oki, "A novel HBT distributed amplifier design topology based on attenuation compensation techniques", *IEEE Transactions on Microwave Theory and Techniques*, vol. 42, no. 12, pp. 2583–2589, December 1994.

[10] E. Cohen, Y. Betser, B. Sheinman, S. Cohen, S. Sidorov, A. Gavrilov, and D. Ritter, "75 GHz InP HBT distributed amplifier with record figures of merit and low power dissipation", *IEEE Transactions on Electron Devices*, vol. 53, no. 2, pp. 392–394, February 2006.

[11] J. Bellon and M. J. N. Sibley, "Frequency response compensation of transit time limited pin photodiode", *Electronics Letters*, vol. 36, no. 14, pp. 1222–1223, July 2000.

Chapter 6

Summary and Conclusions

The results of the research are summarized and discussed in this last chapter. This work is compared to the works introduced in chapter 1, and conclusions and directions for further research are suggested.

6.1 Photodiode

Performance of the PDs were below expectations due to a problem in the layer growth (refer to section 3.3.3 and figure 3.15). Recently this problem was solved by means of adding another stage to layer growth sequence. C–V measurements were made on large area devices fabricated from a wafer with 6000 Å collector. The results verify that the problem has been solved, as seen in figure. The collector is almost completely depleted from a reverse voltage of 2 V, which is adequate for proper operation of the PDs. New 6000 Å and 8000 Å wafers were grown and are currently processed into receivers.

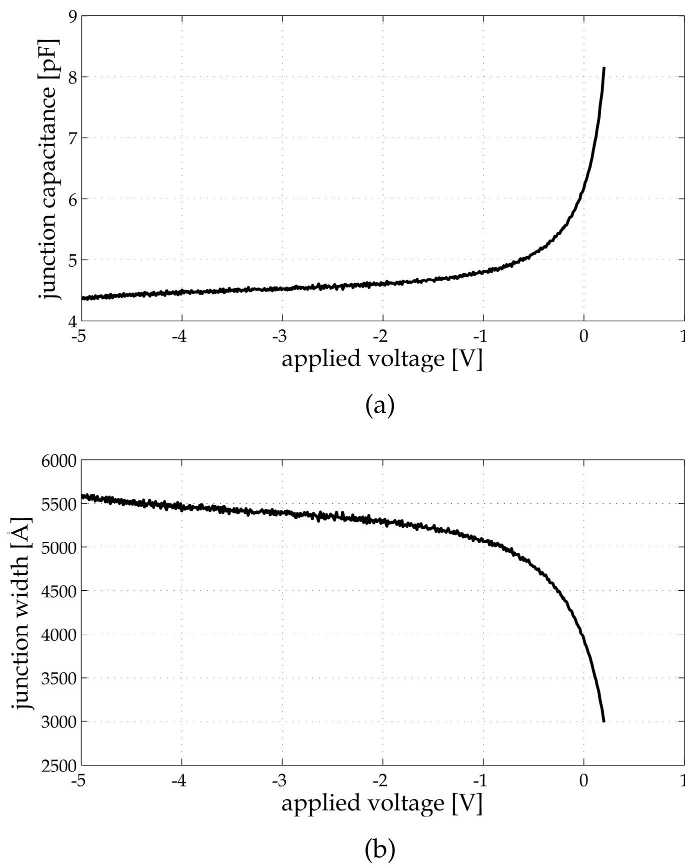


Figure 6.1: C–V measurements of a 6000 Å wafer: (a) junction capacitance vs. applied voltage (b) junction width vs. applied voltage (measured on a large area device)

6.2 Lumped Circuits

Two topologies of lumped TIA were introduces - version A (with a CB input stage) and version B (without input stage). Version A exhibits wider bandwidth when connected to a PD due to the input stage, but version B has better noise performance due to the lack of input stage. Best optoelectronic results were achieved with version A and a 10 μm photodiode in the 8000 Å wafer, and these results are used for comparison in table 6.1. Table 6.1 implies that should the circuit be fabricated from properly grown layers and with ARC, it would exhibit the highest optoelectronic gain-bandwidth product in its class.

Team $\longrightarrow$	ETHZ Switzerland [1]	ETHZ Switzerland [2]	Notre Dame IN, USA [3]	KAIST Korea [4]	This Work
Transistor	HBT	HBT	HEMT	HBT	HBT
Photodiode	top illuminated	top illuminated	top illuminated	refractive facet	top illuminated
Anti-reflective Coating	yes	yes	yes	yes	no
Optical Bandwidth [GHz]	30	50	8.3	6.9	14.2 (expected to reach 20)
Optoelectronic Gain [V/W]	48	N/A	410	85	61.2 (can exceed 120 with ARC)
Gain-Bandwidth Product $\left[\mathrm{THz} \cdot \frac{\mathrm{V}}{\mathrm{W}}\right]$	1.44	N/A	3.40	0.59	0.47 (expected to exceed 1.7)

Table 6.1: Comparison of this work to recent works on InP-based lumped photoreceivers

Further improvements may be applied to the circuit, as follows:

- Better biasing can reduce the current flowing in the input stage (Q1) and reduce the circuit's noise
- Different topologies may be examined, such as omitting the first emitter follower (Q4) and connecting the feedback resistor directly from the cascode's output to its input
- Phase margin can be designed in a self-adjustment topology, e.g. using capacitors made of the base–collector layers

In addition, improvements in the fabrication and process technology will enable to achieve better results. These improvements include better layer structure, reduction in lithography size [2], and implementation of capacitance reduction techniques (such as undercut [2] and base air bridge [5]).

6.3 Distributed Circuits

Distributed amplifiers designed in the attenuation compensation architecture were demonstrated. The circuits have input impedance of approximately 25 Ω and bandwidth in excess of 50 GHz. Best optoelectronic results were achieved with a 4-stage weak-peaking amplifier in the 6000 Å wafer. The results of this amplifier appear in table 6.2 and compared to the other works. In fact, this is the ever first distributed receiver that consists of HBT and top illuminated PD. If the circuit is fabricated from a better wafer with ARC it will offer a cost-effective solution for communications in 40 Gbps and beyond.

Team $\longrightarrow$	NTT Corp. Japan [6]	HHI Berlin Germany [7]	This Work
Transistor	HEMT	HEMT	HBT
Photodiode	waveguide	waveguide	top illuminated
Anti-reflective Coating	yes	yes	no
Optical Bandwidth [GHz]	46.5	72.0	20.2 (expected to reach 48)
Optoelectronic Gain [V/W]	55	45	26.3 (can exceed 50 with ARC)
Gain-Bandwidth Product $\left[\mathrm{THz} \cdot \frac{\mathrm{V}}{\mathrm{W}}\right]$	2.56	3.24	0.53 (expected to exceed 2.4)

Table 6.2: Comparison of this work to recent works on InP-based distributed photoreceivers

Further research may focus on examining other architectures. Amplifiers based on cascode stages are good candidates as they offer high bandwidth and high input capacitance, that can help to reduce the input impedance.

References for Chapter 6

[1] Martin Bitter, Raimond Bauknecht, Werner Hunziker, and Hans Melchior, "Monolithic InGaAs–InP p-i-n/HBT 40-Gb/s optical receiver module", *IEEE Photonics Technology Letters*, vol. 12, no. 1, pp. 74–76, January 2000.

[2] D. Huber, R. Bauknecht, C. Bergamaschi, M. Bitter, A. Huber, T. Morf, A. Neiger, M. Rohner, I. Schnyder, V. Schwarz, and H. Jäckel, "InP–InGaAs single HBT technology for photoreceiver OEICs at 40 Gb/s and beyond", *IEEE Journal of Lightwave Technology*, vol. 18, no. 7, pp. 992–1000, July 2000.

[3] P. Fay, C. Caneau, and I. Adesida, "High-speed MSM/HEMT and p-i-n/HEMT monolithic photoreceivers", *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 1, pp. 62–67, January 2002.

[4] Bangkeun Lee, Yongjoo Song, and Kyounghoon Yang, "InP-based OEIC photoreceivers using shared layer integration technology of heterojunction bipolar transistors and refracting-facet photodiodes", *2003 International Conference on Solid State Devices and Materials*, pp. 182–183, September 2003.

[5] Mathias Kahn, *Transistor Bipolaire À Hététrojonction GaInAs/InP Pour Circuits Ultra-Rapides : Structure, Fabrication et Caractérisation*, PhD thesis, Universite Paris XI Orsay, June 2004.

[6] Kiyoto Takahata, Yoshifumi Muramoto, Hideki Fukano, Kazutoshi Kato, Atsuo Kozen, Shunji Kimura, Yuhki Imai, Yutaka Miyamoto, Osaake Nakajima, and Yutaka Matsuoka, "Ultrafast monolithic receiver OEIC composed of multimode waveguide p-i-n photodiode and HEMT distributed amplifier", *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 6, no. 1, pp. 31–37, January/February 2000.

[7] Gebre Giorgis Mekonnen, Heinz-Gunter Bach, Andreas Beling, Reinhard Kunkel, Detlef Schmidt, and Wolfgang Schlaak, "80-Gb/s InP-based waveguide-integrated photoreceiver", *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 11, no. 2, pp. 356–360, March/April 2005.

Appendix A

A Control GUI for C–V Measurements

During the work on this research I developed a GUI (graphical user interface) for controlling a HP 4280A C-V Plotter and performing a C–V measurement using a PC. Details are discussed hereinafter as it can be of further help for anyone who uses a 4280A for C–V measurements and wishes to control it from a PC.

A.1 Introduction

The GUI runs on MATLAB 6.5 and consists of a figure file (*cv_main.fig*) and MATLAB M-file (*cv_main.m*). The GUI is invoked from an envelope function (*cv.m*) that also sets variables to defaults and establishes connection to the instrument through GPIB interface.

Requirements from the PC are as follows:

- MATLAB 6.5 with instrumentation toolbox installed
- GPIB card — supported by MATLAB

In the following example a National Instruments GPIB card is assumed. In addition, the GPIB address of the 4280A is 27. This can be set by two dials at the back of the 4280A, or changed in the program to match the instrument's address. The program was tested and works with MATLAB 6.5.0 (R13) and National Instrument USB-GPIB-B card and its driver.

It should be noted that every button or text editing box in a figure has its own name (*tag*) and is connected to a *callback* function in the corresponding M-file. Tags can be derived from the M-file.

A.2 MATLAB Figures

cv_main.fig

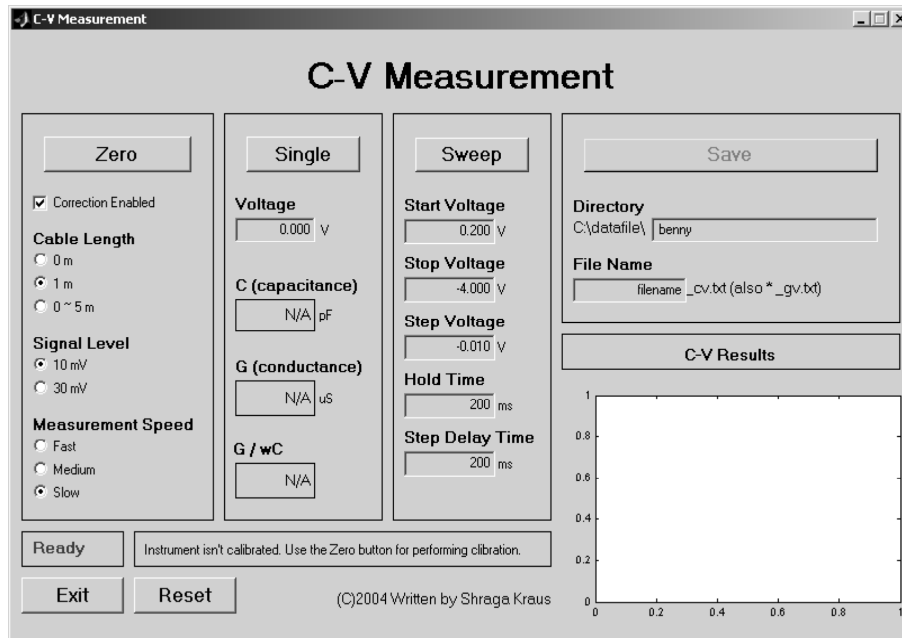


Figure A.1: MATLAB figure of cv_main

gpib_error.fig

Figure A.2: MATLAB figure of gpib_error

save_error.fig

Figure A.3: MATLAB figure of save_error

save_ok.fig

Figure A.4: MATLAB figure of save_ok

A.3 MATLAB Functions

cv.m

```
function cv

global v vh gh vcg
close all

% Defaults
v.gpib_err = 0;
```

```
v.calibrated = 0;
v.data_to_save = 0;
v.single_err = 0;
v.sweep1_err = 0;
v.sweep2_err = 0;
v.sweep3_err = 0;
v.sweep4_err = 0;
v.sweep5_err = 0;
v.step_err = 0;
v.save_err = 0;
v.filename_err = 0;
v.correct_en = 1;
v.cable_length = 1;
v.signal_level = 10;
v.meas_speed = 'slow';
v.single_v = 0;
v.single_c = -1;
v.single_g = -1;
v.single_gwc = -1;
v.v_start = 0.2;
v.v_stop = -4;
v.v_step = -0.01;
v.t_hold = 200;
v.t_step_delay = 200;
v.save_dir = 'benny';
v.save_filename = 'filename';

% Establish GPIB connection
gh = gpib('ni', 0, 27); % National Instrument card, card No.0, GPIB address = 27
fopen(gh);
ghstatus = gh.Status;
if (ghstatus(1:4) == 'clos') % error in GPIB connection
   gpib_error_message;
   return
end
set(gh,'EOSMode','read&write') %'write')
clrdevice(gh) % restore instrument defaults

% Open CV window
hfig = openfig('cv_main', 'reuse'); %, 'visible');
set(hfig, 'handlevisibility', 'on', 'doublebuffer', 'on')
vh = guihandles(hfig);
```

cv_main.m

```
function varargout = cv_main(varargin)

global v vh gh vcg

if nargin == 0  % LAUNCH GUI
fig = openfig(mfilename,'reuse');
% Use system color scheme for figure:
```

```
set(fig,'Color',get(0,'defaultUicontrolBackgroundColor'));
% Generate a structure of handles to pass to callbacks, and store it.
handles = guihandles(fig);
guidata(fig, handles);
if nargout > 0
varargout{1} = fig;
end
elseif ischar(varargin{1}) % INVOKE NAMED SUBFUNCTION OR CALLBACK
try
[varargout{1:nargout}] = feval(varargin{:}); % FEVAL switchyard
catch
disp(lasterr);
end
end

% --------------------------------------------------------------------
function gpib_zero (h, eventdata, handles, varargin)
global v vh gh vcg
set(vh.status_cv, 'foregroundcolor', [0 0 1], 'string', 'Busy');
set(vh.zero_cv, 'foregroundcolor', [0.5 0.5 0.5], ...
    'enable' ,'inactive');
set(vh.message, 'string', 'Calibration in progress...');
drawnow
go_zero
set(vh.zero_cv, 'enable' ,'on', 'foregroundcolor', [0 0 0]);
set(vh.single_c, 'string', 'N/A');
set(vh.single_g, 'string', 'N/A');
set(vh.single_gwc, 'string', 'N/A');
status_message

% --------------------------------------------------------------------
function gpib_single (h, eventdata, handles, varargin)
global v vh gh vcg
set(vh.status_cv, 'foregroundcolor', [0 0 1], 'string', 'Busy');
set(vh.single_cv, 'foregroundcolor', [0.5 0.5 0.5], ...
    'enable' ,'inactive');
set(vh.message, 'string', 'Single measurement in progress...');
drawnow
go_single
if ((v.single_c == -1) | (v.single_g == -1) | (v.single_gwc == -1))
    c_str = 'N/A';
    g_str = 'N/A';
    gwc_str = 'N/A';
else
    c_str = sprintf('%g', v.single_c * 1e12);
    g_str = sprintf('%g', v.single_g * 1e6);
    gwc_str = sprintf('%g', v.single_gwc);
end
set(vh.single_c, 'string', c_str);
set(vh.single_g, 'string', g_str);
set(vh.single_gwc, 'string', gwc_str);
status_message
```

```
% ------------------------------------------------------------------
function gpib_sweep (h, eventdata, handles, varargin)
global v vh gh vcg
button_string = get(vh.sweep_cv, 'string');
if (length(button_string) == 5) % sweep button pressed
    set(vh.sweep_cv, 'string', 'Stop');
    set(vh.zero_cv, 'foregroundcolor', [0.5 0.5 0.5], 'enable' ,'inactive');
    set(vh.single_cv, 'foregroundcolor', [0.5 0.5 0.5], 'enable' ,'inactive');
    set(vh.save_cv, 'foregroundcolor', [0.5 0.5 0.5], 'enable' ,'inactive');
    set(vh.exit_request, 'foregroundcolor', [0.5 0.5 0.5], 'enable' ,'inactive');
    set(vh.status_cv, 'foregroundcolor', [0 0 1], 'string', 'Busy');
    set(vh.message, 'string', 'Sweep measurement in progress...');
    drawnow
    go_sweep
    set(vh.sweep_cv, 'string', 'Sweep');
    correct_en
    set(vh.single_cv, 'foregroundcolor', [0 0 0], 'enable' ,'on');
    set(vh.save_cv, 'foregroundcolor', [0 0 0], 'enable' ,'on');
    set(vh.exit_request, 'foregroundcolor', [0 0 0], 'enable' ,'on');
else % stop button pressed
    stop_sweep
    set(vh.sweep_cv, 'string', 'Sweep');
    correct_en
    set(vh.single_cv, 'foregroundcolor', [0 0 0], 'enable' ,'on');
    set(vh.save_cv, 'foregroundcolor', [0 0 0], 'enable' ,'on');
    set(vh.exit_request, 'foregroundcolor', [0 0 0], 'enable' ,'on');
end
status_message

% ------------------------------------------------------------------
function save_cv (h, eventdata, handles, varargin)
global v vh gh vcg
go_save

% ------------------------------------------------------------------
function exit_request (h, eventdata, handles, varargin)
global v vh gh vcg
fclose(gh)
delete(gh)
clear gh
close
exit

% ------------------------------------------------------------------
function gpib_reset (h, eventdata, handles, varargin)
global v vh gh vcg
% Check if a sweep is currenly running
button_string = get(vh.sweep_cv, 'string');
if (length(button_string) == 4) % button is "Stop" labeled,
                                % sweep is running
    stop_sweep
```

```
    pause(0.5)
end
clrdevice(gh) % restore device defaults
v.calibrated = 0;
set(vh.correct_en, 'value', 1);
correct_en
status_message

% ------------------------------------------------------------------
function correct_en(h, eventdata, handles, varargin)
global v vh gh vcg
v.correct_en = get(vh.correct_en, 'value');
if (v.correct_en == 1)
    set(vh.zero_cv, 'enable' ,'on', 'foregroundcolor', [0 0 0]);
    set(vh.m0, 'enable' ,'on', 'foregroundcolor', [0 0 0]);
    set(vh.m1, 'enable' ,'on', 'foregroundcolor', [0 0 0]);
    set(vh.m0_5, 'enable' ,'on', 'foregroundcolor', [0 0 0]);
    set(vh.cable_length, 'foregroundcolor', [0 0 0]);
else
    set(vh.zero_cv, 'enable' ,'inactive', 'foregroundcolor', [0.5 0.5 0.5]);
    set(vh.m0, 'enable' ,'inactive', 'foregroundcolor', [0.5 0.5 0.5]);
    set(vh.m1, 'enable' ,'inactive', 'foregroundcolor', [0.5 0.5 0.5]);
    set(vh.m0_5, 'enable' ,'inactive', 'foregroundcolor', [0.5 0.5 0.5]);
    set(vh.cable_length, 'foregroundcolor', [0.5 0.5 0.5]);
end
go_correct_en

% ------------------------------------------------------------------
function cl0(h, eventdata, handles, varargin)
global v vh gh vcg
v.cable_length = 0;
set(vh.m1, 'value' ,0);
set(vh.m0_5, 'value' ,0);

% ------------------------------------------------------------------
function cl1(h, eventdata, handles, varargin)
global v vh gh vcg
v.cable_length = 1;
set(vh.m0, 'value' ,0);
set(vh.m0_5, 'value' ,0);

% ------------------------------------------------------------------
function cl05(h, eventdata, handles, varargin)
global v vh gh vcg
v.cable_length = 5;
set(vh.m0, 'value' ,0);
set(vh.m1, 'value' ,0);

% ------------------------------------------------------------------
function sl10(h, eventdata, handles, varargin)
global v vh gh vcg
v.signal_level = 10;
```

```
set(vh.mv30, 'value' ,0);

% --------------------------------------------------------------------
function sl30(h, eventdata, handles, varargin)
global v vh gh vcg
v.signal_level = 30;
set(vh.mv10, 'value' ,0);

% --------------------------------------------------------------------
function msfast(h, eventdata, handles, varargin)
global v vh gh vcg
v.meas_speed = 'fast';
set(vh.medium, 'value' ,0);
set(vh.slow, 'value' ,0);

% --------------------------------------------------------------------
function msmed(h, eventdata, handles, varargin)
global v vh gh vcg
v.meas_speed = 'medi';
set(vh.fast, 'value' ,0);
set(vh.slow, 'value' ,0);

% --------------------------------------------------------------------
function msslow(h, eventdata, handles, varargin)
global v vh gh vcg
v.meas_speed = 'slow';
set(vh.fast, 'value' ,0);
set(vh.medium, 'value' ,0);

% --------------------------------------------------------------------
function single_v(h, eventdata, handles, varargin)
global v vh gh vcg
v.single_v_str = get(vh.single_v, 'string');
v.single_v = str2num(v.single_v_str);
if ((v.single_v < 42) & (v.single_v > (-42)))
    v.single_err = 0;
    single_v_str = sprintf('%2.3f', v.single_v);
    set(vh.single_v, 'string', single_v_str);
else
    v.single_err = 1;
end
status_message

% --------------------------------------------------------------------
function v_start(h, eventdata, handles, varargin)
global v vh gh vcg
v.v_start_str = get(vh.v_start, 'string');
v.v_start = str2num(v.v_start_str);
if ((v.v_start < 42) & (v.v_start > (-42)))
    v.sweep1_err = 0;
    v_start_str = sprintf('%2.3f', v.v_start);
    set(vh.v_start, 'string', v_start_str);
```

```
    v.v_start = str2num(v_start_str); % chopping voltage smaller than 1mV
else
    v.sweep1_err = 1;
end
status_message

% -------------------------------------------------------------------
function v_stop(h, eventdata, handles, varargin)
global v vh gh vcg
v.v_stop_str = get(vh.v_stop, 'string');
v.v_stop = str2num(v.v_stop_str);
if ((v.v_stop < 42) & (v.v_stop > (-42)))
    v.sweep2_err = 0;
    v_stop_str = sprintf('%2.3f', v.v_stop);
    set(vh.v_stop, 'string', v_stop_str);
    v.v_stop = str2num(v_stop_str); % chopping voltage smaller than 1mV
else
    v.sweep2_err = 1;
end
status_message

% -------------------------------------------------------------------
function v_step(h, eventdata, handles, varargin)
global v vh gh vcg
v.v_step_str = get(vh.v_step, 'string');
v.v_step = str2num(v.v_step_str);
if ((v.v_step < 42) & (v.v_step > (-42)))
    v.sweep3_err = 0;
    v_step_str = sprintf('%2.3f', v.v_step);
    set(vh.v_step, 'string', v_step_str);
    v.v_step = str2num(v_step_str); % chopping voltage smaller than 1mV
else
    v.sweep3_err = 1;
end
status_message

% -------------------------------------------------------------------
function t_hold(h, eventdata, handles, varargin)
global v vh gh vcg
v.t_hold_str = get(vh.t_hold, 'string');
v.t_hold = str2num(v.t_hold_str);
if ((v.t_hold < 650) & (v.t_hold > 45))
    v.sweep4_err = 0;
    t_hold_str = sprintf('%3.0f', v.t_hold);
    set(vh.t_hold, 'string', t_hold_str);
    v.t_hold = str2num(t_hold_str); % chopping digits after decimal point
else
    v.sweep4_err = 1;
end
status_message

% -------------------------------------------------------------------
```

```
function t_step_delay(h, eventdata, handles, varargin)
global v vh gh vcg
v.t_step_delay_str = get(vh.t_step_delay, 'string');
v.t_step_delay = str2num(v.t_step_delay_str);
if ((v.t_step_delay < 650) & (v.t_step_delay > 45))
    v.sweep5_err = 0;
    t_step_delay_str = sprintf('%3.0f', v.t_step_delay);
    set(vh.t_step_delay, 'string', t_step_delay_str);
    v.t_step_delay = str2num(t_step_delay_str); % chopping digits after decimal point
else
    v.sweep5_err = 1;
end
status_message

% ---------------------------------------------------------------------
function save_dir(h, eventdata, handles, varargin)
global v vh gh vcg
v.save_dir = get(vh.directory, 'string');
save_dir = strcat('C:\datafile\', v.save_dir);
if (save_dir(length(save_dir)) ~= '\')
    save_dir = strcat(save_dir, '\'); % add backslash to full path string
else  % chop backslash from string on screen
    dir_len = length(v.save_dir) - 1;
    v.save_dir = v.save_dir(1:dir_len);
    set(vh.directory, 'string', v.save_dir);
end
if (exist(save_dir, 'dir'))
    v.save_err = 0;
else
    v.save_err = 1;
end
status_message

% ---------------------------------------------------------------------
function save_filename(h, eventdata, handles, varargin)
global v vh gh vcg
v.save_filename = get(vh.filename, 'string');
if (findstr('\', v.save_filename))
    v.filename_err = 1;
else
    v.filename_err = 0;
end
status_message

% ---------------------------------------------------------------------
function status_message
global v vh gh vcg

% Check if there's step error
if (v.sweep1_err + v.sweep2_err + v.sweep3_err == 0)
    v_diff = v.v_stop - v.v_start;
    if ( (sign(v_diff) ~= sign(v.v_step)) | (v.v_step == 0) | (abs(v.v_step) > abs(v_diff)) )
```

```
        v.step_err = 1;
    else
        v.step_err = 0;
    end
end

if (v.single_err + v.sweep1_err  + v.sweep2_err + v.sweep3_err + ...
        v.sweep4_err + v.sweep5_err + v.step_err + v.save_err + ...
        v.filename_err == 0) % no input error
    set(vh.status_cv, 'foregroundcolor', [0 0.407 0], 'string', 'Ready');
    set(vh.single_cv, 'foregroundcolor', [0 0 0], 'enable' ,'on');
    set(vh.sweep_cv, 'foregroundcolor', [0 0 0], 'enable' ,'on');
    set(vh.save_cv, 'foregroundcolor', [0 0 0], 'enable' ,'on');
    % indicate ready on screen
    if (v.calibrated == 0)
        set(vh.message, 'string', ...
            'Instrument isn''t calibrated. Use the Zero button for performing clibration.');
    else
        set(vh.message, 'string', ...
            strcat('Instrument is calibrated. Connect probe', ...
            's to DUT and measure.'));
    end

else % there is any input error
    if (v.sweep1_err  + v.sweep2_err + v.sweep3_err + v.sweep4_err + ...
            v.sweep5_err + v.step_err > 0) % error in sweep inputs
        set(vh.sweep_cv, 'foregroundcolor', [0.5 0.5 0.5], ...
            'enable' ,'inactive');
    else
        set(vh.sweep_cv, 'foregroundcolor', [0 0 0], 'enable' ,'on');
    end
    if (v.single_err > 0) % error in single input
        set(vh.single_cv, 'foregroundcolor', [0.5 0.5 0.5], 'enable' ,'inactive');
    else
        set(vh.single_cv, 'foregroundcolor', [0 0 0], 'enable' ,'on');
    end

    % indicate error status on screen
    if (v.step_err > 0)
        set(vh.status_cv, 'foregroundcolor', [1 0 0], 'string', 'Error');
        set(vh.message, 'string', 'Illegal step voltage specified.');
    end
    if (v.sweep1_err  + v.sweep2_err + v.sweep3_err + v.sweep4_err + ...
            v.sweep5_err + v.single_err)
        set(vh.status_cv, 'foregroundcolor', [1 0 0], 'string', 'Error');
        set(vh.message, 'string', 'Illegal voltage or time value.');
    end

    if (v.filename_err > 0) % error in save filename
        set(vh.save_cv, 'foregroundcolor', [0.5 0.5 0.5], 'enable' ,'inactive');
        set(vh.status_cv, 'foregroundcolor', [1 0 0], 'string', 'Error');
        set(vh.message, 'string', 'Illegal save filename.');
```

```
    end

    if (v.save_err > 0) % error in save directory
        set(vh.save_cv, 'foregroundcolor', [0.5 0.5 0.5], 'enable' ,'inactive');
        set(vh.status_cv, 'foregroundcolor', [1 0 0], 'string', 'Error');
        set(vh.message, 'string', 'Save directory non-existent.');
    end

end

if (v.data_to_save > 0) % there is data from sweep measurement
    if (v.save_err + v.filename_err > 0) % error in save inputs
        set(vh.save_cv, 'foregroundcolor', [0.5 0.5 0.5], ...
            'enable' ,'inactive');
    else
        set(vh.save_cv, 'foregroundcolor', [0 0 0], 'enable' ,'on');
    end
else
    set(vh.save_cv, 'foregroundcolor', [0.5 0.5 0.5], 'enable' ,'inactive');
end

drawnow

% ---------------------------------------------------------------------
function ok_callback (h, eventdata, handles, varargin)
close
```

go_correct_en.m

```
function go_correct_en

global v vh gh vcg

if (v.correct_en == 1)
    fprintf(gh, 'CE1')
else
    fprintf(gh, 'CE0')
end
pause(0.1)
```

go_save.m

```
function go_save

global v vh gh vcg

cv_name = strcat('c:\datafile\', v.save_dir, '\', v.save_filename, '_cv.txt');
gv_name = strcat('c:\datafile\', v.save_dir, '\', v.save_filename, '_gv.txt');
```

```
vcg_size = size(vcg);
vcg_len = vcg_size(1);
cv_data(1:vcg_len, 1) = vcg(1:vcg_len, 1);
cv_data(1:vcg_len, 2) = vcg(1:vcg_len, 2);
gv_data(1:vcg_len, 1) = vcg(1:vcg_len, 1);
gv_data(1:vcg_len, 2) = vcg(1:vcg_len, 3);
fid = fopen(cv_name, 'wt');
for n = 1:vcg_len
    str_line = sprintf('%e\t%e\n', vcg(n,1), vcg(n,2) );
    fprintf(fid, '%s', str_line);
end
fclose(fid);
fid2 = fopen(gv_name, 'wt');
for n = 1:vcg_len
    str_line = sprintf('%e\t%e\n', vcg(n,1), vcg(n,3) );
    fprintf(fid2, '%s', str_line);
end
fclose(fid2);

if ( (exist(cv_name, 'file')) & (exist(gv_name, 'file')) & ...
        (fid ~= -1) & (fid2 ~= -1) )
    save_ok_message
else
    save_error_message
end
```

go_single.m

```
function go_single

global v vh gh vcg

if (v.signal_level == 30) % set signal level
    fprintf(gh, 'SL2')
else
    fprintf(gh, 'SL1')
end

if (v.meas_speed == 'fast') % set measurement speed
    fprintf(gh, 'MS1')
elseif (v.meas_speed == 'medi')
    fprintf(gh, 'MS2')
else
    fprintf(gh, 'MS3')
end

pause(0.1)
fprintf(gh, 'TR3') % trigger mode = hold/manual
fprintf(gh, 'IB1') % voltage source mode = constant DC
PVnum = sprintf('%g', v.single_v);
PVstring = strcat('PV', PVnum);
fprintf(gh, PVstring) % set voltage of internal voltage source
```

```
fprintf(gh, 'VO1') % connect internal voltage source
fprintf(gh, 'BC') % clear output buffer
fprintf(gh, 'EX') % trigger the measurement
r = fscanf(gh); % a string that contains the results
fprintf(gh, 'VO0') % disconnect internal voltage source
fprintf(gh, 'TR1') % trigger mode = internal
pause(0.1)

% extract numeric results from string
rl = length(r);
comma_pos = 0;
for n = 1:rl
    if (r(n) == ',')
        comma_pos = n;
    end
end
if ( (comma_pos == 0) | (comma_pos == rl) )
    v.single_c = -1;
    v.single_g = -1;
    v.single_gwc = -1;
    gpib_error_massage
else
    c_str = r(4:comma_pos-1);
    g_str = r(comma_pos+4:rl);
    v.single_c = str2num(c_str);
    v.single_g = str2num(g_str);
    if (v.single_c ~= 0)
        v.single_gwc = v.single_g / (2e6*pi*v.single_c); % instrument uses
                                                          % 1MHz small signal
        temp = floor(v.single_gwc * 100000);
        v.single_gwc = temp / 100000;
    else
        v.single_gwc = -1;
    end
end
```

go_sweep.m

```
function go_sweep

global v vh gh vcg

axes(vh.axes2)

if (v.signal_level == 30) % set signal level
    fprintf(gh, 'SL2')
else
    fprintf(gh, 'SL1')
end

if (v.meas_speed == 'fast') % set measurement speed
    fprintf(gh, 'MS1')
```

```
elseif (v.meas_speed == 'medi')
    fprintf(gh, 'MS2')
else
    fprintf(gh, 'MS3')
end

pause(0.1)
fprintf(gh, 'TR3') % trigger mode = hold/manual
fprintf(gh, 'IB2') % voltage source mode = single-staircase sweep mode
PSnum = sprintf('%g', v.v_start);
PPnum = sprintf('%g', v.v_stop);
PEnum = sprintf('%g', abs(v.v_step));
PLnum = sprintf('%g', v.t_hold/1000);
PDnum = sprintf('%g', v.t_step_delay/1000);
PSstring = strcat('PS', PSnum);
PPstring = strcat('PP', PPnum);
PEstring = strcat('PE', PEnum);
PLstring = strcat('PL', PLnum);
PDstring = strcat('PD', PDnum);
fprintf(gh, PSstring) % set start voltage
fprintf(gh, PPstring) % set stop voltage
fprintf(gh, PEstring) % set step voltage
fprintf(gh, PLstring) % set hold time
fprintf(gh, PDstring) % set setup delay time
current_v = v.v_start;
meas_num = 1;
vcg = 1; % reset data matrix
fprintf(gh, 'BC') % clear output buffer
fprintf(gh, 'SW1') % start measurement
%fprintf(gh, 'EX') % trigger the measurement

while ( (current_v ~= v.v_stop) & (v.data_to_save > -1) )
    r = fscanf(gh); % a string that contains the results
    rl = length(r);
    comma1_pos = 0;
    comma2_pos = 0;
    for n = 1:rl
        if ( (r(n) == ',') & (comma1_pos == 0) )
            comma1_pos = n;
        elseif (r(n) == ',')
            comma2_pos = n;
        end
    end
    if ( (comma1_pos == 0) | (comma1_pos == rl) | ...
            (comma2_pos == 0) | (comma2_pos == rl) )
        fprintf(gh, 'SW0') % stop measurement
        fprintf(gh, 'TR1') % trigger mode = internal
        v.data_to_save = -1;
        gpib_error_massage
        return
    else
        c_str = r(4:comma1_pos-1);
```

```
        g_str = r(comma1_pos+4:comma2_pos-1);
        v_str = r(comma2_pos+2:rl);
        current_v = str2num(v_str);
        vcg(meas_num, 1) = current_v;
        vcg(meas_num, 2) = str2num(c_str);
        vcg(meas_num, 3) = str2num(g_str);
    end

    % plot the C-V data on window
    plot(vcg(:,1), vcg(:,2))
    if (v.v_start < v.v_stop)
        xmin = v.v_start;
        xmax = v.v_stop;
    else
        xmin = v.v_stop;
        xmax = v.v_start;
    end
    ymin = min(vcg(:,2));
    ymax = max(vcg(:,2));
    if (ymax == ymin)
        ymax = ymin + 1e-14;
    end
    axis([xmin xmax ymin ymax])

    % display numeric data in the "Single" zone of the window
    v.single_c = str2num(c_str);
    v.single_g = str2num(g_str);
    if (v.single_c ~= 0)
        v.single_gwc = v.single_g / (2e6*pi*v.single_c); % instrument uses
                                                         % 1MHz small signal
        temp = floor(v.single_gwc * 100000);
        v.single_gwc = temp / 100000;
    else
        v.single_gwc = -1;
    end
    if ((v.single_c == -1) | (v.single_g == -1) | (v.single_gwc == -1))
        c_str_disp = 'N/A';
        g_str_disp = 'N/A';
        gwc_str_disp = 'N/A';
    else
        c_str_disp = sprintf('%g', v.single_c * 1e12);
        g_str_disp = sprintf('%g', v.single_g * 1e6);
        gwc_str_disp = sprintf('%g', v.single_gwc);
    end
    set(vh.single_c, 'string', c_str_disp);
    set(vh.single_g, 'string', g_str_disp);
    set(vh.single_gwc, 'string', gwc_str_disp);

    drawnow
    meas_num = meas_num + 1;
end
```

```
v.data_to_save = v.data_to_save + 1;
```

go_zero.m

```
function go_zero

global v vh gh vcg

fprintf(gh, 'TR3') % trigger mode = manual

if (v.cable_length == 0) % set cable length
    fprintf(gh, 'LE1')
elseif (v.cable_length == 5)
    fprintf(gh, 'LE3')
    pause(0.1)
    fprintf(gh, 'CA')
    pause(2)
else
    fprintf(gh, 'LE2')
end

if (v.signal_level == 30) % set signal level
    fprintf(gh, 'SL2')
else
    fprintf(gh, 'SL1')
end

if (v.meas_speed == 'fast') % set measurement speed
    fprintf(gh, 'MS1')
elseif (v.meas_speed == 'medi')
    fprintf(gh, 'MS2')
else
    fprintf(gh, 'MS3')
end

pause(0.1)
zo_cmd = sprintf('ZO\n');
fprintf(gh, zo_cmd, 'async') % zero open
a = gh.TransferStatus;
while (a(1:4) ~= 'idle')
    a = gh.TransferStatus;
end
pause(1)

fprintf(gh, 'TR1')

go_correct_en

v.calibrated = 1;
```

gpib_error.m

```
function varargout = gpib_error(varargin)

% Begin initialization code - DO NOT EDIT
gui_Singleton = 1;
gui_State = struct('gui_Name',       mfilename, ...
                   'gui_Singleton',  gui_Singleton, ...
                   'gui_OpeningFcn', @gpib_error_OpeningFcn, ...
                   'gui_OutputFcn',  @gpib_error_OutputFcn, ...
                   'gui_LayoutFcn',  [] , ...
                   'gui_Callback',   []);
if nargin & isstr(varargin{1})
    gui_State.gui_Callback = str2func(varargin{1});
end

if nargout
    [varargout{1:nargout}] = gui_mainfcn(gui_State, varargin{:});
else
    gui_mainfcn(gui_State, varargin{:});
end
% End initialization code - DO NOT EDIT

function gpib_error_OpeningFcn(hObject, eventdata, handles, varargin)
handles.output = hObject;
guidata(hObject, handles);

function varargout = gpib_error_OutputFcn(hObject, eventdata, handles)
varargout{1} = handles.output;

% --- Executes on button press in pushbutton1.
function pushbutton1_Callback(hObject, eventdata, handles)
```

gpib_error_message.m

```
function gpib_error_message

global v vh3 gh

hfig = openfig('gpib_error', 'reuse');
set(hfig, 'handlevisibility', 'on', 'doublebuffer', 'on');
vh3 = guihandles(hfig);
```

save_error.m

```
function varargout = save_error(varargin)

% Begin initialization code - DO NOT EDIT
gui_Singleton = 1;
gui_State = struct('gui_Name',       mfilename, ...
                   'gui_Singleton',  gui_Singleton, ...
                   'gui_OpeningFcn', @save_error_OpeningFcn, ...
```

```
                   'gui_OutputFcn',  @save_error_OutputFcn, ...
                   'gui_LayoutFcn',  [] , ...
                   'gui_Callback',   []);
if nargin & isstr(varargin{1})
    gui_State.gui_Callback = str2func(varargin{1});
end

if nargout
    [varargout{1:nargout}] = gui_mainfcn(gui_State, varargin{:});
else
    gui_mainfcn(gui_State, varargin{:});
end
% End initialization code - DO NOT EDIT


function save_error_OpeningFcn(hObject, eventdata, handles, varargin)
handles.output = hObject;
guidata(hObject, handles);

function varargout = save_error_OutputFcn(hObject, eventdata, handles)
varargout{1} = handles.output;

% --- Executes on button press in pushbutton1.
function pushbutton1_Callback(hObject, eventdata, handles)
```

save_error_message.m

```
function save_error_message

global v vh4 gh

hfig = openfig('save_error', 'reuse');
set(hfig, 'handlevisibility', 'on', 'doublebuffer', 'on') ;
vh4 = guihandles(hfig);
```

save_ok.m

```
function varargout = save_ok(varargin)

% Begin initialization code - DO NOT EDIT
gui_Singleton = 1;
gui_State = struct('gui_Name',       mfilename, ...
                   'gui_Singleton',  gui_Singleton, ...
                   'gui_OpeningFcn', @save_ok_OpeningFcn, ...
                   'gui_OutputFcn',  @save_ok_OutputFcn, ...
                   'gui_LayoutFcn',  [] , ...
                   'gui_Callback',   []);
if nargin & isstr(varargin{1})
    gui_State.gui_Callback = str2func(varargin{1});
end
```

```
if nargout
    [varargout{1:nargout}] = gui_mainfcn(gui_State, varargin{:});
else
    gui_mainfcn(gui_State, varargin{:});
end
% End initialization code - DO NOT EDIT

function save_ok_OpeningFcn(hObject, eventdata, handles, varargin)
handles.output = hObject;
guidata(hObject, handles);

function varargout = save_ok_OutputFcn(hObject, eventdata, handles)
varargout{1} = handles.output;

% --- Executes on button press in pushbutton1.
function pushbutton1_Callback(hObject, eventdata, handles)
```

save_ok_message.m

```
function save_ok_message

global v vh2 gh

full_path_cv = strcat('C:\datafile\', v.save_dir, '\', v.save_filename, '_cv.txt');
full_path_gv = strcat('C:\datafile\', v.save_dir, '\', v.save_filename, '_gv.txt');

hfig = openfig('save_ok', 'reuse'); %, 'visible');
set(hfig, 'handlevisibility', 'on', 'doublebuffer', 'on')
vh2 = guihandles(hfig);

set(vh2.cv_filename, 'string', full_path_cv);
set(vh2.gv_filename, 'string', full_path_gv);
drawnow
```

stop_sweep.m

```
function stop_sweep

global v vh gh vcg

fprintf(gh, 'SW0') % stop measurement
fprintf(gh, 'TR1') % trigger mode = internal
v.data_to_save = -1;
```

Appendix B

Miller Effect

Bandwidth of common emitter stages is limited by Miller effect. A cascode topology is used in this work to eliminate the effect (sections 4.2.1 and 5.2.1). This appendix outlines the Miller effect.

B.1 Summary of Miller Effect

Miller effect was discovered by John M. Miller on 1919 during his research on vacuum tubes and published in 1920 [1]. The outcome of Miller effect is related to input and output impedances of an amplifier with feedback network. Consider the circuit of figure B.1. The circuit consists of a voltage amplifier with voltage gain of a_v, and a negative feedback network consisting on an impedance that shunts the amplifier's input and output, denoted by Z_f. The input impedance of the circuit is given by

$$Z_{in} = Z_f \cdot \frac{1}{a_v + 1} \parallel z_{in.a} \approx \frac{Z_f}{a_v} \tag{B.1}$$

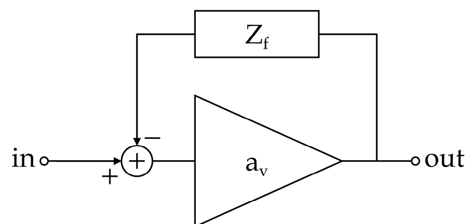


Figure B.1: Voltage amplifier with feedback network

where $z_{in.a}$ is the input impedance of the amplifier. The approximation in (B.1) is valid when $a_v \gg 1$ and $z_{in.a} \longrightarrow \infty$. The output impedance is

$$Z_{out} = Z_f \cdot \frac{a_v}{a_v - 1} \parallel z_{out.a} \approx Z_f \parallel z_{out.a} \tag{B.2}$$

where $z_{out.a}$ is the output impedance of the amplifier. Two important cases should be discussed: (a) capacitive feedback and (b) resistive feedback. If the feedback network is a capacitor (denoted C_f), it can be replaced by two capacitors: $C_{in} = C_f a_v$ and $C_{out} = C_f$, as illustrated in figure B.2a. The outcome of Miller effect in this case is the multiplica-

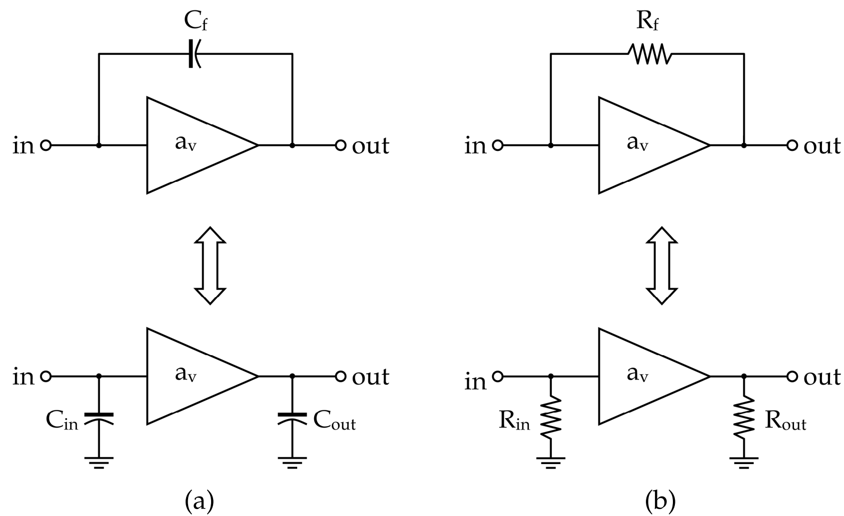


Figure B.2: Miller effect with (a) capacitive and (b) resistive feedback networks

tion of C_f by a_v, introducing a very large capacitance at the input. This capacitance is the main limit of the bandwidth of a common emitter stage, since C_μ acts as a feedback capacitor (see figure B.3), and the stage has significant voltage gain [2]. A resistive feedback network, denoted R_f (shown in figure B.2b), can be replaced by two resistors: $R_{in} = R_f / a_v$ and $R_{out} = R_f$. The meaning of Miller effect here is that the closed loop amplifier has a very low input resistance. Actually, Miller effect is the phenomenon that makes an opamp with PIPO feedback network to act as a transimpedance amplifier [2] (refer to section 4.1.2).

Finally, an important note should be indicated: calculations based on Miller effect are valid only through the first pole of the voltage amplifier. At higher frequencies the voltage gain drops, and Miller effect changes as well [3].

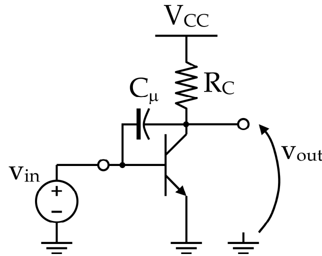


Figure B.3: C_μ as a feedback capacitor in common emitter stage

References for Appendix B

[1] John M. Miller, "Dependence of the input impedance of a three-electrode vacuum tube upon the load in the plate circuit", *Scientifc Papers of the Bureau of Standards*, vol. 15, no. 351, pp. 367–385, 1920.

[2] Paul R. Gray and Robert G. Meyer, *Analysis and Design of Analog Integrated Circuits*, John Wiley & Sons, third edition, 1993.

[3] Behzad Razavi, *Design of Analog CMOS Integrated Circuits*, McGraw-Hill, 2001.

Hebrew Section

עד למדידות – חשמליות ואופטיות-אלקטרוניות – ולתוצאותיהן. הפרק החמישי דן במגברים המפולגים – עקרונות הפעולה שלהם, תכן המגברים המפולגים, האינטגרציה האופטית-אלקטרונית, מדידת המעגלים וביצועיהם. את העבודה חותם הפרק השישי, שמסכם את תוצאות המחקר, מציג מסקנות, ומעלה אפשרויות וכיוונים להמשך המחקר בתחום.

שני נספחים מופיעים בסוף העבודה: האחד מציג תוכנית MATLAB עם ממשק משתמש גרפי, לביצוע מדידות מתח-קיבול בעזרת מכשיר HP 4280A ותקשורת GPIB. הנספח השני סוקר בקיצור נמרץ את אפקט מילר, בשל חשיבותו הרבה לתכן המעגלים בעבודה זו.

אלקטרוני של 14.2 גיגה-הרץ, ולהגבר של 61.2 וולט/וואט. בפיסות החדשות, הנמצאות כעת בייצור, אנו מצפים למעגלים בעלי רוחב סרט של יותר מ-20 גיגה-הרץ, ולהגבר אופטי-אלקטרוני של כ-120 וולט/וואט (עם ציפוי מונע החזרות). ביצועים אלו יהיו שיא של מכפלת הגבר-רוחב סרט במעגלים מסוג זה.

למגברים מפולגים, המכונים גם מגברי גל מתקדם (TWAs), רוחב סרט גדול בהשוואה למגברים מקובצים – רוחב הסרט נושק לתדר הקיטעון להגבר הזרם, f_T. מאידך, הם צורכים יותר הספק, משתרעים על פני שטח גדול יותר, בעלי ליניאריות נמוכה, ותגובת התדר שלהם גלית. בעבודה זו התגובה הגלית נוצלה לטובת הגדלת רוחב הסרט האופטי-אלקטרוני. טכנולוגיית הטרנזיסטורים הביפולריים מעורבי-הצומת של הטכניון הוכחה כמתאימה לייצור מעגלים מפולגים רחבי סרט, בצורה של מגברי הספק. מגברים כאלה, שתוכננו בעבר על ידי מר עמנואל כהן, השיגו רוחב סרט של 75 גיגה-הרץ, ושיא ביחס בין רוחב הסרט לתדר הקיטעון להגבר הזרם: 0.92. בעבודה זו, בחלק העוסק במגברים מפולגים, המגברים תוכננו מחדש כך שיתפקדו כמגבר טרנסאימפדנס, בעל אימפדנס כניסה נמוך ככל האפשר. דגש הושם על הגדלת התפוקה (yield), ונדרשו תכן חדש ומסכות חדשות. המגברים מבוססים על ארכיטקטורת קיזוז הנחתה (attenuation compensation). רוחב הסרט האופטי-אלקטרוני שהושג הוא 20.2 גיגה-הרץ, ובעזרת השכבות המתוקנות אנו מצפים להגיע ל-48 גיגה-הרץ. ההגבר האופטי-אלקטרוני עומד על 26.3 וולט/וואט, ובעזרת ציפוי מונע החזרות ניתן להגדילו לכ-50 וולט/וואט. מחקרים שנעשו בתחום עד כה עשו שימוש בטרנזיסטורי HEMT ובפוטודיודות מסוג גל-בו, הדורשים תהליך ייצור מסובך וזיווד מורכב ויקר. המעגלים שתוכננו בעבודה זו מהווים אפוא הפתרון הראשון מסוגו למקלטים מונוליתיים פשוטים וזולים, המיועדים לתקשורת אופטית בקצב של 40 גיגה-ביט לשניה, ואף יותר מכך.

בעבודה ששה פרקים: הפרק הראשון מהווה מבוא, ובו הקדמה על כל הנושאים הנדונים במחקר, סקר ספרות, והסבר על מטרות המחקר. הפרק השני עוסק בטרנזיסטור הביפולרי מעורב-הצומת – החל בתופעות הפיסיקליות שמשפיעות באופן ישיר על תכנון המעגלים, תהליך ייצור הטרנזיסטורים, מדידות הנעשות על הטרנזיסטורים, אפיון הטרנזיסטורים, ובניית מודלים – לאות קטן ולאות גדול – של הטרנזיסטור לצורך סימולציית מחשב. את הפרק חותמת סקירה קצרה על מנגנוני הרעש בטרנזיסטור ביפולרי מעורב-צומת. הפרק השלישי עוסק בפוטודיודה – תכונותיה הפיסיקליות, תהליך הייצור שלה, מדידת הפוטודיודה ואפיונה, ומנגנוני הרעש בדיודה. בפרק הרביעי נדון תכן המעגלים המקובצים – מאבני הבניין המשמשות לתכנון מעגל, דרך סקירה נקיפה על הארכיטקטורה שנבחרה למגבר, סקירת ביצועי הרעש וחסינות מפני רעשי ספקים, האינטגרציה האופטית-אלקטרונית,

גיגה-הרץ. בטכנולוגיית הייצור של הטכניון ניתן לייצר טרנזיסטורים בעלי f_T של יותר מ-200 גיגה-הרץ, כאשר השכבות מתוכננות עבור טרנזיסטור אופטימלי, ללא התחשבות בפוטודיודות.

לחומר GaInAs, כאשר הוא מתואם גביש ל-InP, פער אנרגיה של 0.75 eV, המתאים לקליטה של גלים באורך גל של 1.55 מיקרון. עובדה זו הופכת את מערכת החומרים InP/GaInAs, ובפרט את טרנזיסטורי ה-HBT המיוצרים מחומרים אלו, לאטרקטיבית עבור ייצור יישומים אופטיים-אלקטרוניים מונוליתיים. כדי לנצל את יתרונות טרנזיסטורי ה-HBT, ואת קלות ייצור הפוטודיודות מאותן השכבות, הדיודות בהן נעשה שימוש בעבודה זו הן מסוג הארה מלמעלה (top illuminated). דיודות אלה קלות לייצור, אך איטיות יחסית לסוגים אחרים של פוטודיודות, ובמיוחד יחסית לדיודות גל-בו. כמו בטרנזיסטורים, עובי הפוטודיודות שייוצרו הוא 6000 Å או 8000 Å, כדי להבטיח תגובתיות טובה ונצילות קוונטית טובה. לא ייושם ציפוי מונע החזרות על הפוטודיודות, אך ניתן ליישמו בעתיד. בעיה בגידול השכבות, שנתגלתה בעת מדידת הדיודות, מקטינה את רוחב הסרט של הדיודות. הבעיה תוקנה לאחרונה, וכעת מיוצרים מעגלים חדשים, בעלי פוטודיודות מהירות יותר.

קצה קדמי של מקלטים אופטיים מורכבים מפוטודיודה וממגבר. האור הפוגע בפוטודיודה מומר על ידה לזרם חשמלי, שמוגבר על ידי המגבר ומומר למתח חשמלי. מכל מקום, החיבור הפיסי בין הפוטודיודה למגבר מהווה מכשול משמעותי בפני הביצועים בתדר גבוה. כאשר הפוטודיודה והמגבר מיוצרים בנפרד, הם מקובעים על תושבת, ומחוברים ביניהם על ידי חוטי קישור (העשויים על פי רוב מזהב). לחוטי הקישור השראות גדולה (מסדר גודל של ננו-הנרי בודדים), שפוגעת קשות בהתנהגות המעגל בתדרים גבוהים. בנוסף, ייצור שתי יחידות נפרדות וזיווגן מגדיל את עלויות הייצור באופן משמעותי. כדי להתגבר על הבעיות דלעיל, מייצרים מעגלים מונוליתיים. במעגל מונוליתי הפוטודיודה והמגבר מייצרים על אותה משבצת מוליך למחצה (die), וניתן לתכנן את החיבור ביניהם באופן אופטימלי.

מעגלים מקובצים ניחנים בביצועים יציבים, היענות תדר שטוחה, ליניאריות גבוהה, צריכת הספק נמוכה ושטח קטן. חסרונם העקרוני הוא רוחב הסרט שלהם, המוגבל לכשליש מתדר הקיטעון להגבר הזרם, f_T. עקב היותם איטיים יחסית, מגברים מקובצים מתאימים, באופן טבעי, לפוטודיודות מוארות מלמעלה, בפרט כאשר עוסקים בטרנזיסטורים ביפולריים מעורבי-צומת. המגברים המקובצים שתוכננו בעבודה זו מבוססים על תכן קודם (של ד״ר בני שיינמן) עם שינויים מקיפים, במטרה להגדיל את רוחב הסרט ולהשיג ביצועים יציבים יותר ועקביים יותר. הסכמאות, כמו-גם המסכות, תוכננו מחדש מן ההתחלה. המגברים מגיעים לרוחב סרט אופטי-

תקציר

טרנזיסטורים ביפולריים מעורבי-צומת (HBTs), המבוססים על מערכת החומרים של אינדיום פוספיד (InP), מצטיינים בתדרי קיטעון גבוהים, מעבר ל-200 גיגה-הרץ. אותה מערכת החומרים משמשת לייצור פוטודיודות עבור תקשורת אופטית באורך גל של 1.55 מיקרון. לכן, לטרנזיסטורים ביפולריים מעורבי-צומת המבוססים על אינדיום פוספיד תפקיד חשוב בתקשורת סיבים אופטיים. עבודה זו מציגה תכן של מעגלים משולבים אופטיים-אלקטרוניים, מבוססי InP, למערכות תקשורת אופטית.

המחקר מתמקד בקצה קדמי (front end) של מקלטים אופטיים המבוססים על פוטודיודה ומגבר טרנסאימפדנס. תוכננו שני סוגים של מגברים: מקובצים ומפולגים. המעגלים המקובצים ניחנים בשטח קטן, צריכת הספק נמוכה ורעש נמוך. לעומתם, למגברים המקובצים רוחב סרט גדול יותר, תוך שימוש באותם הטרנזיסטורים.

המחקר בעולם בתחום טרנזיסטורים ביפולריים מעורבי צומת, העשויים מהחומרים InP ו-GaInAs מתרכז בהשגת שתי מטרות: (א) מעגלים אופטיים-אלקטרוניים לתקשורת סיבים אופטיים באורך גל של 1.55 מיקרון; (ב) מעגלי RF ומעגלי המרת מידע (data conversion) מהירים ודלי רעש. טרנזיסטורי SHBT, בעלי צומת מעורב אחד בלבד, מאפשרים לייצר פוטודיודות מאותן השכבות המשמשות לבסיס ולקולקטור של הטרנזיסטורים. לכן, טרנזיסטורי SHBT משמשים במעגלים אופטיים-אלקטרוניים. ביישומים אחרים ניתן להשתמש בטרנזיסטורי DHBT, בעלי שני צמתים מעורבים, שמתח הפריצה שלהם גבוה יותר. שתי תכונות של הטרנזיסטור קובעות את ביצועיו בתדר גבוה: f_T, תדר הקיטעון להגבר הזרם, ו-f_{MAX}, תדר הקיטעון להספק היונילטרלי. שני התדרים נקבעים על ידי הקיבולים שבטרנזיסטור, ההתנגדויות שלו וזמן המעבר של נושאי מטען בקולקטור. הטרנזיסטורים, בהם נעשה שימוש בעבודה זו, הם טרנזיסטורי SHBT עם קולקטור בעובי 6000 Å או 8000 Å. קולקטור כה עבה הכרחי עבור הפוטודיודות, המיוצרות מאותן השכבות, אך טומן בחובו זמן מעבר ארוך, וכתוצאה – תדרי קיטעון נמוכים. לטרנזיסטורי ה- 8000 Å תדר f_T של 67 גיגה-הרץ, ולטרנזיסטורי ה- 6000 Å f_T של 90

5.2 ערכים שנעשה בהם שימוש לחישוב מאפייני CPW 150
5.3 ערכים בפועל של קווי תמסורת במגברי TWA 156
5.4 NEP של מעגלים אופטיים-אלקטרוניים מפולגים 161
5.5 ביצועים חשמליים מסומלצים ומדודים של מגברי TWA 164
5.6 ביצועים אופטיים-אלקטרוניים מסומלצים ומדודים של מגברי TWA 165

6.1 השוואת עבודה זו לעבודות שנעשו לאחרונה על מקלטים אופטיים מקובצים מבוססי InP 169
6.2 השוואת עבודה זו לעבודות שנעשו לאחרונה על מקלטים אופטיים מפולגים מבוססי InP 170

רשימת טבלאות

1.1 עבודות שנעשו לאחרונה על מקלטים אופטיים מקובצים מבוססי InP 18
1.2 עבודות שנעשו לאחרונה על מקלטים אופטיים מפולגים מבוססי InP 19

2.1 מאפייני גידול שכבות 35
2.2 שלבי ייצור HBT 43
2.3 פרמטרי HBT מחולצים 69
2.4 המרת פרמטרים מדודים לפרמטרי VBIC 74
2.5 סיכום סוגי רעש 79

3.1 תכונות פיסיקליות של GaInAs ופרמטרים לפיהם חושב תדר ה-3dB של פוטודיודה 90
3.2 מאפייני גידול שכבות 92
3.3 שלבי ייצור פוטודיודה 96
3.4 תגובתיות ונצילות קוונטית חיצונית מדודות של פוטודיודות שונות 100
3.5 פרמטרי אות קטן מחולצים של פוטודיודות שונות 101

4.1 מונחי מודל כלאיים-π מבוטאים על ידי מונחי מודל T 109
4.2 הערכים המעשיים שנעשה בהם שימוש בתכנון TIA 126
4.3 נקודות עבודה מסומלצות של כל טרנזיסטורי ה-TIA 128
4.4 ביצועי רעש של TIA 130
4.5 NEP של מעגלים אופטיים-אלקטרוניים 137
4.6 ביצועים מסומלצים ומדודים של מגברי TIA 140
4.7 רוחב סרט אופטי-אלקטרוני של מגברי TIA עם פוטודיודות שונות 141

5.1 גדלים של קו תמסורת מבוטאים על ידי פרמטרי המודל של הקו 148

6.1 מדידות C-V של פיסת 6000 Å: (a) קיבול הצומת כנגד מתח הצומת (b) רוחב הצומת כנגד מתח הצומת 168

A.1 איור MATLAB של cv_main 174
A.2 איור MATLAB של gpib_error 174
A.3 איור MATLAB של save_error 175
A.4 איור MATLAB של save_ok 175

B.1 מגבר מתח עם רשת משוב 193
B.2 אפקט מילר עם (a) רשת משוב קיבולית ועם (b) רשת משוב התנגדותית 194
B.3 C_μ כקבל משוב בדרגת אמיטר משותף 195

4.8 סכמה של רשת משוב שלילי 116
4.9 סכמה של מגבר שרת עם משוב PIPO 118
4.10 הטופולוגיה של מגבר השרת 123
4.11 טופולוגיה של TIA בחוג סגור (*גרסה B*) 124
4.12 TIA עם דרגת כניסה בסיס משותף (*גרסה A*) 125
4.13 TIA גרסה A והמסלול שלאורכו מחושב V_x 127
4.14 V_x כנגד β וכנגד R_{C3} 128
4.15 זרם רעש מסומלץ 130
4.16 PSRR מסומלץ 132
4.17 מתאר של מגברי TIA: (a) גרסה A (b) גרסה B 134
4.18 TIA גרסה B עם פוטודיודה בקוטר 15 מיקרון ונגד ממתוח: (a) סכמה (b) מתאר 136
4.19 מדידת מעגל 139
4.20 עקומה מותאמת לנתונים שנמדדו ממעגל משולב אופטי-אלקטרוני 139
4.21 ביצועים מסומלצים ומדודים של מגברי TIA שונים 141

5.1 חיבור חשמלי של קו תמסורת 146
5.2 מודלים של קו תמסורת (a) ללא הפסדים (b) ועם הפסדים 147
5.3 קו תמסורת עם אי-רציפות קיבולית מחזורית 149
5.4 מבנה של מוליך גלים קו-פלנרי 150
5.5 Z_0 כפונקציה של W ו-G 150
5.6 מבנה של TWA טיפוסי 151
5.7 שני מסלולי אות בעלי השהיות שונות 153
5.8 סכמה של דרגת הגבר 154
5.9 אימפדנס כניסה של TWA 155
5.10 סכמת ממתוח של TWA 157
5.11 זרם רעש מסומלץ של מגבר 4 דרגות ושל מגבר 5 דרגות 158
5.12 מתאר של מגברי TWA: (a) הגבהה חזקה, 4 דרגות, גרסה חשמלית (b) הגבהה בינונית, 5 דרגות, עם פוטודיודה (c) הגבהה חלשה, 4 דרגות, עם פוטודיודה 159
5.13 תגובת תדר מנורמלת של פוטודיודה, TWA ומעגל אופטי-אלקטרוני 160
5.14 הגבר טרנסאימפדנס מסומלץ ומדוד של מגבר TWA 163
5.15 השהיית חבורה מסומלצת ומדודה של מגבר TWA 163
5.16 הפסדי החזרת יציאה מסומלצים ומדודים 164

2.48 מודל גאמל-פון 73
2.49 מעגל סימולציית ADS לאימות מודל VBIC 75
2.50 השוואה בין פרמטרי S מסומלצים ופרמטרים מדודים 76
2.51 מעגל תמורה לאות קטן של צומת PN עם מקורות רעש 80
2.52 מקורות רעש ב-HBT 81

3.1 תהליך גילוי פוטון על ידי פוטודיודה 87
3.2 חתך של פוטודיודה מוארת מלמעלה 88
3.3 תדר 3dB- של פוטודיודה מ-GaInAs כתוצאה מזמזן המעבר כנגד רוחב הצומת 90
3.4 מודל לאות קטן של פוטודיודה 91
3.5 מבנה השכבות על הפיסה כפי שגודלו על ידי מערכת ה-MOMBE 92
3.6 שכבות הפיסה לאחר איכול מזת הדיודה 93
3.7 שכבות הפיסה לאחר מימוש המגעים 93
3.8 שכבות הפיסה לאחר מסכת בידוד 94
3.9 שכבות הפיסה לאחר מסכת כיסוי פולימיד 94
3.10 חתך של פוטודיודה מוגמרת עם ציפוי מונע החזרות 95
3.11 תמונות SEM של פוטודיודה בשלבי ייצור שונים: (a) נידוף מתכת המגעים והסרתה (b) כיסוי פולימיד – מבט עין-ציפור (c) כיסוי פולימיד – מבט על (d) נידוף מתכת ראשונה והסרתה 97
3.12 דיאגרמה של מערכת המדידה האופטית-אלקטרונית 98
3.13 תמונות של מערכת המדידה האופטית-אלקטרונית: (a) מבט כולל (b) פרובים אופטיים וחשמליים למדידת פוטודיודה 99
3.14 חילוץ C_j 101
3.15 מדידות C-V והנתונים המעובדים של שתי הפיסות 102
3.16 מקור רעש בפוטודיודה 104

4.1 חיבורי טרנזיסטור ביפולרי: (a) אמיטר משותף (b) בסיס משותף (c) קולקטור משותף 108
4.2 מודל כלאיים-π לאות קטן של טרנזיסטור ביפולרי 108
4.3 דרגת אמיטר משותף: (a) טופולוגית המעגל (b) ייצוג באות קטן 110
4.4 דרגת אמיטר משותף עם נגד ניוון: (a) טופולוגית המעגל (b) ייצוג באות קטן 111
4.5 דרגת בסיס משותף: (a) טופולוגית המעגל (b) ייצוג באות קטן 112
4.6 דרגת קולקטור משותף: (a) טופולוגית המעגל (b) ייצוג באות קטן 114
4.7 טופולוגית המעגל של דרגת קסקוד 115

2.19 תמונות SEM של טרנזיסטור בשלבי ייצור שונים: (a) איכול אמיטר (b) נידוף מתכת בסיס והסרתה (c) הגנת אמיטר (d) הגנת קולקטור (e) בידוד (f) חשיפת אמיטר 44

2.20 תמונות SEM של טרנזיסטור ושל מעגל בשלבי ייצור שונים: (a) מעבר בסיס-קולקטור (b) נידוף מתכת ראשונה והסרתה (c) קבל מוגמר (d) חיבורים מוגמרים עם הצטלבויות ונגד 45

2.21 מערך מדידות DC: (a) מערכת המדידה (b) פרובים עבודת יד 46

2.22 עקומת זרם-מתח של צומת PN 47

2.23 עקומות ממדידת אמיטר משותף 48

2.24 עקומות ממדידת בסיס משותף 49

2.25 עקומות ממדידת שרטוט גאמל 50

2.26 חילוץ I_0 ו-n מעקומת זרם-מתח של צומת PN 51

2.27 עקומות β_F ו-β_{AC} מחושבות מנתוני מדידת שרטוט גאמל 52

2.28 V_{TO} מסומן על עקומות ממדידת אמיטר משותף 52

2.29 V_{CEO} מסומן על עקומות (a) α_F (b) ו-I_B 53

2.30 אביזר בדיקה ל-TLM 54

2.31 מסלול הזרם במדידת TLM 55

2.32 חילות התנגדות השכבה, התנגדות המגע ומרחק ההעברה מעקומת TLM 55

2.33 מערכת מדידת C-V 56

2.34 (a) עקומת C-V (b) ויחס המוליכויות 57

2.35 פרופיל הסימום של צומת PN חד צדדית 58

2.36 מערכת מדידת RF 59

2.37 טרנזיסטור כרשת זוגיים בחיבור אמיטר משותף 59

2.38 מודל T לאות קטן של HBT 61

2.39 HBT עם הרכיבים הפרזיטיים של רפידות המדידה כלולים ברשת זוגיים 62

2.40 מתאר של (a) רפידות מדידת HBT (b) אביזר מדידה פתוח (c) אביזר מדידה מפולש 62

2.41 חילוץ C_{BC} 64

2.42 חילוץ (a) L_e (b) $r_e + r_{ee}$ (c) r_b 65

2.43 חילוץ τ_D 66

2.44 חילוץ f_T 67

2.45 חילוץ f_{MAX} 68

2.46 מודל אברס-מול (גרסת הזרקה) 70

2.47 מודל אברס-מול (גרסת מעבר) 71

רשימת איורים

1.1 סכמה של מקלט בעל רכיבים נפרדים עם חוטי קישור 16
1.2 תגובת תדר של מקלט אופטי עם חוטי קישור באורך 0.55 מ״מ ו-0.70 מ״מ 17

2.1 (a) מבנה סכמתי חד-מימדי של BJT (b) ומסלול האלקטרונים במצב פעיל קדמי 25
2.2 ריכוז נושאי המטען לפי המקום לאורך BJT במצב פעיל קדמי 26
2.3 זרמים חשמליים מסומנים על סימן סכמתי של BJT 27
2.4 דיאגרמת פסי אנרגיה של (a) BJT בשיווי משקל (b) ובמצב פעיל קדמי, (c) ו-HBT בשיווי משקל (d) ובמצב פעיל קדמי 28
2.5 דיאגרמת פסי אנרגיה של (a) HBT מדורג-בסיס בשיווי משקל (b) ובמצב פעיל קדמי 29
2.6 מבנה סכמתי חד-מימדי של BJT עם תת-קולקטור 29
2.7 HBT חד-מימדי: (a) מבנה סכמתי (b) ורכיבי אות קטן 32
2.8 מבנה השכבות על הפיסה כפי שגודלו על ידי מערכת ה-MOMBE 34
2.9 שכבות הפיסה לאחר נידוף מתכת האמיטר והסרתה 36
2.10 שכבות הפיסה לאחר איכול האמיטר 37
2.11 שכבות הפיסה לאחר מסכת מתכת בסיס 37
2.12 שכבות הפיסה לאחר איכול הבסיס 38
2.13 שכבות הפיסה לאחר איכול הקולקטור 38
2.14 שכבות הפיסה לאחר נידוף מתכת הקולקטור 39
2.15 שכבות הפיסה לאחר איכול תת-הקולקטור 39
2.16 שכבות הפיסה לאחר פסיבציה של הטרנזיסטור 40
2.17 מתכת ראשונה כפי שמחוברת למגעי הטרנזיסטור 41
2.18 תהליך ייצור חיבורים וקבל: (a) מסכת מתכת ראשונה (b) מסכת קבל (c) מסכת הצטלבות (d) מסכת מתכת שניה 41

3.4 רעש בפוטודיודה 103
מקורות לפרק 3 105

4 מעגל אופטי-אלקטרוני מקובץ 107
4.1 אבני הבניין של מגברים 108
4.2 ארכיטקטורת מגבר טרנסאימפדנס 122
4.3 אינטגרציה אלקטרו-אופטית 135
4.4 מדידת המעגלים ואפיונם 138
4.5 ביצועי המעגלים 140
מקורות לפרק 4 143

5 מעגל אופטי-אלקטרוני מפולג 145
5.1 עקרונות של מגברים מפולגים 146
5.2 ארכיטקטורת מגבר מפולג 154
5.3 אינטגרציה אלקטרו-אופטית 160
5.4 ביצועי המעגלים 162
מקורות לפרק 5 166

6 סיכום ומסקנות 167
6.1 פוטודיודה 168
6.2 מעגלים מקובצים 169
6.3 מעגלים מפולגים 170
מקורות לפרק 6 171

A תוכנה בעלת ממשק גרפי לשליטה על מדידות C-V 173
A.1 מבוא 173
A.2 איורים של MATLAB 174
A.3 פונקציות של MATLAB 175

B אפקט מילר 193
B.1 תקציר של אפקט מילר 193
מקורות לנספח B 195

תוכן העניינים

תקציר **1**

רשימת קיצורים **3**

רשימת סימנים **5**

1 מבוא **13**

1.1 טרנזיסטור ביפולרי מעורב-צומת 14

1.2 פוטודיודה 15

1.3 מקלטים מונוליתיים 16

1.4 מקלט מונוליתי מקובץ 18

1.5 מקלט מונוליתי מפולג 19

מקורות לפרק 1 20

2 טרנזיסטור ביפולרי מעורב-צומת **23**

2.1 תכונות פיסיקליות של HBT ותופעות בו 24

2.2 ייצור HBT והמבנה שלו 34

2.3 מדידת ה-HBT ואפיונו 46

2.4 חילוץ מודל אות קטן של HBT 61

2.5 בניית מודל VBIC ל-HBT 70

2.6 רעש ב-HBT 77

מקורות לפרק 2 82

3 פוטודיודה **85**

3.1 תכונות פיסיקליות של פוטודיודה 86

3.2 ייצור פוטודיודה והמבנה שלה 91

3.3 מדידת הפוטודיודה ואפיונה 98

המחקר נעשה בהנחיית פרופ׳ דן ריטר בפקולטה להנדסת חשמל ובמרכז מחקר
למיקרואלקטרוניקה

תודות

ברצוני להודות לאנשים הרבים שסייעו לביצוע המחקר, ושבלעדיהם לא ניתן היה להוציאו לפועל:

1. תודה לטכניון, לפקולטה להנדסת חשמל ולמרכז מחקר למיקרואלקטרוניקה, שבמסגרתם נעשה המחקר, ושהעמידו לרשותי את מיטב הציוד והטכנולוגיות לצורך המחקר.
2. תודה לפרופ׳ דן ריטר על ההנחיה המסורה ועל התמיכה, העזרה והעידוד לכל אורך הדרך.
3. תודה לפרופ׳ גדי אייזנשטיין ולפרופ׳ מאיר אורנשטיין על העזרה והתמיכה בכל המדידות ועל העמדת מעבדותיהם לרשותי, על ציודן.
4. תודה למר אורי קרני ולמר יוחאי סבירסקי על הקמת מערכת המדידות האלקטרו-אופטיות, בהנחייתו של פרופ׳ גדי אייזנשטיין, ועל התמיכה הצמודה במהלך מדידות.
5. תודה לחבריי לקבוצת המחקר של פרופ׳ דן ריטר: ארקדי גברילוב, שמעון כהן, דורון כהן-אליאס, צופית מגריסו וגנאדי בורדו, על אלפי עזרות קטנות שלא יסולאו בפז.
6. תודה למשרד התעשייה, המסחר והתעסוקה של מדינת ישראל על מימון הפרוייקט, שבמסגרתו התבצע המחקר.
7. תודה להוריי, לאה ויהושע קראוס, ולחמותי וחמי, שרה ואלימלך וסטרייך, על התמיכה, העזרה והסיוע הכה-חיוניים לכל אורך הדרך.
8. לבסוף, תודה מיוחדת לאשתי, נעמה, ולילדיי – מתן, נועה ויובל – על הסבלנות וההבנה, ועל היותם משפחה נפלאה, שנתנה לי את כוחות הנפש לעשות את המחקר ולהגישו.

מעגל משולב אופטי-אלקטרוני, מבוסס InP, למערכות תקשורת אופטית

חיבור על מחקר

לשם מילוי חלקי של הדרישות לקבלת התואר מגיסטר למדעים
בהנדסת חשמל

שרגא קראוס

הוגש לסנט הטכניון - מכון טכנולוגי לישראל

אייר תשס״ו חיפה מאי 2006