High Performance Photodetectors for Multimode Optical Data Links

by

Wojciech Piotr Giziewicz

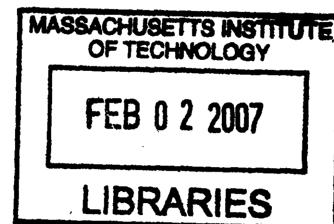
Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2006



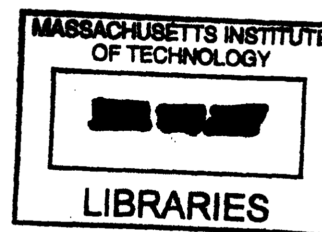
© Massachusetts Institute of Technology 2006. All rights reserved.

Author ....
Department of Electrical Engineering and Computer Science
August 9, 2006

Certified by ....
Lionel C. Kimerling
Thomas Lord Professor of Materials Science
Thesis Supervisor

Accepted by ....
Arthur C. Smith
Chairman, Department Committee on Graduate Students

ARCHIVES



High Performance Photodetectors for Multimode Optical Data Links

by

Wojciech Piotr Giziewicz

Submitted to the Department of Electrical Engineering and Computer Science
on August 9, 2006, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The majority of photodetectors presented in the literature, or available commercially, have dimensions on the order of 50 μm or smaller, suitable for glass multimode or single mode fibre applications. The recent successful commercialisation of very large core diameter plastic optical fibre in systems based around 650 nm emitters, as well as the recent emergence of new polymer materials enabling relatively low loss at the more standard 780 nm and 850 nm wavelengths, has exposed the need for integrated photodetectors with dimensions well above 100 μm and capable of bitrates from 250 Mb/s for low-cost consumer applications to multiple Gb/s for high performance short reach interconnects. This size-performance regime has been largely ignored until now. This work examines interdigitated detector structures in multiple material systems by measurement and simulation. An optoelectronic frequency response measurement system was designed and implemented for this work, allowing measurement up to 8 GHz using 850 nm or 1550 nm sources. The full expression for frequency response of diffusion current under different illumination scenarios was derived, a topic normally omitted in the discussion of photodetectors, and applied to the analysis of device measurements. Silicon detectors of various geometries were fabricated, with measured bandwidths at 5 V reverse bias up to 2 GHz for 200 μm diameter devices and 4 GHz for 50 and 100 μm diameter devices. The latter is the highest bandwidth reported for a silicon detector fabricated in a CMOS-compatible process and biased at a practically accessible voltage. Device performance was confirmed by simulation, and a novel structure is proposed featuring a buried junction on SOI determined by simulation to have twice as high a responsivity-bandwidth product as the best reported devices fabricated on high resistivity SOI. The silicon device structure was modified for epitaxial germanium wafers, and devices were fabricated. The germanium devices were simulated to determine the appropriate technology scaling direction and maximum device dimensions for desired performance specifications.

Thesis Supervisor: Lionel C. Kimerling
Title: Thomas Lord Professor of Materials Science

Acknowledgments

Although every thesis has only one formal author, a great many people contribute directly or indirectly to various stages of its development; from the formulation of the problem, followed by the conception of a possible solution, development of a method to reach the solution, experimental and theoretical verification, and (more often than not) several repetitions of this process. There are many people that deserve my great thanks for their contributions to these cycles.

First must come my thesis advisor Prof. Lionel "Kim" Kimerling and direct supervisor Dr. Jurgen Michel. They were able to identify a problem, bring together all of the elements necessary to tackle the problem (including but not limited to an interested sponsor, infrastructure, and a fantastic knowledge of the field), and then guide the project to a successful completion. I am very grateful to have had the chance to work with them and to learn from their experience, and I hope that I will be able to live up to the example that they have set. I would also like to thank Prof. Clifton Fonstad and Prof. Sheila Prasad for their guidance, leadership, and friendship. I learnt a great deal working with them, and that knowledge is certainly reflected in this work.

Dr. Mikihiko Kato of Fuji Photo Film Inc. sponsored this work, and I am greatly indebted to him for trusting us to do good work, and for providing us with all of the necessary resources to do so - probably even more than we expected. Dr. Kato and Yukiya Miyachi, also of Fuji Photo Film Inc, did an excellent job of keeping me on schedule and focused on the goal of the work, but also approached the problem in a fresh and open way that was very exciting.

Experimental work is also never really the work of one person alone, and this is particularly true in semiconductor processing. The staff of the Microsystems Technology Laboratories work extremely hard to keep all of the equipment working properly and consistently for many users, often with conflicting goals and agendas, and I would like to thank each and every one of them for their efforts and for putting up with my questions, complaints, and (at times of quickly approaching deadlines) outright demands. Likewise the staff of the Centre for Materials Science and Engineering should be thanked and congratulated for maintained a first-rate characterisation facility that was critical to this work.

I have had many co-workers over the years, and would like to thank all of them for their teaching, ideas, stimulating discussion, and help in getting important things done that I could not

manage by myself. It is difficult, and perhaps unfair, to single out a select few since all of their contributions have been so valuable. However, I will make three exceptions in this case and thank in particular Jifeng Liu for this tireless work on germanium epitaxy (including most importantly providing me with material necessary for my work!), Dr. Daniel Sparacin for proof-reading the draft of this document and for his wonderful organisation and help (or more precisely for allowing me to help) in making the Electronic Materials Research Group device characterisation laboratory a great place to do measurements, and last but not least Dr. Henry Choy for teaching me more than I could have imagined about epitaxy, material characterisation, device characterisation, and most importantly how to do good science. All of those hours spent scraping arsenic out of the MBE finally paid off.

Finally, and most importantly, I would like to thank my family and friends for their unwavering support over the past several years. Without the support of my parents in particular, it would have been impossible to swallow some big pills and make some tough decisions that led me to be able to write these acknowledgements today.

Contents

1 Introduction **25**

1.1 Areas of Interest for Optoelectronic Integration 26

1.2 Integrated or Si CMOS-compatible Photodetectors 30

1.2.1 Si Lateral Photodetectors on bulk substrates 30

1.2.2 Lateral Photodetectors on SOI substrates 32

1.2.3 Integrated Silicon Photodetectors involving epitaxy and process modifications 33

1.3 Germanium photodetectors 35

1.3.1 Vertical p-i-n photodetectors 36

1.3.2 Lateral p-i-n photodetectors 37

1.4 Thesis Goals and Organisation 38

2 Photodiode and Receiver Fundamentals **41**

2.1 Absorption 41

2.2 Basic Photodiodes 43

2.3 Small-Signal Model 46

2.3.1 Diode Small-Signal Model 47

2.3.2 Transit time dependence 49

2.3.3 Diffusion time dependence 52

2.3.4 Parasitic Elements 60

2.4 Receivers 63

2.4.1 Amplifier Stability 66

2.5 Noise 67

3 Photodetector Frequency Response Measurement 69

3.1 Impulse Response . . . 70

3.1.1 Impulse Response Examples . . . 73

3.1.2 Impulse Response Summary . . . 76

3.2 Eye Diagram Measurement . . . 77

3.2.1 Bit Error Rate . . . 78

3.2.2 Data encoding . . . 79

3.2.3 Eye Diagram Apparatus . . . 82

3.2.4 Eye Diagram Examples . . . 83

3.2.5 Eye Diagram Summary . . . 85

3.3 Frequency Domain Measurement . . . 86

3.3.1 Frequency Domain Measurement Apparatus . . . 87

3.3.2 Measurement Procedure . . . 88

3.3.3 Calibration and Operating Point Considerations . . . 90

3.3.4 Frequency Domain Measurement Examples . . . 99

3.3.5 Frequency Domain Measurement Summary . . . 101

4 Lateral Photodetector Design and Fabrication 103

4.1 Device Design . . . 103

4.1.1 Maximum possible bandwidth . . . 105

4.1.2 Realistic first-order bandwidth estimate . . . 107

4.2 Silicon Photodetector Fabrication Procedure . . . 110

4.2.1 Detailed Description of Process . . . 110

4.2.2 Silicon Device Process Comments . . . 112

4.3 Germanium Photodetector Fabrication Procedure . . . 118

4.3.1 Germanium growth and preparation . . . 118

4.3.2 Detailed Description of Process . . . 121

4.3.3 Germanium Device Process Comments . . . 123

5 Photodetector Measurement Results 133

5.1 Lateral p-i-n devices in a standard commercial GaAs VLSI process . . . 135

5.1.1 DC Electrical Characteristics 137
5.1.2 Capacitance 137
5.1.3 Sensitivity 139
5.1.4 AC Characteristics 140
5.1.5 GaAs Device Conclusions 141
5.2 CMOS-manufacturable Si lateral p-i-n photodetectors 142
5.2.1 DC Electrical Characteristics 142
5.2.2 Capacitance 149
5.2.3 Sensitivity 153
5.2.4 AC Characteristics 156
5.3 Germanium-on-Silicon lateral p-i-n photodetectors 166
5.3.1 DC Electrical Characteristics 166
5.3.2 AC Characteristics 167

6 Simulation and Discussion **173**
6.1 Depletion Region Shape, Capacitance 174
6.2 Isolation Junction Behaviour 180
6.3 Frequency Response 182
6.4 Novel device structure: buried junction on SOI 186
6.5 Germanium-on-silicon lateral p-i-n devices 189

7 Photodetector and System Design Summary **195**
7.1 Ideal system - Quantum Limit of Detection 197
7.2 Realistic System - Detector and Receiver Noise 198
7.3 Detector Design for Particular Applications 200
7.3.1 10 Gb/s over glass MMF 201
7.3.2 10 Gb/s over POF with 100 μm core diameter 203
7.3.3 1.5 Gb/s over POF with 200 μm core diameter 204
7.4 Detector and System Design Charts 205

8 Conclusions and Future Work **213**

8.1 Summary 213

8.2 Future Work 217

8.2.1 Process development 217

8.2.2 Device development 218

8.2.3 Further work outside the scope of photodiode devices 218

A Silicon and Germanium Wafer Cleaning Techniques **221**

B Silicon and Germanium Device Detailed Process Description **227**

B.1 Silicon Process 227

B.2 Germanium Process 229

References 233

List of Figures

1-1 32 channel transmitter/receiver module using flip-chip bonded VCSEL and photodiode arrays from [3] . . . 26
1-2 Photo of a Byteflight bidirectional transceiver from [7] . . . 28
1-3 Schematic drawing of a spatially modulated photodetector from [15] . . . 31
1-4 Schematic drawing of a silicon lateral trench photodetector from [16]. . . . 32
1-5 Schematic drawing of a vertical photodetector integrated in a BiCMOS process, from [25]. . . . 34
1-6 Schematic drawing of a lateral p-i-n photodetector fabricated from germanium on a thin SOI substrate, from [40]. . . . 38

2-1 Absorption coefficients for several materials as a function of wavelength, from [45] 42
2-2 Schematic representation at thermal equilibrium of charge density, electric field, and potential (relative to intrinsic material) for an abrupt p-n junction (left) and a p-i-n structure (right) . . . 44
2-3 Schematic representation of photodiode properties with and without illumination, indicating important device parameters (dark current, photocurrent, open circuit voltage). The device may be represented by the equivalent circuit on the right. . . . 45
2-4 Schematic representation of photocurrent: A) generation in p-region and diffusion of minority electrons to SCR, B) generation in SCR and flow of electron and hole by drift, C) generation in n-region and diffusion of minority holes to SCR. . . . 45
2-5 Small signal model of the photodiode including the incremental model of the diode. 50
2-6 Schematic representation of the one-dimensional carrier diffusion problem with uniform generation. . . . 54

2-7 Plot of J_{photo} according to equation 2.32 for an arbitrary G_L as a function of $\omega\tau_e$ for different ratios a/L_e. The 3-dB frequency is indicated by the dotted line. 56

2-8 Schematic representation of the one-dimensional carrier diffusion problem with varying generation. 56

2-9 Plots of expression 2.40 for varying $\omega\tau_e$ for various unitless ratios αL_e and a/L_e. The 3-dB levels is indicated by the dotted lines. The vertical scale of plot c is different than that of plots a and b. 58

2-10 Equivalent circuit elements of a p-i-n diode. The circuit of Figure 2-5 may be inserted as the i region equivalent. The parasitic elements of a different geometry (e.g. lateral/MSM) are identical but not as clearly presentable graphically. 60

2-11 Photodetector incremental circuit schematic including device parasitic elements. . . 61

2-12 Plots of the system function response (equation 2.42) for different series inductance L_s at the load, and typical values $C_j = 1$ pF, $C_{px} = 40$ fF, $R_s = 10\,\Omega$, $R_L = 50\,\Omega$ . . . 62

2-13 Full incremental photodiode circuit model 63

2-14 Basic receiver consisting of a load resistor and a voltage buffer. A bias current may be operating in parallel to adjust the DC offset of the output signal. 64

2-15 Transimpendance amplifier with a feedback resistor, decoupling the voltage across the diode and the photocurrent. 65

2-16 SPICE simulation example of a TIA circuit with a single-pole operational amplifier. As the photodiode capacitance is raised, the phase margin is decreased causing overshoot/ringing of the output voltage. 66

3-1 Schematic representation of an impulse response measurement system. The figure as drawn includes free-space coupling of a mode-locked laser beam, such as Ti:Sapphire operating around 850 nm. The same general principle applies to fibre coupling at the laser output, including then a fibre splitter rather than the beam splitter. 71

3-2 Vertical Ge-on-Si photodiode illuminated by 1040 nm 1 ps-pulse. Measured by the author and Jifeng Liu, reported in [59]. 74

3-3 Lateral GaAs photodiode fabricated in a standard VLSI process, illuminated by 850 nm sub-ps pulse. Measured by the author, reported in [66]. The device bandwidth was measured independently to be 2 GHz, and the pulse response is dominated by series inductance. 74
3-4 Lateral Ge-on-SOI photodiode reported by Dehlinger et al. [40] 75
3-5 Eye diagram of a 12.5 Gbps NRZ 2^{31} PRBS signal generated by an Agilent N4906B BERT and received by an Agilent 86100C sampling oscilloscope with a 50 GHz 86117A electrical plugin. The generated and receiver were connected by a 50 cm cable with 2.4 mm connectors. 78
3-6 24 bit sequence encoded in RZ and NRZ format. 80
3-7 FFT spectra of a data stream encoded in NRZ and RZ formats. The clock rate is 2 bps in this calculated example. 80
3-8 Simulated eye diagrams for a NRZ bitstream under increasingly severe filter corner frequencies. 81
3-9 Simulated eye diagrams for a RZ bitstream under increasingly severe filter corner frequencies. 82
3-10 Schematic representation of the eye diagram measurement system designed and assembled by the author. The switches represent manual replacement of a unit, cable, or fibre. Certain details have been omitted, in particular the polarisation controller between the 1550 nm laser and the external modulator, and the DC electrical connections to the external modulator. 83
3-11 Vertical Ge-on-Si detector illuminated by an 850 nm VCSEL driven with a 660 Mb/s NRZ PRBS, the minimum bitrate of the N4906B BERT. The detector bandwidth was independently measured to be below 400 MHz. The eye diagram exhibits a low Q with substantial amplifier noise. Some electrical reflection may be seen on the falling data edge. 84
3-12 Lateral Si detector illuminated by an 850 nm VCSEL driven with a 1.25 Gb/s NRZ PRBS, from [14] 84
3-13 Silicon ring modulator driven by 400 Mb/s NRZ and 1.5 Gb/s RZ PRBS signals with light output measured by a high speed photodetector and amplifier, from [73] 85

3-14 Schematic representation of the design of the frequency domain measurement apparatus . . . 86

3-15 Detailed schematic diagram of the frequency domain measurement apparatus. The switches as shown in the diagram are implemented as manual fibre or cable connections, but may also be constructed as semi and fully automated electrical and optical switches/splitters. . . . 89

3-16 Frequency domain measurement apparatus (Figure 3-15) showing calibration reference planes. The elements of the apparatus that are corrected for in the calibration are in faint grey. The cable between the network analyser and either the reference photodetector or the output of the DUT (or amplifier) is physically the same cable. . 91

3-17 Manufacturer calibration/verification data for the bias T and microwave probe used in the measurement apparatus . . . 92

3-18 Photograph of the AOC 10 Gbps VCSEL and connector, as well as of the completed source package connected to a bias T and network analyser. . . . 94

3-19 Measured S11 (reflection) and S12 (transmission) of the packaged VCSEL. There are two reflections clearly visible in both traces at multiples of approximately 1.8 GHz and 2.5 GHz. These interference peaks correspond to a reflection at the cable-PCB junction (smaller mismatch/reflection) and the PCB-VCSEL connector/jack junction (larger mismatch/reflection). The S21 trace is very low due to the coupling loss between VCSEL-coupled MMF and reference detector-coupled SMF which results in approximately 40-50 dB of photocurrent loss). . . . 94

3-20 Leakage signal for a large and small sample size, as seen in the inset photos. The smaller sample reduces leakage by up to 20 dB in the 150-400 MHz range, and by 10-15 dB in the 1-3 GHz range. In the measurement setup, the DUT is probed while the VCSEL is modulated and coupled to the optical multimeter. . . . 96

3-21 Frequency response due to standing waves in the electrical transmission line. . . . 96

3-22 Schematic diagram of the Darlington gain block and the bias configuration of the class A amplifier including blocking capacitors that determine the low-frequency bandwidth limit of the amplifier. . . . 98

3-23 Frequency domain trace of a Ge-on-Si vertical photodiode. The trace exhibits significant noise and interference/leakage. 99

3-24 Frequency domain trace of the same Ge-on-Si vertical photodiode as in Figure 3-23 (bottom trace). Noise and leakage have been eliminated. The traces follow curve fits to approximately the 8 GHz, the measurement limit of the apparatus with the 850 nm VCSEL source. 100

3-25 Frequency domain trace of a SiGe resonant cavity photodetector taken from [75] . . 100

4-1 Schematic representation of a simple vertical and lateral detector structures. The dimensions indicated on the lateral structure unit cell are inter-finger spacing s and finger width w. 104

4-2 Side and top schematic views of the silicon lateral p-i-n device structure, including some typical or expected width, spacings, and thickness measurements. 105

4-3 Estimated bandwidth and metal fill factor-bandwidth product for a silicon lateral p-i-n device assuming complete depletion of the semiconductor between the detector fingers, assuming 1 μm wide fingers. 106

4-4 Estimated bandwidth a bulk germanium lateral p-i-n device assuming complete depletion of the semiconductor between the detector fingers, assuming 1 μm wide fingers. 106

4-5 Schematic diagram of the device structure used to estimate the lateral p-i-n device bandwidth based on carrier diffusion and RC effect. 107

4-6 Estimated bandwidth of silicon lateral p-i-n devices with only partial active region depletion and carrier flow by diffusion in the QNR. All subfigures include a baseline curve corresponding to 10^{16} cm^{-3} doping, 1.75 μm inter-finger spacing, 1 μm finger width, and 5 V reverse bias. 108

4-7 Silicon lateral detector process flow schematic diagram indicating selected important steps. 111

4-8 TSUPREM4 simulation of doping profiles following implantation and annealing . . 114

4-9 Exposure test patterns showing the results of slight overexposure (distinguishable features: 0.45 μm in dark field, 0.5 μm in clear field) and significant underexposed (distinguishable features: 0.5 μm in dark field, 0.4 μm in clear field). 117
4-10 Germanium lateral detector process flow schematic diagram from growth to after the annealing step for different annealing methods. 119
4-11 (004) rocking curve of 700 nm as-grown epitaxial UHVCVD germanium on silicon . 120
4-12 (004) rocking curve of 700 nm cyclic annealed epitaxial UHVCVD germanium on silicon 121
4-13 Germanium lateral detector process flow schematic diagram starting with annealed wafers. 122
4-14 Optical microscope pictures of defects formed during the cyclic annealing process. . 125
4-15 Optical microscope photo of undercutting of germanium epilayer due to polysilicon etch misalignment. The undercut area is to the right of each finger. 126
4-16 Optical microscope photo of metal fingers that did not adhere to the polysilicon contacts, but remained fixed to the substrate by the wider metal bus. 127
4-17 SEM Images of Ge Lot 2 devices following blanket AlSi(2%) deposition and anneal (400 °C, 30 minutes). The images suggest the formation of large craters in, or general disappearance of, the metal film in and around the devices. 127
4-18 SEM Images of Ge Lot 2 devices following blanket AlSi(2%) deposition, anneal (400 °C, 30 minutes), and metal etch. 128
4-19 SEM Images of Ge Lot 2 devices following blanket AlSi(2%) deposition, anneal (400 °C, 30 minutes), metal etch, and BOE dip exposing the Ge surface. 129
4-20 SEM Images of Ge Lot 2 devices processed according to the original process flow, and following a BOE dip exposing the Ge surface. Small sections of metal remained on the polysilicon electrodes and in the field. 130
4-21 Al-Ge phase diagram from [89] 131
4-22 Polysilicon test pattern with square boron-implanted region following metal deposition, anneal, metal etch, and surrounding dielectric etch. Grains of different size and distribution are visible on the implanted and non-implanted areas of the polysilicon. 132

5-1 Schematic of the side and tops view of the lateral p-i-n device structure, including ion implant and ohmic contact structures but excluding interconnect metal. The width of the i region is defined photolithographically as the separation between the p and n implant regions. 136
5-2 Capacitance (circles) and Capacitance^{-2} (triangles) as a function of reverse bias for devices of similar layout, but fabricated in two different GaAs VLSI processes and with different i-region widths. 138
5-3 Schematic diagram of simulated depletion region behaviour at low (a) and higher (b) reverse bias. The dashed arrows indicate electric field direction (not magnitude). The dotted line represents the depletion region edge. 139
5-4 Photocurrent as a function of reverse bias and incident light power for a 0.9 μm device fabricated in the H-GaAs-V process. 140
5-5 AC photocurrent response as a function of frequency at different reverse bias levels. The fitting parameter "c" increases with reverse bias (going from a value of 1.2 at 0 V to 2.7 at 7V) 141
5-6 Si lot 1 DC device characteristics. 143
5-7 SEM pictures of lot 1 devices. 144
5-8 I-V properties of isolated devices. The three lines are the measurement of three devices of different dimensions (200 μm diameter, electrode spacings of 2-4 μm). The n resistor is a long narrow implantation region used to diagnose doping levels in the process. 145
5-9 I-V properties of lot 3 isolated devices. The three lines are the measurement of three devices of different dimensions (200 μm and 50 μm diameter, electrode spacings of 2-4 μm). 146
5-10 I-V properties of lot 4 isolated devices. The three lines are the measurement of three devices of different dimensions (200 μm and 100 μm diameter, electrode spacings of 2-4 μm). 146
5-11 Schematic diagrams representing the I-V characteristics as measured in lots 2-4, and the fabrication conditions that could lead to such a condition. 147

5-12 SEM pictures of misalignment of test patterns in the first and second mask sets used to fabricate the Si lateral p-i-n devices. 148

5-13 SEM pictures of n and p contact vias with pad metal. The roughness and dust-like particles are related to the metal deposition and anneal which left some residue on the field oxide surface. 149

5-14 C-V properties of isolated lot 3 devices. The three lines indicate different electrodes spacings as shown in the legends. 150

5-15 C-V properties of isolated lot 3 devices. The three lines indicate different electrodes spacings as shown in the legends. 151

5-16 C-V properties of isolated lot 4 devices. The different lines indicate different electrodes spacings as shown in the legends. 152

5-17 Lot 4, $1/C^2$ vs. reverse bias data, 200 μm diameter devices. 153

5-18 Photocurrent vs. reverse bias, normalised to account for electrode shadowing and surface reflections. The devices has 1 μm wide electrodes, with electrode spacing as shown in the figure. 153

5-19 Sensitivity vs. illumination for the isolated Si lot 2 devices. The inset shows the raw photocurrent data, demonstrating linear sensitivity over three orders of magnitude followed by an abrupt change in behaviour. The spot diameter was approximately 70 μm. 154

5-20 Sensitivity vs. illumination for the isolated Si lot 3 devices. The inset shows the raw photocurrent data, demonstrating linear sensitivity over two orders of magnitude followed by a strong change in behaviour. The spot diameter was approximately 70 μm. 155

5-21 Ebers-Moll model representation of the lateral p-i-n device as a phototransistor. The model may be simplified as shown assuming that the detector is in reverse bias. . . . 155

5-22 Schematic diagram of the load line operating point of the illuminated isolated devices. 156

5-23 Frequency response of the control and isolated devices at different average illumination powers as indicated in the figures (70 μm spot size). 157

5-24 Frequency response of the control and isolated devices at different average illumination powers as indicated in the figures (70 μm spot size). The devices geometry was 200 μm diameter, 2 μm electrode spacing. 158

5-25 Frequency response of isolated lot 3 devices at different reverse biases and low illumination (i.e. as in lowest curves of Figure 5-24). The measured data is shown in the dark lines, and the curve fit in the grey overlaid lines. 159

5-26 Extracted 3 dB frequencies and fitting coefficients of lot 3 devices with geometry as noted below and within each diagram. 159

5-27 Extracted 3 dB frequencies and fitting coefficients of 200 μm diameter lot 4 devices with electrode spacing as noted below and within each diagram. In subfigures a-d, the data is shown by the black lines and the curve fit is shown by the grey overlaid lines. 163

5-28 Extracted 3 dB frequencies and fitting coefficients of 200 μm diameter lot 4 devices with electrode spacing as noted below and within each diagram. The data is shown by black lines and the curve fit is shown by the grey overlaid lines. 165

5-29 IV characteristics of typical devices fabricated in Ge lot 1 and lot 2. 167

5-30 Typical frequency response measurements of Ge lot 1 devices. The measurement was performed using a Minicircuits ZKL-2R7 amplifier with a gain of 24 dB. 168

5-31 Photocurrent as a function of epilayer thickness for different reverse biases. The data is taken from measurement of 100 μm diameter, 4 μm finger spacing devices from three different wafers in the fabrication lot. 169

5-32 Schematic representation of the device measurement, with the DUT assumed to be a diode in parallel with a parasitic resistor. 170

5-33 Frequency response measurement of selected devices, and extracted bandwidth as a function of electrode spacing for constant device diameter and as a function of device diameter for constant electrode spacing. 171

6-1 Simulation of a lateral p-i-n structure assuming a bulk p-type substrate of different doping concentration. The dotted lines are the edges of the depletion region for reverse bias from 0 to 5 V. 175

6-2 Simulation of a lateral p-i-n structure assuming an isolated structure with different active region doping concentration and constant bulk doping concentration of 4×10^{15} cm^{-3}. The dotted lines are the edges of the depletion region for reverse bias from 0 to 5 V. 175

6-3 Simulation of bulk device capacitance as a function of voltage for different bulk doping levels. 176

6-4 Simulation of isolated device capacitance as a function of voltage for 2 μm finger spacing and a fixed substrate doping of 4×10^{15} cm^{-3} (a), a fixed active region doping of 1×10^{16} cm^{-3} (b), and fixed active region and substrate doping (1×10^{16} cm^{-3} and 4×10^{15}cm^{-3} respectively) (c). 178

6-5 Measured capacitance curves of 200 μm diameter, 2 μm finger spacing devices from Si lots 3 and 4 (dark lines) and simulated capacitance for similar structures (dotted lines). 179

6-6 Simulation of a lateral p-i-n structure assuming an isolated structure with different active region doping concentration and constant bulk doping concentration of 4×10^{15} cm^{-3}. The dotted lines are the edges of the depletion region at reverse bias from 0 to 5 V. 179

6-7 Simulation of device intensity dependence. The substrate equilibrium potential changes with illumination and the behaviour is strongly dependent on carrier lifetime. 181

6-8 Simulated effect of substrate current on bandwidth as a function of illumination and bias. 182

6-9 Measured (dark) and simulated (smooth overlay) characteristics of Si lot 3 and 4 devices. The device geometry in both cases is 200 μm diameter, 2 μm electrode spacing. 183

6-10 Simulation of device bandwidth sensitivity to various fabrication parameters. 184

6-11 Simulated SOI detectors based on the structure and results reported in [18], and including different finger spacings and doping levels. The three groups of devices based on doping level are offset for clarity. 185

6-12 New device structure with a simulation of the depletion region edge (dotted lines) as a function of reverse bias. 187

6-13 Simulation results of novel proposed buried junction SOI detectors. 188

6-14 Simulation of depletion region edge (dotted line) in the lightly p-type Ge epilayer for increasing reverse bias. The depletion of the entire inter-finger space occurs at approximately 1 V reverse bias for the n-type substrate case. 190

6-15 Simulation of lateral Ge-on-Si devices with p- epilayers (10^{15} cm^{-3}) on p-type Si substrates (10^{16} cm^{-3}) 191

6-16 Simulation of lateral Ge-on-Si devices with p- epilayers (10^{15} cm^{-3}) on n-type Si substrates (10^{16} cm^{-3}). Three different finger spacings exhibit different degrees of RC and drift time influence on bandwidth. 193

6-17 Simulation of lateral Ge-on-I devices with p- epilayers (10^{15} cm^{-3}). The finger spacing is 2 μm. These curves may be directly compared with the 2 μm finger spacing curves of Figure 6-16. 194

7-1 Quantum detection limits compared to maximum incident power for different wavelengths and efficiencies. 197

7-2 Detection limits of an ADN2821 TIA assuming $Q = 7$ and a 1 dB penalty for a typical extinction ratio. The dark lines are varying system external quantum efficiency curves as noted on the right of the figure. 199

7-3 Design space (clear area) for a 10 Gb/s optical link over glass MMF. 202

7-4 Design space (clear area) for a 10 Gb/s optical link over 100 μm diameter POF. . . . 203

7-5 Design space (clear area) for a 1.5 Gb/s optical link over 200 μm diameter POF using 650 nm VCSEL sources. 204

7-6 Design chart of lateral germanium and buried junction SOI devices. The germanium devices assume a 850 nm transmission system, and the SOI devices assume a 650 nm transmission system. A factor of 1 represents the class I limit, and a factor of 0.1 represents the class IM limit (also shown as a darker shade of grey). 207

7-7 Design chart for an 850 nm data transmission system including different photodetector and fibre geometries. A factor of 1 represents the class 1 limit, and a factor of 0.1 represents the class 1M limit (also shown as a darker shade of grey). 209
7-8 Design chart for an 650 nm data transmission system including different photodetector and fibre geometries. A factor of 1 represents the class I limit, and a factor of 0.1 represents the class IM limit (also shown as a darker shade of grey). 210
7-9 Design chart for an 850 nm data transmission system including germanium lateral and vertical detectors, and different fibre geometries. A factor of 1 represents the class I limit. 211

A-1 The unpolished surface of a silicon wafer following step 1 of the Ge RCA cleaning procedure. Both surfaces of the silicon wafer are covered with rounded etched areas distributed over the surface 224
A-2 Profilometer traces of two different patterns on the same Ge-on-Si wafer following a modified Ge RCA step 1, followed by oxide mask removal. The jagged feature on the right falling edge of a) is due to a damaged profilometer tip. 225

List of Tables

5.1 Electrical characteristics of the 0.9 μm lateral p-i-n devices. 137

5.2 Comparison of measured bandwidths and fitting coefficients with calculated RC bandwidth based on measured capacitance and 50 Ω load. 161

5.3 Comparison of measured bandwidths and fitting coefficients with calculated RC bandwidth based on measured capacitance and 50 Ω load. 162

5.4 Comparison of measured bandwidths and fitting coefficients with calculated RC bandwidth based on measured capacitance and 50 Ω load. 166

5.5 Responsivity figures extracted from the optoelectronic frequency response measurement of Ge lot 2 devices. 170

7.1 Eye safety levels for Class 1 and Class 1M systems in mulitmode optical fibre (diameter $< 150\ \mu$m) [95]. 196

A.1 Standard Si RCA clean recipe 221

A.2 Ge RCA clean recipe 223

B.1 Silicon lateral p-i-n detailed process description 227

B.2 Epitaxial germanium growth and annealing process description 229

B.3 Germanium lateral p-i-n detailed process description 230

Chapter 1

Introduction

The field of optoelectronic integration, for communications or interconnect applications, has seen a tremendous amount of research over the last two decades. Academic groups and companies have been motivated either by the promise of low cost and high functionality that comes with Moore's Law or by the need to develop ways of extending Moore's Law into the future due to apparent limits of electrical scaling, particularly in the field of interconnects. The former has yet to be realised due to the relatively low volumes of the real-world optical interconnect market, and the need for the latter has been gradually displaced by innovative driver, receiver, packaging, and circuit board designs that have shifted the transition point where optical communication becomes attractive relative to electrical communication to ever higher bitrate-distance products. Optical communication over single-mode fibre is the preferred choice for high data rate, long distance (hundreds of kilometres) communication, but the situation becomes less clear over shorter distances. For example, using 4-level electrical pulse amplitude modulation, 10 Gb/s has been implemented on printed circuit board (PCB) backplanes [1]; the sort of bitrate and distance where optical communication had once been considered the more viable option. The expectation is for electrical bitrates to continue to rise in the future, which has prompted some groups to re-evaluate the driving factor for high-volume optical interconnect. One conclusion is that the advantage of optical interconnect may come from its bandwidth density ($Gb/s/cm^2$) as electrical connector scaling becomes a limiting factor. Alternately, there exist applications where electrical interconnect is undesirable for relatively mundane reasons such as ease of installation or cabling. In either case, it is most desirable for the optoelectronic driver and receiver circuitry to be based on a silicon

platform for low cost in expected high-volume applications.

1.1 Areas of Interest for Optoelectronic Integration

High Density Interconnect

Work on very high density interconnect has centred around vertical cavity surface emitting lasers (VCSELs) due to their projected low cost and relatively easy fabrication of large array. The path towards the state of the art was conceived by King *et al.* [2, 3], where 4x8 arrays of VCSELs and photodiodes were flip-chip bonded to a silicon complementary metal oxide semiconductor (CMOS) driver and receiver chip. The VCSELs characterised by microwave probing demonstrated a modulation bandwidth of 12.5 GHz, while the system bandwidth was limited by the driver circuit to 1 Gb/s/channel in the initial report and 2.5 Gb/s/channel subsequently over fibre ribbon cable. Coarse division multiplexing of VCSEL wavelengths in multimode glass fibre (MMF) was also demonstrated at 4x10 Gb/s. The authors identified the future scaling, assuming an improvement in driver circuitry, to 20 Tb/s/cm^2 for a 250x250 μm unit cell.

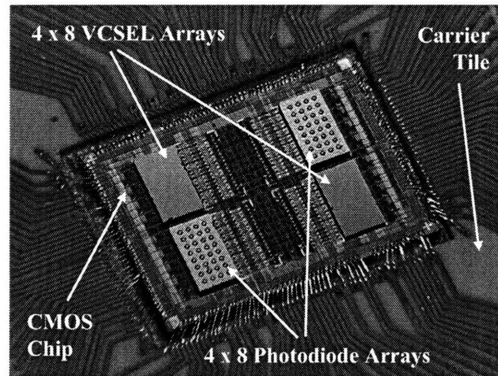


Figure 1-1: 32 channel transmitter/receiver module using flip-chip bonded VCSEL and photodiode arrays from [3]

The promise of high interconnect density was also recognised by IBM and realised in their ongoing Terabus project [4]. The project follows essentially the same line as the aforementioned effort (flip-chip bonded VCSEL and detector arrays on silicon CMOS chip), but benefits from additional years of progress in fabrication and circuit technology enabling the demonstration of a 14 Gb/s

link between transmitter and receiver. Additionally, the CMOS/optoelectronic unit is aligned and attached to a so-called Optocard, featuring planar polymer waveguides as well as electrical interconnection traces. The polymer waveguides have a large core (approximately 30 μm dimensions) on a 62.5 μm pitch. Coupling is achieved through 90 degree gold-covered polymer mirrors at the VCSELs and detectors. The system is envisioned for high-bitrate communication over distances of well under 1 m, for example between modules on a common card or between chips in a module. It is envisioned that ever-increasing bandwidths will be required in multi-core applications and between CPU and GPU units in future computer and gaming systems.

Both of the above systems operate in the 980-990 nm regime to allow for backside emission through the GaAs growth substrate. Also in both cases the photodiodes are vertical III-V p-i-n structures. Integration of the photodetectors onto the silicon chip substrate would be attractive from the point of view of cost, reduced parasitics, and relability. The desired device size is on the order of 40-60 μm diameter.

Polymer Optical Fibre

Typical optical communication systems use glass optical fibre, either single mode for long distances (many kilometres) or multimode for short distances (hundreds of metres). The bandwidth-distance product of the glass fibres is determined by their loss characteristics ($\ll$ 1dB/km) and the modal dispersion (the limiting factor in MMF). An alternate type of fibre is polymer (or plastic) optical fibre (POF). The most common varieties are made of a poly(methyl methacrylate) (PMMA) core and a fluorinated polymer cladding. The PMMA material has much higher loss than glass fibre, with relative transparency windows around 540, 560, and 650 nm at which the loss is on the order of 100 dB/km. Due to this high loss, POF is not suitable for long distance communication. The high attenuation allows for an interesting design decision: since modal dispersion is not the limiting factor for the transmission channel, the fibre may be produced with a very large core, typically 0.5-1 mm. This in turn makes the fibre compatible with inexpensive moulded plastic connectors, relaxes alignment tolerances, and relaxes the requirements for fibre cleaving and polishing. Additionally, the fibre itself is mechanically robust, resistant to vibration and bending damage. The wavelength of choice for these applications is 650 nm at which a range of light

emitting diodes (LEDs), resonant cavity LEDs (RCLEDs), and recently VCSELs operate [5, 6]. The maximum bitrates for these sources are approximately 50 Mb/s, 250 Mb/s, and commercially available greater than 2 Gb/s respectively.

Current applications of PMMA fibre include localised and decorative illumination (for example in keyboards and automobile instrument clusters), sensing in electrically noisy environments, and low bitrate communication in noisy and/or mechanically harsh environments. This includes trains and, in particular, automobiles. The wiring harness in the latter is an increasingly complex assembly as manufacturers integrate on-board computers, entertainment systems, satellite sensors for anti-lock braking and traction control, rear-view camera, local radar, etc., into their products. The data bus connecting all of these elements must be resistant to significant electromagnetic interference, a strongly corrosive environment, sharp bending radii, and strong vibrations necessitating an expensive and heavy cabling system. In 1998, BMW developed a POF-based bus for its 7-series automobiles called "Byteflight" [7]. The system defines a 10 Mb/s half-duplex bidirectional protocol over a single POF physical layer, primarly designed to provide the on-board computer with safety data and to allow the on-board computer to deploy safety measures including airbags. A photo of the transceiver designed by Infineon, including an LED flip-chip bonded on top of a large detector and housed in an inexpensive moulded plastic package, is shown in Figure 1-2.

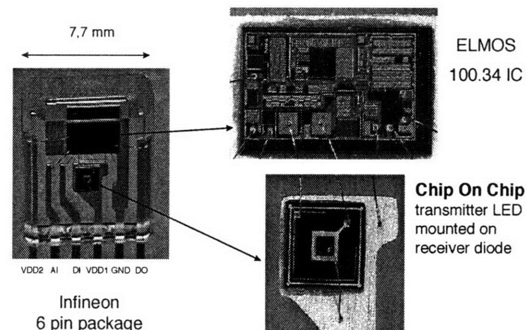


Figure 1-2: Photo of a Byteflight bidirectional transceiver from [7]

The lessons of Byteflight and the parallel development of the D2B bus by Daimler-Chrysler led to the formation of the media oriented system transport (MOST) standards group [8] in 2000, which defines a unidirectional protocol on a POF physical layer, operating initially at 22.5 Mb/s and recently moving to 50 Mb/s with the intention of 150 Mb/s and beyond in upcoming years.

The physical transceivers include red LEDs and discrete photodetectors, with RCLEDs poised to enter as the standard moves to higher bitrates. At the end of 2003, 13 European car models came equipped with some form of MOST bus, with transceiver shipments approaching 10 millions units per year. As a complementary effort, an IEEE 1394b-over-POF effort was launched by the industry with 250 Mb/s operation and the possibility to use existing control ICs as a building block of the transceiver. Finally, there is interest in moving 1394b-over-POF to office or home use, as POF cable installation is less expensive than glass MMF systems (e.g. ethernet over MMF) with a higher bandwidth than CAT5 twisted pair wiring. Transceiver manufacturers even at the 250 Mb/s data rate have found difficulty in selecting appropriate photodiodes for this application that provide a reasonable bandwidth for a large area, important because the sensitivity is directly related to the modal overlap of the fibre mode and the detector [9].

Finally, other POF materials, precisely perfluorinated PMMA, have been shown to exhibit a much lower absorption loss, particularly in the interesting regimes of 780 and 850 nm where more robust and higher bandwidth VCSEL sources could be used in the transmiitter [10, 11, 12] . In light of the lower attenuation, these fibres are manufactured with 120-300 μm diameter cores, and the cores have a graded index structure to reduce dispersion. They exhibit a highly multimode behaviour, but over short to moderate distances (25-50 m) are able to maintain relatively high bandwidths with only several dB of loss. For example, there was recent demonstration of 10 Gbps transmission over 50 m of GI-POF using an 850 nm VCSEL source [13]. These fibres are intended to be used in future short-distance home, business, and entertainment networks (for example, replacing electrical HDMI cables in HDTV systems).

In the context of such high bandwidths as are achievable with GI-POF, the problem of an appropriate photodetector matched to the large core diameters well above 100 μm becomes very serious. A detector is desired with the following minimum properties: a large diameter for high modal overlap and liberal misalignment tolerance, and low capacitance and intrinsic bandwidth limit. Additionally, it is desirable that the detector be integrated in a CMOS fabrication process for low-cost volume manufacturing. Depending on the application, the responsivity of the detector (measured in A/W) may be of secondary importance, e.g. industrial or automotive applications may not require the system to confirm to eye safety limits, as opposed to a commercial or household application.

Unfortunately such photodetectors do not exist commercially, with approximately a 100 μm diameter being the largest available detector at this time. These largest detectors are GaAs-based, unsuitable for silicon CMOS integration, and specified to bitrates below 2 Gb/s. Additionally, there has been almost no published research in the last 10 years related to large-area detectors: the focus has been on achieving bandwidth records for very small devices, or occasionally on achieving several Gb/s bandwidths for glass MMF applications (device diameters around 50 μm). The purpose of this work is to examine this missing area of knowledge in order to enable high bandwidth systems over GI-POF.

1.2 Integrated or Si CMOS-compatible Photodetectors

The majority of work on integrated Si detectors has focused on a lateral or interdigitated geometry because such a structure offers a lower capacitance/area than vertical structures, as will be discussed in chapter 4. However, there has also been some effort in vertical structures with special epitaxial growth structures.

1.2.1 Si Lateral Photodetectors on bulk substrates

The simplest form of lateral photodetector is a structure consisting of interdigitated p and n electrodes on a bulk substrate [14], a structure that was also fabricated in this work and will be called a control device (section 5.2.1). In this structure, an applied reverse bias results in an expansion of the depletion region around one of the finger polarities vertically and laterally. The capacitance of such a structure is low. However, as wavelengths of interest (particularly 850 nm) are weakly absorbed by silicon, the bandwidth is strongly affected by minority carrier diffusion over a long distance. Alternatively, the structure may be more successfully applied to shorter wavelengths, for example 650 nm, which is applicable to POF applications. In this case the diffusion distance is shorter, therefore improving the bandwidth. The authors report on structures with dimensions of 100 $\times$ 100 μm, finger width of 1 μm, and finger spacing of 4 μm. Responsivity of 0.37 A/W is achieved at 850 nm and 0.36 A/W at 660 nm (external quantum efficiency of 54% and 69% respectively). The small-signal bandwidth is below 100 MHz for both wavelengths at reasonable reverse biases of 0-5 V. However, the authors present an eye diagram measurement

claiming greater than 1 Gb/s performance at 20 V reverse bias and 850 nm wavelength. The latter is not implementable for an integrated solution, but the work does exhibit interesting properties of RC versus diffusion behaviour.

Due to the large substrate diffusion current, the aforementioned authors envisioned a spatially modulated photodetector [15]. The structure is shown schematically in Figure 1-3. It is essentially an identical structure as described above, but with half of the detector area shielded by an upper metal layer. It is assumed that substrate-generated carriers diffuse equally to shadowed and exposed electrodes. Therefore the signal is read differentially between these two sets of electrodes, and the resulting signal is the fast lateral current generated in the exposed regions. The small-signal bandwidth of the differential signal is approximately 1 GHz at 2 V reverse bias, and the responsivity is not stated exactly but is admittedly very low. Using an integrated TIA and limiting amplifier on-chip, the authors demonstrate a 2 Gb/s open eye, albeit at a rather high optical power (over -8 dBm). The device and circuitry were fabricated in a standard CMOS process.

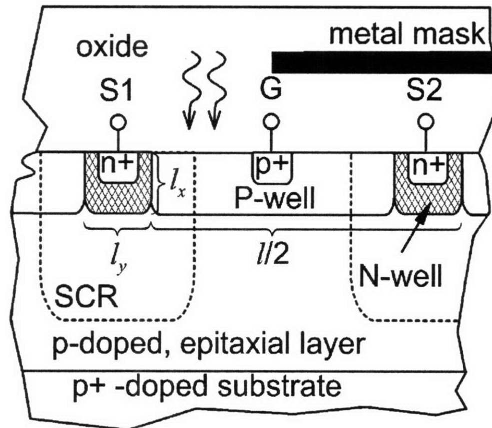


Figure 1-3: Schematic drawing of a spatially modulated photodetector from [15]

In an effort to increase sensitivity and bandwidth by collecting more of the carrier generated deep in the substrate, authors at IBM presented a silicon lateral trench photodetector [16]. A schematic of the structure is shown in Figure 1-4. It is a simple lateral geometry, with the exception that the interdigitated electrodes are in fact deep (7 μm) trenches formed by RIE and filled with doped polycrystalline silicon. Therefore a uniform lateral depletion region and electric field is established far from the semiconductor surface. The devices are circular with a diameter of 75 μm, exhibiting a responsivity of 0.47 A/W (68% efficiency) at 850 nm. The small-signal bandwidth

for 670 nm light at reverse bias of 3-5 V is approximately 3 GHz. For 850 nm light, the bandwidth is approximately 100 MHz, though the authors prefer to quote the 6 dB frequency of 1.5 GHz. The low frequency response is due to further carrier diffusion, as admitted by the authors. The suggested solutions are deeper and wider trenches or isolation oxide [17]. The former is not a scalable solution as device capacitance grows with finger depth, and indeed the devices as presented may not be scaled to larger sizes due to the high capacitance of the deep fingers. The latter suggestion is not industrially applicable due to the relative scarcity of very thick (> 8 μm) SOI. The authors further present a 2.5 Gb/s eye diagram at 845 nm, however admitting that it was achieved by AC-coupling of the detector to a TIA thereby filtering the low frequency component of the response.

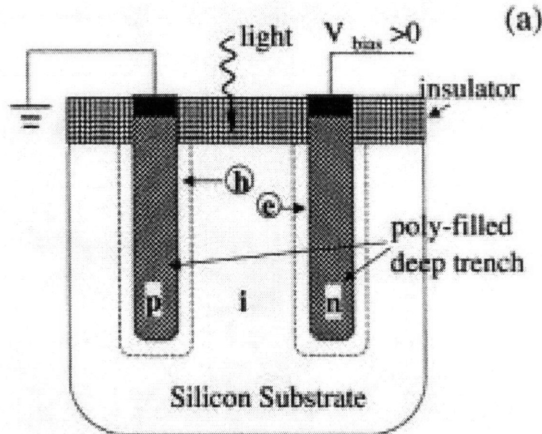


Figure 1-4: Schematic drawing of a silicon lateral trench photodetector from [16].

1.2.2 Lateral Photodetectors on SOI substrates

Significant progress was made on lateral photodetectors on SOI substrates, primarily by the research group of Campbell at the University of Texas in Austin. The insulating buried oxide blocks diffusion of carriers generated deep in the substrate, such that they are not collected and do not contribute a slow component to the photocurrent. These devices are simulated in the discussion portion of this work (section 6.3) to confirm the reported results and gain further insight into the structure.

Discrete devices were fabricated on very low resistivity SOI substrates (50-100 Ω·cm), with SOI thickness ranging from 2.7 to 6.2 μm and a device area of 50 × 50 μm [18]. Due to the index of

refraction step at the buried silicon-oxide interface, a weak resonant cavity is formed in the top SOI layer which affects measured responsivity as a function of wavelength. Peak external efficiency is independent of the SOI thickness due to the cavity effect, at 26-29%. The implant fingers of the devices are not metallised which raises sensitivity but precludes scaling to larger device sizes due to the effect of series resistance. Bandwidth was up to 3.4 GHz at 5 V reverse bias for the thinnest SOI structure. Bandwidth decreased with increasing thickness as less of the inter-finger space was depleted for a given reverse bias. It is interesting to note that the frequency response decay slopes for thicker SOI devices in this citation correspond to the results of the derivations in section 2.3.3, though the authors do not identify the frequency-limiting behaviours explicitly.

The authors above subsequently integrated the SOI detectors with amplifier circuitry in a CMOS-compatible processes, and subsequently a standard 130 nm CMOS flow [19, 20, 21]. The different reports feature similar low resistivity SOI substrates with the device structures discussed above (with varying finger spacing between reports), but different transistor structures and therefore amplifier bandwidths. The detectors also exhibited lower bandwidth than in the previous report due to shallow n and p finger that were implemented by source/drain implants. To attain comparable bandwidths, reverse bias was raised to 10 and 20 V. Bitrates at low bit error rate (BER) of 1 and 2 Gb/s were attained with high sensitivity, and bitrates up to 5 and 8 Gb/s were demonstrated with 20 V reverse bias and high input power of 2 dBm.

The highest bandwidth achieved on an SOI substrate was reported by Liu *et al.* for a 5×5 μm MSM structure, fabricated using electron beam lithography and metal lift-off to achieve 100 nm electrode widths and spacings [22] . The SOI thickness was also approximately 100 nm, resulting in efficiency at 780 nm of 0.9%. Bandwidth, measured by impulse response, was determined to be 140 GHz at an unspecified reverse bias. The authors identified that any scaling of the device to larger dimensions would strongly degrade bandwidth.

1.2.3 Integrated Silicon Photodetectors involving epitaxy and process modifications

An innovative epitaxial silicon photodiode integrated in a bipolar or BiCMOS fabrication process was first proposed and implemented by Yamamoto *et al.* [23]. This structure was implemented and studied in detail by Zimmerman *et al.* [24, 25]. The detector structure has since been

integrated into a new standard analog BiCMOS process at Austria Microsystems. A schematic diagram of the device structure and integration with bipolar transistors is shown in Figure 1-5.

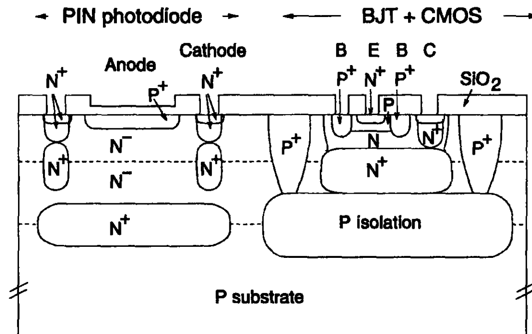


Figure 1-5: Schematic drawing of a vertical photodetector integrated in a BiCMOS process, from [25].

The fabrication process begins with a p-type substrate into which an n+ region is diffused. Next, a two-step epitaxial growth is carried out to deposit lightly n-type material (doping level $10^{13} - 10^{14}$ cm^{-3}), and after each step n+ plugs are diffused into the epitaxial material. The end result is a 5-10 μm tall n- region that serves as the i region of the p-i-n diode. The cathode is defined by the n+ diffusion at the bottom of the epi stack, and n+ diffused plugs through the epitaxial layers (the same structure as a subcollector in an npn transistor). The anode is defined by a p+ diffusion at the wafer surface. In the BJT area, a p isolation region is diffused into the epitaxial growth, as well as an n+ region that serves as the subcollector. The modification to the standard fabrication process comes in the form of the additional epitaxy and diffusion steps to prepare the virtual substrate for the transistors. However, after the epitaxy the process continues along the same steps as the standard BiCMOS process, and the epitaxial structure does not degrade the f_T of the bipolar transistors.

The authors have applied this detector structure to several areas. A low-power 300 Mb/s optoelectric integrated circuit (OEIC) with octagonal detectors 400 μm in diameter and biased at 2.5 V was demonstrated for POF applications [26]. The detector included an anti-reflection coating and achieved an external quantum efficiency of 96%. The TIA and detector were fabricated monolithically, with total chip area of 890 $\times$ 515 μm dominated by the detector, and power consumption under 6 mW. It has also been demonstrated as the basis on an integrated optical receiver for CD/DVD/blue-laser applications [27]. In this case, large size and highest bandwidth are not

required, and the device is run biased at 2.5 V.

The device performance is limited in part to capacitance and in part to a large drift delay through the long nominally i region (discussed in section 2.3.2). By biasing the detector at a higher voltage than the supply (12 V and 5 V respectively), the detector bandwidth was increased substantially [24], demonstrating 1.8 Gb/s performance at high sensitivity. Further refinements led to 2.5 Gb/s operation with a 12 V detector bias, and 300 μm detector diameter [28]. Finally due to the disadvantage of two power supplies, and voltage up-converter was implemented on-chip (5 V supply to 11 V detector bias) and a 5 Gb/s OEIC was demonstrated for a 50 μm diameter detector [29]. It was found that there was no problem of noise coupling between the converter and TIA circuit as had been feared in the literature.

It should be noted that all of the aforementioned results were achieved using a 660 nm light source, the characteristic absorption length of which is approximately 3.5 μm . Therefore, all incident light was absorbed in the long epitaxial i region (fully depleted at an appropriate bias), and the problem of carrier diffusion would be ignored entirely. The measured bitrates for longer wavelengths such as 780 or 850 nm are much lower due to the weaker absorption properties and lack of any carrier blocking materials in the structure. The devices are however very well suited to red light and short wavelength applications. The only disadvantage arises from the need for a second high-voltage supply if the integrated converter is not used and high bandwidth is required, and the not insignificant added expense and complication of the silicon epitaxy and dopant diffusion required to define the detector structure.

1.3 Germanium photodetectors

Germanium has many attractive properties compared to silicon while maintaining basic Si processing compatibility. The narrow band gap results in an absorption length of 350 μm for 850 nm light and absorption to beyond 1.6 μm . Additionally, the mobility is higher by a factor of 3 for electrons and a factor of 4 for holes, though this fact is only helpful in low-field conditions since the saturation drift velocities are lower in germanium. There have been reports of various structures, for example MBE-grown germanium on silicon detectors [30], and germanium diodes grown on graded SiGe buffers[31]. The silicon substrate is interesting in order to maintain process capability

and the ability to integrate with standard CMOS devices. In recent years, work by Luan *et al.* [32, 33] resulted in significant progress in the field as a two-step growth process followed by *in situ* annealing allowed for the direct epitaxy of germanium on silicon with much lower dislocation density than previously achieved. The difference in thermal expansion coefficients may induce a tensile strain in the germanium epilayer, since it is mostly or fully relaxed during growth due to the large lattice mismatch. Tensile strain results in a decrease in direct band gap, increasing the absorption of wavelengths around the interesting telecommunication C and L bands [34, 35].

1.3.1 Vertical p-i-n photodetectors

There have been several reports of vertical Ge-on-Si vertical p-i-n structures. Liu *et al.* presented a structure built by patterned growth in oxide wells on p+ silicon substrates. The devices exhibited calculated internal efficiencies of 90% in the wavelength range 650-1340 nm. The epilayer, nominally i region, thickness was 2.35 μm, and transit time was the dominant performance limit at that height. Bandwidth was determined by impulse measurement to be 8.5 GHz for a 10×70 μm device. Further device area scaling would have had a negligible effect on bandwidth, as the device was transit time limited and half of the measured capacitance originated from the measurement pads (as discussed in section 2.3.4). A thinner epilayer would extend the transit time bandwidth, at the cost of reduced responsivity in the C and L bands.

A high responsivity for a thinner epilayer region may be obtained by inserting the absorbing region into a resonant cavity [36]. The cavity is conveniently formed by growth on an SOI substrate and metallisation of the top surface. In the case of the citation, the growth substrate was a double SOI substrate with silver metallisation of the top surface to increase reflectivity of both facets. Peak simulated quantum efficiency was expected to be approximately 90%, though the measured response was much poorer due to germanium epilayer quality (i.e. a lack of annealing step in the processing). A subsequent report of a similar structure used an SOI growth substrate, gold mirror on the top surface (forming a Schottky contact rather than a p-i-n structure), and illumination through the SOI [37]. For a 1.4 μm thick germanium layer, the measured quantum efficiency at 1.535 μm was 60%. Measured bandwidth at reverse bias of 2-5 V was approximately 10 GHz for 10 μm diameter devices, and 9 GHz for 28 μm diameter devices (comparable area to the previous

citation).

Recently, high performance MBE-grown Ge-on-Si photodetectors have been reported [38] . The growth procedure is somewhat different than in the UHVCVD method, though it also comprises a thin initial Ge layer for strain relaxation. The i region thickness is only 300 nm thick, and the entire structure is only 1 μm tall. These thicknesses, along with a lack of tensile strain due to a low growth temperature of 300°C result in a relatively low responsivity: external quantum efficiency is 15% at 850 nm and between 10 and 1% in the range 1500-1600 nm. These devices set a bandwidth record of 39 GHz at 2 V reverse bias for 10 μm diameter devices. At that size, they are approximately transit-time limited. As size is increased, the bandwidths follow the theoretical RC limit as capacitance rises due to the thin i region: 16 GHz and 8 GHz at 20 and 30 μm diameters respectively. These bandwidths are slightly higher than comparably sized devices cited above, and demonstrate the large design space available to maximise responsivity-bandwidth product for a certain required device property.

1.3.2 Lateral p-i-n photodetectors

The first recent report of interdigitated Ge p-i-n photodetectors presented the results of a structure on a bulk germanium wafer, with dimensions of 50×50 μm . The substrate was Sb-doped resulting in a resistivity of 1-5 Ω·cm, corresponding to a doping concentration of $1 - 2 \times 10^{15}$ cm^{-3}. Fingers were 1 μm wide with 2 μm spacing, defined by ion implantation into the wafer [39] . The finger buses were metallised with Ag, and silicon dioxide was used as anti-reflection coating. External quantum efficiency was over 60% in the range 1.0-1.5 μm . A bandwidth of 1.8 GHz was measured at 5 V reverse bias, rising to 3 GHz at 15 V. This structure demonstrated the potential of an interdigitated geometry to provide high performance devices on germanium.

The above work was followed by ground-breaking results Dehlinger, Koester *et al.* of IBM, who demonstrated extremely high bandwidths for epitaxial germanium detectors grown on ultra-thin SOI substrates [40, 41, 42]. The structure consisted of 400 nm of germanium grown on a 15 nm SOI substrate (and an additional 35 nm of Si buffer) by a two-step UHVCVD method as described above, followed by cyclic thermal annealing. After annealing, Hall measurements indicated that the material was p-type with mobility of 1200 cm^2/Vs and carrier concentration of

1×10^{16} cm^{-3}. The germanium film was patterned to form device mesa of 10×10 to 30×30 μm dimensions. Fingers were fabricated by implantation into the germanium through photoresist masks, and Ti/Al electrodes were deposited by lift-off. All photolithography was carried out by electron beam lithography. A device schematic is shown in Figure 1-6.

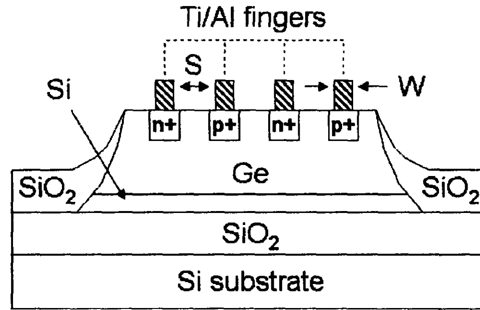


Figure 1-6: Schematic drawing of a lateral p-i-n photodetector fabricated from germanium on a thin SOI substrate, from [40].

A weak resonant cavity was formed by the SOI substrate, and the peak external efficiency was 52% at 895 nm. Bandwidth was measured by the impulse response method, with a peak of 29 GHz at -2 V for 0.4 μm finger spacing and the same bandwidth at -4 V for 0.6 μm spacing. The peak efficiency-bandwidth product was 13.2 GHz at 850 nm. Interestingly, from a review of the literature there has been no further device development since mid-2004. The Terabus project discussed earlier in this section uses III-V photodiodes despite participation by many of the authors of the Ge detector work. The only report of a receiver using the Ge detector appeared recently [43], and the authors present a 17 Gb/s receiver using the Ge detector described above, but connected by wirebonds to a CMOS TIA and limiting amplifier circuit.

1.4 Thesis Goals and Organisation

The goal of this work is to investigate device structures that could be applicable to large-area photodetectors suitable for above 1 Gb/s communications systems. As may be seen from the literature review in the previous sections, the vast majority of photodetector research in the last 10 years has concentrated on high data rates with device areas not exceeding 2500 μm^2. The exception is the work of Zimmerman *et al.* who recognised the value of a red-light POF system, but

whose design is certainly unsuitable for the 780 or 850 nm wavelength regime. This work will investigate far larger device areas, with the minimum goal of demonstration of a bandwidth above 1 GHz for a 200 μm diameter device (31500 μm^2). Lateral geometry devices will be applied to reach this goal. Device scaling beyond this nominal area and bandwidth will be examined by measurement and simulation.

Objectives:

1. Design and implementation of a small signal frequency response measurement system for optoelectronic devices. Investigation of alternate bandwidth measurement techniques, and implication of parasitic elements and instrumentation choices on the measured or inferred frequency response.

2. Design, fabrication, and measurement of silicon lateral photodetectors fabricated in a fully CMOS-compatible process using standard processing steps. Investigation of different frequency response regimes by the use of different finger spacings and device areas, thus enabling the detemination of the physical bandwidth-limiting mechanism. This is the first study to apply an empirical curve fitting of measured results to directly determine the dominant physical mechanism.

3. Simulation of the silicon device structure to confirm measurement results and physical mechanisms. Investigation of device sensitivity to fabrication parameters and ultimate performance limits assuming various forms of device scaling.

4. Application of simulation structure to alternative materials, specifically epitaxial germanium on silicon substrates. First attempts at fabrication and measurement of germanium lateral devices.

The thesis is organised as follows:

Chapter 2 Explanation of fundamental photodiode behaviour. Derivation of basic small signal model, discussion of drift delay bandwidth. Derivation of diffusion current frequency response for different illumination conditions, a topic previously not addressed in textbooks

or in the literature. Discussion of parasitic circuit elements, their influence on apparent device frequency response, and interaction with basic amplifier/receiver topologies.

Chapter 3 Discussion of three bandwidth measurement techniques, the appropriate apparatus for each technique, and examples of good and bad measurements. Full discussion of the frequency-domain apparatus designed and implemented by the author in this work, including proper calibration techniques and the influence of parasitic circuit elements.

Chapter 4 Discussion of the design of silicon and germanium lateral p-i-n devices, including estimated maximum bandwidth and a more realistic first-order estimation of device performance. Presentation of fabrication flows for silicon and germanium devices, and discussion of specific processing techniques and insights.

Chapter 5 Presentation of measurement results of 5 fabrication lots of silicon devices. The best devices with 100 μm diameter have a 3 dB bandwidth above 4 GHz for 850 nm illumination, the highest value ever reported for a device fabricated using standard bulk substrates (i.e. not SOI or specialised epitaxial growth). Presentation of measurement results of 2 fabrication lots of germanium on silicon devices. Observation of a previously unreported and unexpected illumination dependence of responsivity and frequency response.

Chapter 6 Presentation of simulation results of both silicon and germanium devices on silicon devices. Discussion of measurement results with reference to the simulation results. Discussion of ultimate performance possibilities and scaling of devices.

Appendix A Discussion of RCA wafer cleaning techniques as applied to germanium on silicon wafers. Discussion of the inadequacy of current cleaning technique, deleterious side effect of germanium cleaning on exposed silicon surfaces, germanium etching mechanisms, and possible alternative cleaning techniques.

Appendix B Detailed process listing of silicon and germanium device process flows.

Chapter 2

Photodiode and Receiver Fundamentals

This chapter will develop the full incremental model of a photodiode, including intrinsic and extrinsic device components. A distinctive feature, as opposed to essentially all reference texts on this subject, will be a thorough discussion of the frequency dependance of diffusion photocurrent. Basic receiver topologies will also be presented in the context of noise, bandwidth, and parasitic circuit element effects. These concepts will form the basis of the discussion in Chapters 3 - 5.

2.1 Absorption

In all light detection mechanisms, optical fields excite electrons and the resulting current or current change is proportional to the excitation rate of the electrons. When excited, the electrons pass from a bound state to an excited state in which they are mobile. An initial state may include, for example, an electron in a full valence band (for an n-type semiconductor) or in a localised donor state. With a transition between states, a photon of energy $h\nu$ is absorbed, corresponding to the energy difference between states. The mobile carriers in free end states contribute to photocurrent flow.

Direct band-to-band absorption occurs when bound carriers interact with a photon whose energy is greater than the band gap of the absorbing material. The momentum of a photon (h/λ) is small compared to the crystal momentum of the lattice (h/a, where a is on the order of Å, and λ is on the order of $10^3 - 10^4$ Å). The photon-absorption process should conserve the momentum of the excited electron, therefore indirect transitions are much less efficient since they require a two-

step process, whereby momentum is conserved via an optical phonon interaction. The absorption process is described in detail in [44].

Absorption in a material is generally expressed as an inverse length $\alpha(\lambda)$ (also called a penetration depth), defined as the relative rate of decrease in light intensity $L(\lambda)$ along its propagation path.

$$\alpha(\lambda) = \frac{1}{L(\lambda)} \frac{d[L(\lambda)]}{dx} \tag{2.1}$$

The absorption coefficients for several materials of interest are shown in Figure 2-1. The sharp decline in absorption near a direct band gap energy is clearly visible. Silicon exhibits a steady increase of absorption with decreasing wavelength due to its indirect bandgap, whereas germanium exhibits a transition from indirect to direct absorption around $\lambda = 1.55$ μm. The historical advantage of III-V materials over group IV materials is evident in the three wavelength regimes of interest (0.78-0.85 μm - GaAs over Si, 1.3 μm quaternary lattice-matched to InP over Ge, 1.5 μm ternary lattice-matched to InP over Ge).

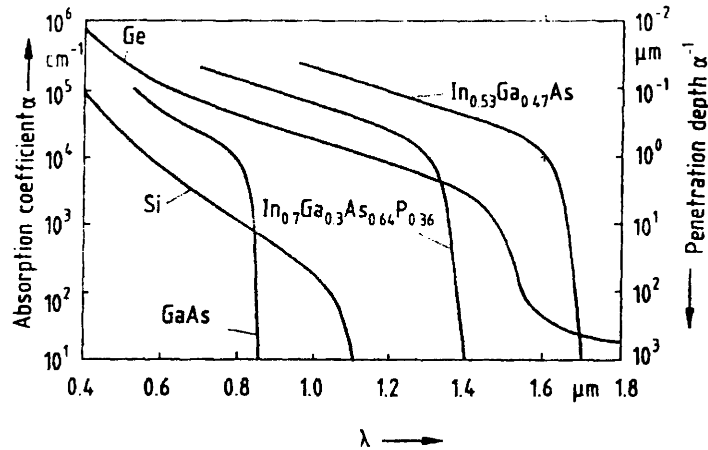


Figure 2-1: Absorption coefficients for several materials as a function of wavelength, from [45]

The absorption coefficient may be improved by using structures that introduce strain into the absorbing material [46]. There has recently been a significant and successful effort to extend the direct absorption of germanium grown on silicon substrates to longer wavelengths in order to compete with InP-based solutions [35].

2.2 Basic Photodiodes

Essentially any p-n junction may act as a photodiode assuming that the semiconductor material is sensitive to the wavelength of light of interest. The analysis of a semiconductor material with a given carrier density (and/or gradient) and a given applied electric field is governed by five basic differential equations accounting for electron and hole current (by drift and diffusion), electron and hole continuity (relating carrier concentration to net carrier flux), and Gauss's law (relating net charge to the gradient in electric field) [47]. However, these equations are coupled, non-linear, and difficult to solve in general, even numerically. There are two important assumptions/approximations that transform the system of equations into a set of linear, non-coupled equations. A full derivation of the properties of a diode may be found in any good microelectronics text, for example [47, 48], and only several important aspects are repeated here.

- Minority carriers flow by diffusion. Therefore, the total minority carrier current density depends solely on the carrier concentration gradient, and non-linear terms relating the minority carrier concentration and the electric field are eliminated.
- Depletion region approximation. At an abrupt p-n junction, an area is formed on both sides of the junction where the mobile charge populations are zero. There are no electric fields outside of this region, and the potential gradient from the p- to n- regions is restricted to the depletion region. This approximation is valid since the depletion region width is generally many times larger than the Debye length on either side of the junction.

The electric field in the depletion region comes as a result of the fixed (ionised) charge from p and n dopant atoms on either side of the junction (since mobile charge is zero), i.e. $\rho(x) = N_A$ on the p side assuming uniform doping. The depletion width and field profile vary with doping levels on either side of the junction following Gauss's law.

$$-\frac{d^2\phi(x)}{dx^2} = \frac{dE(x)}{dx} = \frac{\rho(x)}{\epsilon} \tag{2.2}$$

In the case of a p-i-n structure, in the i region $\rho(x) = 0$, therefore there is a uniform electric field and most of the change in potential occurs over the i-region. Likewise an applied reverse bias results in an increase in the uniform electric field in the i region, much like a parallel plate

capacitor. This is shown in schematic in Figure 2-2, and the advantages of this arrangement will become apparent in later sections of this chapter.

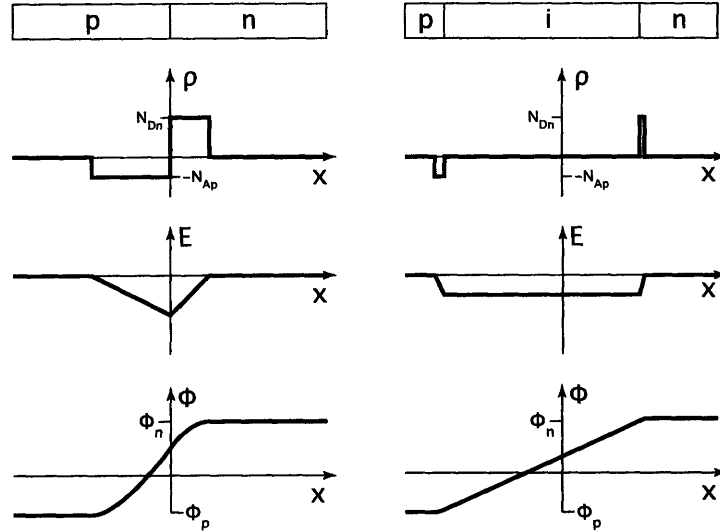


Figure 2-2: Schematic representation at thermal equilibrium of charge density, electric field, and potential (relative to intrinsic material) for an abrupt p-n junction (left) and a p-i-n structure (right)

As assumptions were made to derive a linear set of equations for the p-n or p-i-n junctions, the solutions must also be linear, meaning that the problem may be solved by superposition. Assuming a diode with applied bias v_D and light generation profile $G(x)$, the diode current may be calculated directly or may be solved more quickly as shown in Figure 2-3.

$$i_D(v_D, G(x)) = i_D(v_D, 0) + i_D(0, G(x)) \tag{2.3}$$

Photocurrent flows by different mechanisms depending on the region in which carriers are generated. If carriers are generated in the quasi-neutral (undepleted) regions (QNR), the mechanism for current flow is the diffusion of minority carriers to the edge of the depletion region, that are then carried across the depletion region by the electric field and emerge at the other end as majority carriers. Carriers generated in the space charge (depletion) region (SCR) travel by drift also emerging as majority carriers on either side of the the SCR. If photogenerated current consists of several components, it may also be written by superposition. For example, there may be contributions of diffusion current in the n region, drift current in the i region, and diffusion current in the p region as shown schematically in Figure 2-4. Linearity holds as long as underlying as-

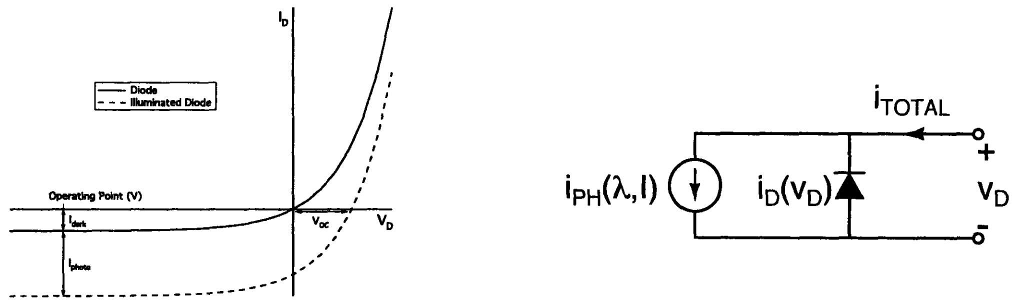


Figure 2-3: Schematic representation of photodiode properties with and without illumination, indicating important device parameters (dark current, photocurrent, open circuit voltage). The device may be represented by the equivalent circuit on the right.

sumptions are not violated, for example the assumption of low-level injection in the quasi-neutral regions (e.g. $n_0 \gg p_n$ for the n region).

$$i_D(0, G(x)) = i_{D,diffn}(0, G_n(x)) + i_{D,drifti}(0, G_i(x)) + i_{D,diffp}(0, G_p(x)) \tag{2.4}$$

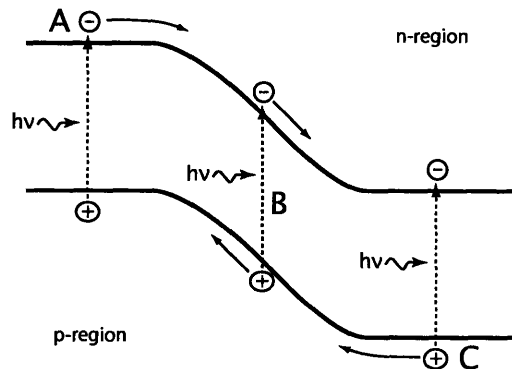


Figure 2-4: Schematic representation of photocurrent: A) generation in p-region and diffusion of minority electrons to SCR, B) generation in SCR and flow of electron and hole by drift, C) generation in n-region and diffusion of minority holes to SCR.

Finally, the problem may be linearised in the normal fashion around an operating point and represented by steady-state and incremental models. The standard notation is shown below.

$$i_D(v_D(t), G(t)) = I_D(V_D, G) + i_d(v_d(t), g(t)) \tag{2.5}$$

2.3 Small-Signal Model

Small-signal (also known as AC or incremental) analysis is used when the signal of interest is a small time-varying fluctuation around a static reference condition. The notation for this state is shown in equation 2.5 for the case of a photodiode. At a given operating point, a small-signal model of the photodiode may be derived in order to describe its behaviour as a function of excitation frequency (optical and/or electrical). Under the assumption of linearity, the two elements of the equivalent circuit in Figure 2-3 may be treated independently. To a first-order approximation, the ideal photocurrent source has no charge storage elements and a flat I-V characteristic ($di/dv = 0$). In general to relate v and i for any general element, a Taylor's series expansion is performed.

$$i_D(v_D) = i_D(V_D) + \left.\frac{di_D}{dv_D}\right|_{V_D}(v_d) + \left.\frac{d^2 i_D}{dv_D^2}\right|_{V_D}(v_d)^2 + \text{higher} - \text{order terms} \tag{2.6}$$

Equation 2.6 was written relating a current dependent on a voltage, but depending on the precise element, current may be related to optical excitation (e.g. photodiode), or voltage may be related to current (e.g. TIA).

The time variation of the driving function, for example generation, is chosen to be of the form

$$G(t) = 1 + A\cos\omega t = Re(1 + Ae^{j\omega t}) \tag{2.7}$$

where in the case of generation $A < 1$ so that $G(t) > 0$ for all t. A solution for sinusoidal variation of the incident light (or other excitation parameter) is sufficient, as the parameter of interest may be written in terms of a frequency response function which allows the calculation of the response to an arbitrary function in time by the use of transform techniques. It is customary to use the complex excitation function rather than its real part. Since the operation of taking the real part of a complex number commutes with differential operator, the solution for the real excitation function is the real part of the solution for the complex excitation function.

2.3.1 Diode Small-Signal Model

Several important points will be mentioned in this discussion, and a full derivation of the model may be found elsewhere [47]. In the Taylor's series expansion (equation 2.6), first-order and higher derivative terms form the incremental model (i.e. relating i_d to v_d). For analysis purposes a linear model is desired, implying operation in a region where the second-order and higher differential terms are much smaller than the first order differential term. This condition is met if the small-signal voltage is small enough, for an ideal diode $|v_d| \leq 2fkT/q$ where f is the proportion of the quadratic to the linear term. A typical value may be $|v_d| \lesssim 5\text{mV}$ for $f = 10\%$. If this condition is satisfied

$$i_d \approx \left.\frac{di_D}{dv_D}\right|_{V_D}(v_d) \tag{2.8}$$

The first-order factor has units of conductance. It is generally represented by the symbol g_d, and is the inverse incremental equivalent diode resistance around the operating point. In the case of the ideal exponential diode

$$\begin{aligned} g_d = \frac{q}{kT} I_s\, e^{qV_D/kT} &\approx \frac{qI_D}{kT} \qquad \text{for forward bias} \\ &\rightarrow \ 0 \qquad \text{for reverse bias} \end{aligned} \tag{2.9}$$

In a real device, current in reverse bias may be dominated by parasitic effects such as sidewall leakage that will result in a finite but generally large conductance value. The incremental parameter g_d is quasi-static, that is to say it does not include a time or frequency dependence. Strictly speaking, to develop a full dynamic model would require revisiting the original equations in the derivation of diode characteristics. A simpler (but nonetheless relatively accurate) approach is to incorporate diode charge stores directly into the small-signal model.

A diode has two charge storage mechanisms. The first is depletion capacitance: the effect of the modulation of the width of the SCR as a function of applied bias

$$w(v_D) = \sqrt{\frac{2\epsilon(\phi_b - v_D)}{q} \frac{N_{Ap} + N_{Dn}}{N_{Ap}N_{Dn}}} \tag{2.10}$$

where ϕ_b is the junction potential and N_{Ap} and N_{Dn} are the doping concentrations on either side of

the junction. A change in depletion region width implies that the amount of fixed charge on either side of the junction has changed. Since the total fixed charge in the SCR is zero due to charge neutrality, one may look at either side of the depletion region, analogous to a discussion of the charge on a parallel plate capacitor. The depletion charge on the p side of an abrupt junction is given by

$$\begin{aligned} q_{DP}(v_D) &= -AqN_{Ap}x_p \quad ; \quad x_p = w\frac{N_{Dn}}{N_{Ap}+N_{Dn}} \\ &= A\sqrt{2\epsilon q(\phi_b - v_D)\frac{N_{Ap}N_{Dn}}{N_{Ap}+N_{Dn}}} \end{aligned} \tag{2.11}$$

where A is the junction area. Applying a method analogous to equations 2.6 and 2.8

$$\begin{aligned} C_{dp}(V_D) &= \left.\frac{dq_{DP}}{dv_D}\right|_{V_D} \\ &= A\sqrt{\frac{\epsilon q}{2(\phi_b - V_D)}\frac{N_{Ap}N_{Dn}}{N_{Ap}+N_{Dn}}} \end{aligned} \tag{2.12}$$

$$\equiv \frac{\epsilon A}{w} \tag{2.13}$$

The equivalence of equation 2.13 is simply the expression for a parallel plate capacitor. This makes perfect sense by considering that for a small signal voltage, mobile charge is added or removed at the edges of the depletion region, and by stating that v_d is small enough to assume linearity in the Taylor's series, one is also stating that the depletion region width does not appreciably change due to the applied small signal voltage. Two important observations arise from equation 2.12. First, the capacitance falls with increasing reverse bias (i.e. negative V_D) in an inverse square root relationship. Second, the magnitude and rate of change of the capacitance is approximately dependent on the square root of the doping concentration of the more lightly-doped side of the junction. This derivation applies for an abrupt p-n junction, but the situation is similar in a p-i-n junction. Recalling Figure 2-2, the electric field in the i region is constant and the SCR width on either side of the i-region is relatively small. Rather than solving for the capacitance directly, equation 2.13 may be applied with the additional assumption that $w \approx d$ where d is the thickness of the i region. This relationship is particularly useful in photodiode design, as it states that the width of the depletion region is effectively constant for a given structure, and that applied

bias may therefore directly be used to very simply calculate the approximate electric field in the depletion region.

The capacitance relationship may be used to measure the profile of a junction. This may be seen by rewriting equation 2.12

$$\begin{aligned}
\frac{1}{C_{dp}^2} &= \frac{2}{\epsilon q A^2} \frac{N_{Ap} + N_{Dn}}{N_{Ap} N_{Dn}} (\phi_b - V_D) \\
&\approx \begin{cases} \frac{2}{\epsilon q A^2 N_{Dn}} (\phi_b - V_D) & \text{for p}^+\text{-n} \\ \frac{2}{\epsilon q A^2 N_{Ap}} (\phi_b - V_D) & \text{for p-n}^+ \end{cases}
\end{aligned} \tag{2.14}$$

Therefore a plot of $1/C_{dp}^2$ vs. V_D should yield a linear plot whose slope and intercept are dependent on device size and material parameters. For non-uniform doping, the shape of the curve may be used to extract the doping profile on the lightly-doped side of the junction. This is a very widely used characterisation technique, and it will be applied when analysing measured and simulated capacitance data in later sections.

The second charge storage mechanism in a diode is diffusion capacitance. It is related to the excess minority carrier population when the device is in forward bias. Photodiodes are operated at zero or reverse bias, such that this mechanism does not contribute to the photodiode small signal model. For completeness, the expression is (for a p$^+$-n device, analogous for p-n$^+$)

$$C_{df}(V_D) = \frac{(w_n - x_n)^2}{2D_h} \frac{q}{kT} I_D(V_D) \tag{2.15}$$

where the first factor has units of time and is called the minority carrier transit time t_{tr}. Two important observations are that 1) the capacitance does not depend on diode size but only on diode current, and that 2) the capacitance disappears when the device is run in reverse bias. The small signal model of the photodiode including the incremental circuit elements is shown in Figure 2-5.

2.3.2 Transit time dependence

At high light modulation frequencies, the transit time of carriers generated in or travelling through the SCR has an effect on the photodiode frequency response. The motion of carriers through the SCR introduces a time delay, and when the frequency at which the incident light is

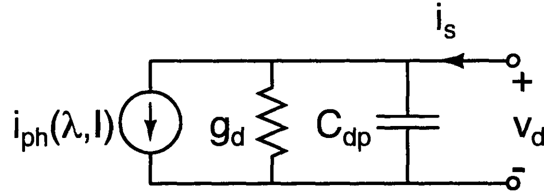


Figure 2-5: Small signal model of the photodiode including the incremental model of the diode.

modulated becomes comparable to the reciprocal of this time delay, the response of the photodiode is degraded.

The definitive derivation of the transit time effect was published by Lucovsky *et al.* [49]. The solution is derived by examining the continuity equations for electrons and holes in the SCR. The equations are coupled and non-linear, but a set of simplifying assumptions may be applied to linearize the system. The assumptions include negligible recombination in the SCR, negligible diffusion current in the SCR, and no perturbation of the electric field due to generated carriers. The latter assumption merits further discussion. Generated holes and electrons are swept to opposite ends of the SCR by the field in this region, resulting in a carrier gradient that induces an electric field in the opposite direction. Therefore, there exists some level of illumination at which the continuity equations may no longer be linearised, or qualitatively the device response degrades due to a reduced electric field in the SCR. This level may be written approximately for a p-i-n device [49], assuming a uniform volume generation rate g, equal hole and electron mobility μ, i region length L, and no recombination (i.e. photocurrent density J_L must equal electronic charge times the number of photons/m^2 absorbed).

$$\begin{aligned} \epsilon\epsilon_0 E_0 &\gg \frac{qL^2 g}{2\mu E_0} \geq \frac{qL^2 g}{2v_{sat}} \\ E_0 &\gg \sqrt{\frac{LJ_L}{2\epsilon\epsilon_0\mu}} \geq \frac{LJ_L}{2\epsilon\epsilon_0 v_{sat}} \end{aligned} \tag{2.16}$$

Since p-i-n devices are typically operated at biases that result in fields on the order of 10^4 V/cm which saturate carrier drift velocities, the second inequality may be used. The inequality is easily maintained in typical devices for photocurrent densities reaching on the order of 10 A/cm^2. It is apparent that the inequality allows for higher photocurrent with shorter i regions, and devices for high power applications are designed with this precise relationship in mind [50]. For real

devices (unequal mobilities, non-uniform electric field, non-uniform generation), there is no intuitive expression analogous to equation 2.16, however the effect of illumination may be calculated numerically [51].

Assuming a more realistic situation including unequal mobilities and nonuniform generation (i.e. exponentially decaying with a characteristic length α), the frequency response of the transit time effect depends on the direction of illumination (i.e. light incident on the p or n side of the device). The full expressions for the two directions were originally published in [49], and have been repeated elsewhere (for example [52]). The expressions were subsequently refined for III-V materials that exhibit a very strong dependence of mobility on electric field, in particular a mobility overshoot due to inter-valley scattering [53]. All of these expressions may be greatly simplified in the limit $\alpha \rightarrow \infty$, in which case all carrier generation takes place at the edge of the SCR ($g = \delta(x - x_0)$). The magnitude of the transfer function becomes equation 2.17, where v is the carrier drift velocity ($v = v_n$ for p-side illumination, $v = v_p$ for n-side illumination) and L is the length of the i region. The half-power bandwidth is stated in equation 2.18. It is sometimes stated as $1.2\pi v/L$ (for example [45]) which refers to the frequency at which the optical power to photocurrent conversion decays to 0.5 (also known as the 6 dB frequency), but this is incorrect from the point of view of electrical small signal circuit analysis.

$$|H(\omega)| = \left|\frac{\sin(\omega L/2v)}{\omega L/2v}\right| \tag{2.17}$$

$$\omega_b \approx 0.9\pi\frac{v}{L} \tag{2.18}$$

This relationship may be used to determine a pessimistic lower bound on transit-time limited bandwidth. Additionally, the real bandwidth for a device with a given illumination profile will fall between the two extremes (i.e. equation 2.17 for $v = v_n, v_p$). A further useful approximation is the treatment of the transit time effect as a first-order effect. The transit time bandwidth (2.18) simply becomes a pole of the total transfer function as proposed in [54]. The frequency behaviour of the sinc function and the single pole are similar for frequencies from DC to the bandwidth. The sinc frequency decay is faster such that this method may over-estimate the magnitude of the frequency response beyond the corner frequency. Conversely, equation 2.17 is generally a conservative estimate such that the net effect is a function closer to the actual behaviour. However, there

is no expression that might prove this statement correct in general, therefore this representation must be considered merely as an estimate.

As a final note, there exist some reports of time-domain models for the transit-time effect. In some cases, this is a restatement of the $\alpha \rightarrow \infty$ condition leading to equation 2.17 [48, 45] and is similarly applicable to other devices functioning on the basis of carrier delay, for example the IMPATT diode. In other cases [55], the model is phenomenological and constructed in order to fit a particular device structure.[1] In such a case, the model often behaves incorrectly in certain limits ($\alpha \rightarrow \infty$, $\alpha \rightarrow 0$) due to simplifying assumptions and should be used with caution if applied to a different device structure.

2.3.3 Diffusion time dependence

In all recent textbooks and publications, to the author's knowledge, the subject of frequency-dependence due to diffusion delay is omitted. Qualitatively, the time constants related to transit delay are

$$\omega_{b,drift} \propto \frac{v}{L} \qquad \omega_{b,diff} \propto \frac{D}{L^2} \tag{2.19}$$

The diffusion bandwidth has a quadratic dependence on distance, and for a comparable distance and generic material parameters for electrons ($v = 10^7$ cm/s, $D = 40$ cm^2/s) the drift bandwidth is an order of magnitude larger. Therefore, most texts concentrate on high-performance (bandwidth) vertical devices where it is possible to avoid the issue of diffusion technologically. The QNR layers may be very thin, resulting in little external photocurrent from those regions, or alternatively (as is done in the case of the vast majority of commercial high-performance devices) p and n layers may be fabricated from a material that is transparent to the wavelength of interest such that all incident light is absorbed in the depleted i region. For example, a detector for 850 nm application may have a p-$Al_{0.1}Ga_{0.9}As$:i-GaAs:n-$Al_{0.1}Ga_{0.9}As$ structure, and a detector for a 1.55 μm application typically has a p-InP:i-$In_{0.53}Ga_{0.47}As$:n-InP structure [55]. In the case of integrated photodetectors, particularly ones with lateral geometry as shall be described in later chapters, it is not generally possible to ensure that the entire absorbing region is depleted and conversely that non-depleted material is also not absorbing, and therefore carrier diffusion must be considered as

[1]It should be noted that the reference [55] is a very important publication in the field, but features several typographical errors in the model development and analysis section, as well as an incorrect normalisation.

a strong limiter of device bandwidth.

The derivation of frequency behaviour below is similar to the method of Sawyer and Rediker [56], although it was conceived independently. There are several important differences, in particular the derivation of the frequency-dependent form of the solution as a special case of the DC solution with a modified coefficient, and the discussion of diffusion in the case of uniform generation.

In the following analysis, the semiconductor material is assumed to be p type, and the results are identical for n type material with a suitable change in notation. Under the assumptions of Section 2.2 the one-dimensional continuity equation including generation in the bulk is

$$\frac{\partial n'}{\partial t} - D_e \frac{\partial^2 n'}{\partial x^2} + \frac{n'}{\tau_e} = g_L(x,t) \tag{2.20}$$

where n' is the excess minority electron density ($n - n_0$), τ_e is the minority electron lifetime in the bulk, D_e is the electron diffusion coefficient, and $g_L(x,t)$ is the net rate of generation of minority electrons due to photon absorption. Generation may be written in the customary fashion as a linear combination of DC and small-signal components. Examining only the AC component, generation and excess minority carriers may be written as complex exponentials

$$g_l(x,t) = G_0(x)e^{j\omega t} \quad , \quad n' = N'e^{j\omega t} \tag{2.21}$$

where G_0 is real and N' has a magnitude and phase. Substituting into equation 2.20

$$j\omega N'e^{j\omega t} - D_e \frac{\partial^2}{\partial x^2} N'e^{j\omega t} + \frac{1}{\tau_e} N'e^{j\omega t} = G_0(x)e^{j\omega t} \tag{2.22}$$

Dividing by the term $e^{j\omega t}$ and rearranging factors results in the continuity equation with a modified minority lifetime factor that includes a complex term.

$$\frac{d^2 N'}{dx^2} - \frac{N'}{D_e \tau_{eff}} = -\frac{G_0(x)}{D_e} \quad ; \quad \tau_{eff} = \frac{\tau_e}{1 + j\omega\tau_e} \tag{2.23}$$

For $\omega \rightarrow 0$, equation 2.23 becomes exactly equation 2.20 without the time differential terms. Therefore, to derive the frequency dependence of carrier diffusion it is sufficient to solve the quasi-static

continuity equation and then substitute the new effective minority carrier lifetime in the place of the quasi-static minority carrier lifetime.

Uniform generation

The first one-dimensional problem of interest is shown in Figure 2-6.

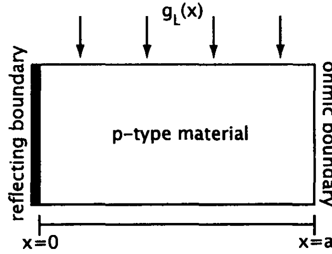


Figure 2-6: Schematic representation of the one-dimensional carrier diffusion problem with uniform generation.

In the case of light incident perpendicular to the direction of minority carrier diffusion, the absorption is decoupled from carrier flow and therefore may be treated as a constant ($g(x) = G_L$). The ohmic and reflecting boundary conditions on the minority carrier distribution correspond to a SCR edge at negative bias ($|V| \gg kT/q$) and majority carrier contact boundary conditions respectively. Specifically

$$n'(a) = 0; \quad \text{negative bias SCR boundary} \tag{2.24}$$

$$\frac{dn'}{dx} = s_1; \quad s_1 \rightarrow 0 \text{ for a reflecting boundary} \tag{2.25}$$

As shown earlier, the frequency-dependent solution may be written directly by substitution of an effective lifetime parameter into the quasi-static solution. Equation 2.20 may be rewritten without the time-derivative term and with some rearrangement for cleaner notation.

$$\frac{d^2n'}{dx^2} - \frac{n'}{L_e^2} = -\frac{g_L(x)}{D_e} \quad ; \quad g_L(x) = G_L,\ L_e \equiv \sqrt{D_e\tau_e} \tag{2.26}$$

This equation is solved by finding a homogeneous and a particular solution.

$$n' = Ae^{x/L_e} + Be^{-x/L_e} + G_L\tau_e \tag{2.27}$$

The boundary conditions are applied and the expression is simplified, resulting in the minority electron profile under uniform generation.

$$n'(x) = G_L\tau_e - \frac{G_L\tau_e \cosh(x/L_e)}{\cosh(a/L_e)} \tag{2.28}$$

The external photocurrent is the minority electron diffusion current at the SCR boundary.

$$J_{photo} = qD_e \frac{dn'}{dx}\bigg|_{x=a} = -qG_L L_e \tanh(a/L_e) \tag{2.29}$$

Considering the properties of the hyperbolic tangent, equation 2.29 is intuitively logical for the limits of sample size much greater than or much smaller than the minority carrier diffusion length.

$$J_{photo} \approx -qG_L L_e \quad ; a \gg L_e \tag{2.30}$$

$$J_{photo} \approx -qG_L a \quad ; a \ll L_e \tag{2.31}$$

The frequency-dependent solution is given by substituting an effective minority carrier lifetime as in equation 2.23

$$J_{photo}(\omega) = -q\frac{G_L L_e}{\sqrt{1+j\omega\tau_e}} \tanh\left(\frac{a}{L_e}\sqrt{1+j\omega\tau_e}\right) \tag{2.32}$$

Equation 2.32 is significant as it shows that the frequency decay of diffusion photocurrent is different than that of a single pole (e.g. RC) system: 10 dB/decade rather than 20 dB/decade. The first factor represents the decay for $a \gg L_e$, therefore the characteristic frequency is $\omega = 1/\tau_e$. The second factor contributes a correction as a approaches and becomes smaller than L_e (i.e. the actual carrier diffusion distance is reduced from L_e to a). This behaviour is shown in Figure 2-7.

The 3dB frequency for $a/L_e \gg 1$ is $\omega_{3dB} \approx 1.7/\tau_e$. As a is decreased, the carrier diffusion time also decreases, and for $a/L_e \ll 1$ the result is $\omega_{3dB} \approx 2.6D_e/a^2$. The difference in the prefactor comes from the shape of the tanh correction factor.

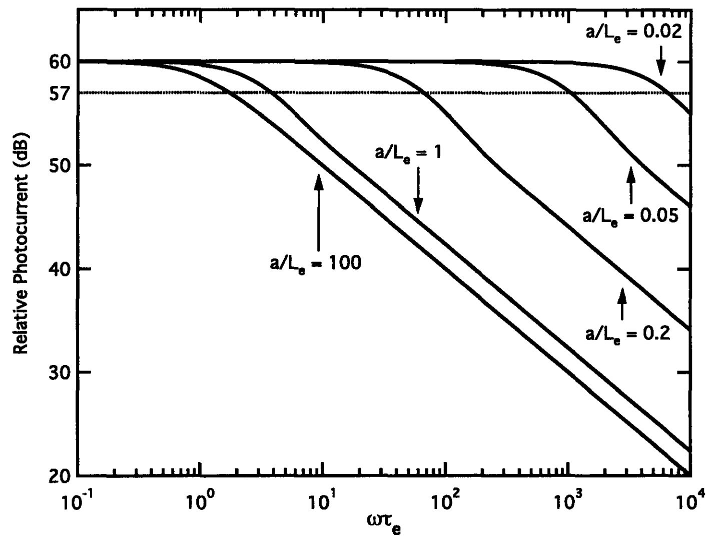


Figure 2-7: Plot of J_{photo} according to equation 2.32 for an arbitrary G_L as a function of $\omega\tau_e$ for different ratios a/L_e. The 3-dB frequency is indicated by the dotted line.

Spatially varying generation

The first one-dimensional problem of interest is shown in Figure 2-8. This solution to this problem will be applicable to the frequency dependency of photocurrent when carriers are generated outside of the SCR in a vertical p-i-n device.

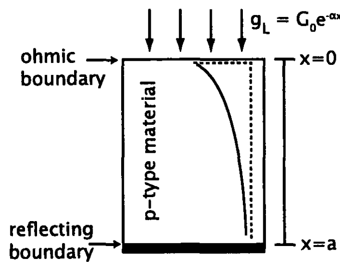


Figure 2-8: Schematic representation of the one-dimensional carrier diffusion problem with varying generation.

The solution method is the same as in the uniform case, but with more involved algebra due to the additional x dependence of the generation rate. The generation factor G_0 may be related to the incident intensity (photons/area^2) by the relationship $G_0 = \alpha I_0$. For vertical illumination with a top-side ohmic boundary as shown in Figure 2-8 and $\alpha^2 \neq 1/L_e{}^2$, the solution[2] for excess carrier

[2]The particular solution for $\alpha^2 = 1/L_e{}^2$ is $\frac{G_0}{2\alpha D_e} x e^{-\alpha x}$.

concentration is

$$n' = Ae^{x/L_e} + Be^{-x/L_e} - \frac{G_0}{D_e} \frac{L_e^2}{\alpha^2 L_e^2 - 1} e^{-\alpha x} \tag{2.33}$$

Applying the boundary conditions and following simplification the expressions for minority electron distribution and diffusion current at the ohmic boundary are

$$n'(x) = \frac{G_0}{D_e} \frac{L_e^2}{\alpha^2 L_e^2 - 1} \frac{\cosh \frac{a}{L_e}(\cosh \frac{x}{L_e} - e^{-\alpha x}) - \sinh \frac{x}{L_e}(\sinh \frac{a}{L_e} + \alpha L_e e^{-\alpha x})}{\cosh \frac{a}{L_e}} \tag{2.34}$$

$$J_{photo} = qD_e \left. \frac{dn'}{dx} \right|_{x=0} = \frac{q\alpha I_0 L_e}{\alpha^2 L_e^2 - 1} \frac{\alpha L_e \cosh \frac{a}{L_e} - \sinh \frac{a}{L_e} - \alpha L_e e^{-\alpha a}}{\cosh \frac{a}{L_e}} \tag{2.35}$$

The interpretation of equation 2.35 is not immediately obvious, but it may be simplified in several intuitively useful limits.

$$L_e \to \infty \ : \ J_{photo} \approx -qI_0(1 - e^{-\alpha a}) \tag{2.36}$$

$$1/\alpha \to 0 \ : \ J_{photo} \approx 0 \tag{2.37}$$

$$a \gg L_e, a \gg 1/\alpha, L_e \gg 1/\alpha \ : \ J_{photo} \approx \frac{-qI_0}{\alpha L_e} \tag{2.38}$$

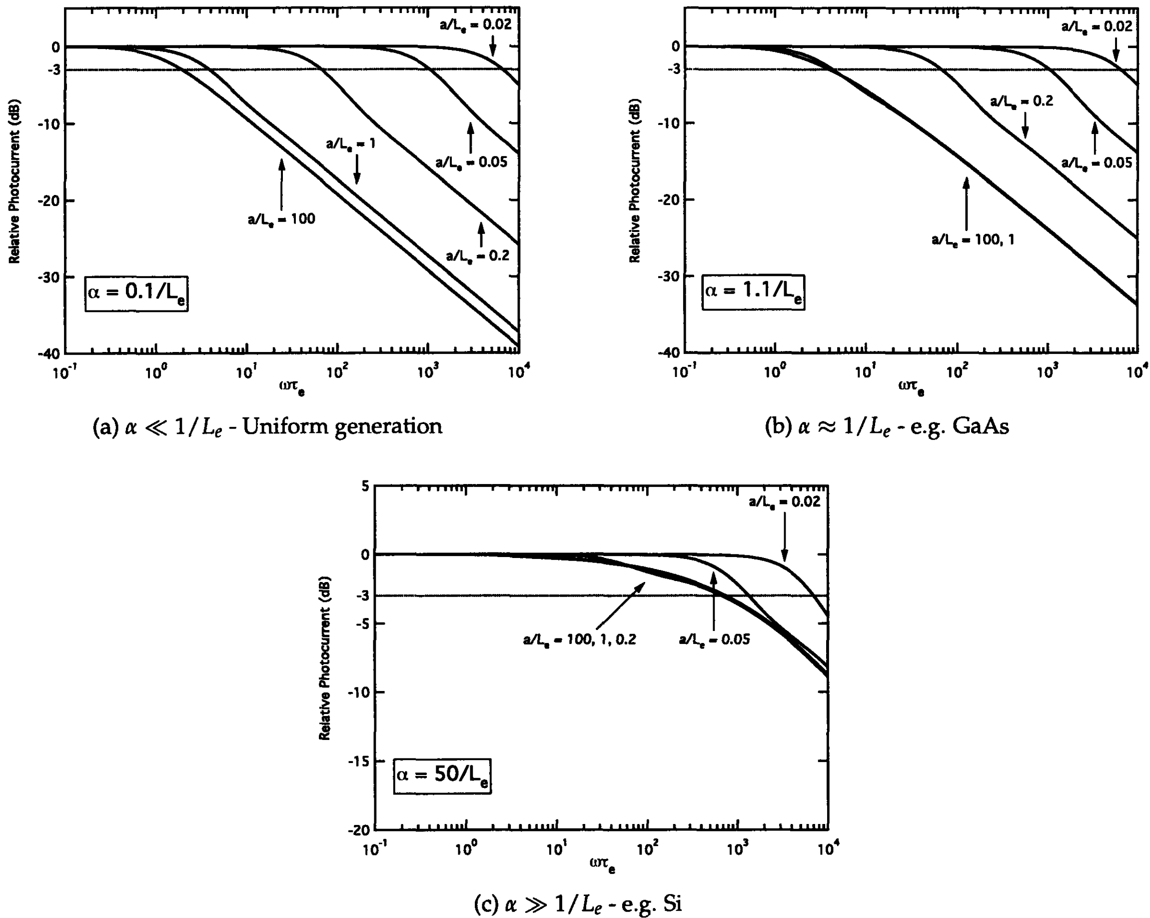
$$a \gg L_e, a \gg 1/\alpha, L_e \ll 1/\alpha \ : \ J_{photo} \approx -qI_0 \alpha L_e \tag{2.39}$$

The frequency-dependent expression for the photocurrent density is determined by substituting an effective minority carrier lifetime into equation 2.35. In this case, the lifetime term is a component of the minority carrier diffusion length.

$$J_{photo} = \frac{qI_0 \alpha^2 L_{eff}^2}{\alpha^2 L_{eff}^2 - 1} \frac{\cosh \frac{a}{L_{eff}} - \frac{1}{\alpha L_{eff}} \sinh \frac{a}{L_{eff}} - e^{-\alpha a}}{\cosh \frac{a}{L_{eff}}} \quad ; L_{eff} = \frac{L_e}{\sqrt{1 + j\omega \tau_e}} \tag{2.40}$$

The shape of expression 2.40 is not intuitively obvious. The properties of interest in this expression include bandwidth and the slope of response decay beyond the corner frequency. The dependencies are shown for a several useful cases in Figure 2-9.

Figure 2-9 exhibits very interesting behaviour. It should be noted that the plots were prepared by assuming $L_e = 1$ and scaling all other factors in relation to that value. This therefore implicitly defines the quantity D_e. Since the x-axis is a unitless quantity $\omega \tau_e$ and $L_e \equiv \sqrt{D_e \tau_e}$, care must be taken if interpreting the plots to extract an absolute number for ω_b. In the limit $\alpha \ll 1/L_e$,



(a) $\alpha \ll 1/L_e$ - Uniform generation

(b) $\alpha \approx 1/L_e$ - e.g. GaAs

(c) $\alpha \gg 1/L_e$ - e.g. Si

Figure 2-9: Plots of expression 2.40 for varying $\omega\tau_e$ for various unitless ratios αL_e and a/L_e. The 3-dB levels is indicated by the dotted lines. The vertical scale of plot c is different than that of plots a and b.

carrier generation is approximately uniform and therefore Figure 2-9a is identical to Figure 2-7. For the case $\alpha \approx 1/L_e$, all conditions $a/L_e \gtrsim 1$ are approximately equal. This is logical since for this absorption case the majority of carriers are generated in approximately the first L_e of material, therefore the remaining bulk plays no role in the photocurrent. For smaller a, the generation over the shorter distance once again appears approximately uniform, such that the right-hand curves of Figure 2-9b are approximately identical to the right-hand curves of Figure 2-9a. A further important observation is that for larger a, the rate of decay beyond the bandwidth appears to be slightly less than the previously observed 10 dB/decade. This case is a realistic one, for example 850 nm light in GaAs where both the absorption and minority electron diffusion lengths are

approximately 1-2 μm (depending on the material quality and doping).

Figure 2-9c shows the case $\alpha \gg 1/L_e$. This is also a realistic case, for example 780 nm light in high-quality, lightly p-type silicon ($\alpha \approx 10\ \mu$m$, L_e \approx 100\ \mu m$)[57, 58]. In this case similarly as before, for very small a the generation in the material appears uniform producing a familiar curve. For $a \gtrsim 1/\alpha$, the curves are all identical and correspond to a diffusion approximately over a distance $1/\alpha$. However, an interesting feature of this relationship is that as the product αL_e increases, particularly to values $\gg 1$, the sharpness of decay around the bandwidth decreases. As seen in Figure 2-9c, the decay may fall to approximately 5 dB/decade for a physically realistic choice of parameters. This behaviour will be observed in the measurement of certain device structures in Chapter 5.

A final detail in the dynamic behaviour of diffusion photocurrent involves the direction of the problem statement. Figure 2-8 realistically applies to the case of the substrate side of a vertical junction; the top boundary is the edge of the SCR, and x increases downwards into the substrate. For the case of diffusion in the top QNR of a vertical device (i.e. the top boundary is the semiconductor surface, and the bottom is the edge of the SCR), the boundary conditions must be reversed. This problem is solved in the familiar fashion.

$$J_{photo} = \frac{qI_0\alpha^2 {L_{eff}}^2}{\alpha^2 {L_{eff}}^2 - 1} \frac{e^{-\alpha a} \cosh \frac{a}{L_{eff}} + \frac{e^{-\alpha a}}{\alpha L_{eff}} \sinh \frac{a}{L_{eff}} - 1}{\cosh \frac{a}{L_{eff}}} \quad ; L_{eff} = \frac{L_e}{\sqrt{1 + j\omega\tau_e}} \tag{2.41}$$

By inspection of the terms, it is apparent that the dynamic behaviour is unchanged (i.e. all terms with a frequency dependence are present in both equations 2.40 and 2.41). For $\alpha a \gg 1$ and $L_e \gg a$ the expressions converge to the same value, otherwise the DC value of the photocurrent is different as a function of the proportion of generated carriers that can reach the ohmic boundary. A further significant conclusion is that the diffusion current frequency response is independent of the generation gradient. This is a result of the initial assumptions of low-level injection and minority carriers travelling only by diffusion. At high illumination levels where $n' \gtrsim p_0$ for p-type material (and the converse for n-type), these assumptions are no longer valid, meaning that the continuity equation cannot be linearised (and is therefore generally solved numerically if an exact solution is required).

2.3.4 Parasitic Elements

The frequency dependence of the photocurrent in the p-n or p-i-n junction itself was discussed in the previous sections. Diode behaviour will additionally depend on details such parasitic circuit elements as shown in the schematic diagram of a vertical p-i-n diode, Figure 2-10.

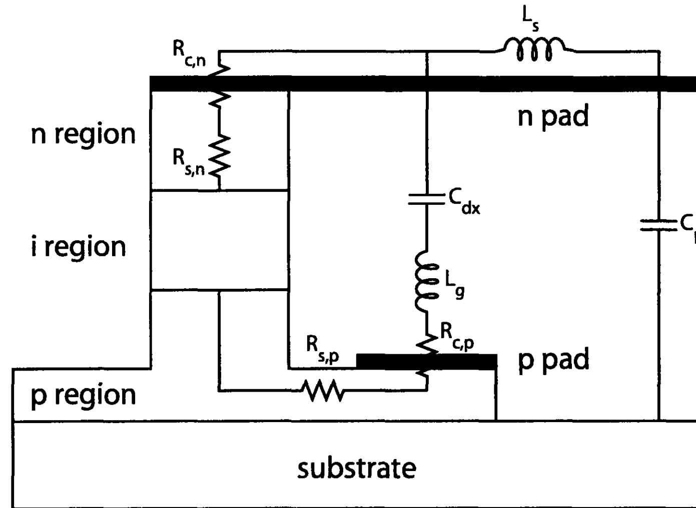


Figure 2-10: Equivalent circuit elements of a p-i-n diode. The circuit of Figure 2-5 may be inserted as the i region equivalent. The parasitic elements of a different geometry (e.g. lateral/MSM) are identical but not as clearly presentable graphically.

The parasitic elements are strongly dependent on material, processing, and layout parameters. Resistances $R_{s,n}$ and $R_{s,p}$ are the series resistances of the QNRs, given by the familiar expression $R = \frac{L}{A}\frac{1}{q\mu n_0}$. High performance III-V devices are generally epitaxially grown on a semi-insulating substrate with thin n and p regions. After the etch to define the diode structure, there remains only a very thin bottom QNR (p in the case of the schematic).[54] Therefore, the spreading/series resistance of the bottom layer will generally dominate in those structures. Resistances $R_{c,n}$ and $R_{c,p}$ are the metal-semiconductor contacts resistances. These values depend on the band structure and area at the interface. For example, in InP-InGaAs structures the contact to p-InP has a high resistivity and an InGaAs cap layer is used to reduce that effect. For small devices, the contact resistance may be an order of magnitude or more greater than the dominant series resistance.[55]. All of the resistances may be lumped into a composite resistance R_s. The inductances L_s and L_g are a strong function of mask layout. In general for properly designed microwave pads close to the

device of interest, these values are in the pH range and may be ignored (i.e. an effect may be seen above 50-75 GHz). Similarly, an integrated device would be excepted to have short connections to an associated amplifier/receiver such that this effect may be ignored. The capacitance C_{dx} is the capacitance between the coplanar n and p pads, generally on the order of 20-40 fF depending on the thickness of the metallisation and pad pitch. The capacitance C_p is the capacitance between the signal pad and the ground plane. For a III-V device grown on a semi-insulating substrate, this capacitance is very low as the substrate acts approximately as a dielectric with a thickness of several hundred μm. However, vertical devices intended for integration often use the substrate as one of the electrodes (for example, vertical Ge on Si p-i-n diodes [59]), in which case the pad capacitance dominates over other parasitic capacitances. Again, the short connection distance between an integrated device and receiver would strongly reduce this effect. Under the assumption of negligible inductance, the capacitances may also be lumped into a composite capacitance C_{px}.

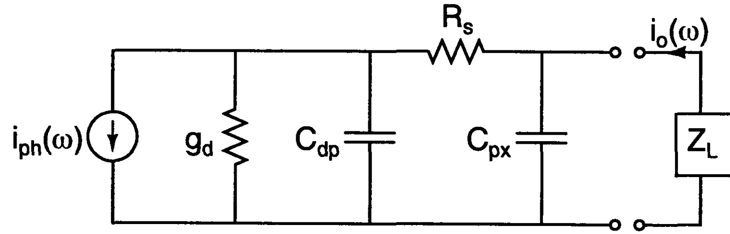


Figure 2-11: Photodetector incremental circuit schematic including device parasitic elements.

The incremental circuit model of the photodiode including parasitic resistance and capacitance elements is shown in Figure 2-11. In general the parallel resistance of the junction is high. Assuming $g_d \approx 0$, the transfer function of the network is easily derived.

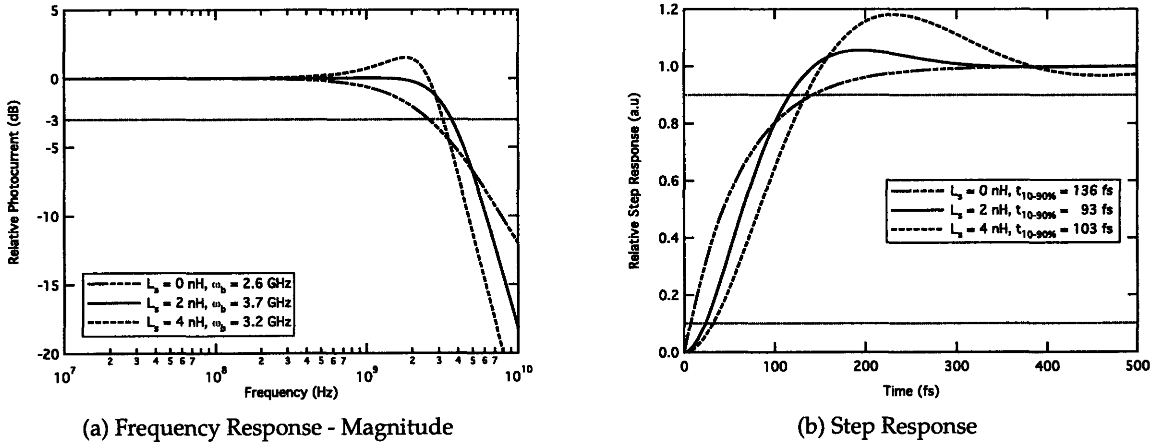
$$H(\omega) = \frac{i_o(\omega)}{i_{ph}(\omega)} = \frac{1}{1 - \omega^2 C_{dp} C_{px} R_s Z_L + j\omega(C_{px} Z_L + C_{dp}(Z_L + R_s))} \tag{2.42}$$

The bandwidth may be estimated, assuming that the two poles of the system function are widely spaced, using the method of open-circuit time constants

$$\omega_b = \frac{1}{\tau_b} \quad ; \quad \tau_b = \tau_{C_{dp}} + \tau_{C_{px}} = C_{dp}(R_s + Z_L) + C_{px} Z_L \tag{2.43}$$

Effect of Series Inductance

In an ideal case, such as the case of a detector integrated on-chip with other devices, the load Z_L will be purely resistive, i.e. $Z_L = R_L$. More often, however, the detector is connected to a package or another chip by wire bonds that contribute series inductance: $Z_L = R_L + j\omega L_s$. This may contribute to the system function by extending the observed bandwidth, as shown in Figure 2-12. For reference, 1 mm of 1 mil diameter bond wire contributes approximately 1 nH of inductance.



(a) Frequency Response - Magnitude

(b) Step Response

Figure 2-12: Plots of the system function response (equation 2.42) for different series inductance L_s at the load, and typical values $C_j = 1$ pF, $C_{px} = 40$ fF, $R_s = 10\ \Omega$, $R_L = 50\ \Omega$

The series inductance transforms the one real pole of the transfer function into two complex poles. As seen in Figure 2-12a, with increasing inductance the system response goes from over-damped to critically damped to underdamped behaviour. In this example, the bandwidth of the system increases by approximately 50% from the zero series inductance to the critically damped case ($L_s \approx 2$ nH). As the series inductance increases past the critically damped condition, a definite gain peak forms and the bandwidth decreases. The consequence of these effects is clearly seen in the system step response, Figure 2-12b with respect to the 10-90% rise time. This is a realistic metric for a transmission system where bits are transmitted as pulses of laser light "on" or "off". The addition of some series inductance decreases the rise time without introducing significant overshoot. As the systems enters the underdamped regime, the rise time is longer and there is significant overshoot corresponding to the gain peak in the frequency domain. This may significantly degrade performance in a transmission system, partly due to the slower rise time and also

due to the negative consequences of the overshoot on the dynamic range of the receiver.

Full Small Signal Model

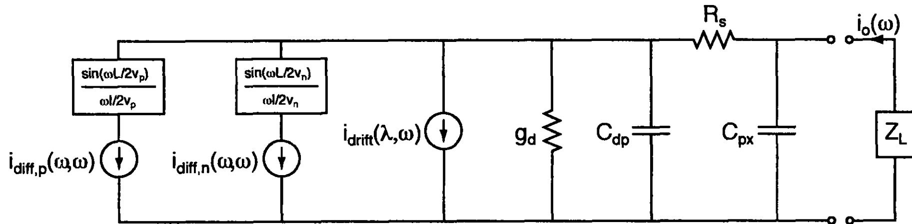


Figure 2-13: Full incremental photodiode circuit model

Figure 2-13 shows the full small signal circuit model of the photodiode including the three sources of photocurrent. Some of these sources may be equal to zero, depending on the physical structure of the device. The diffusion sources have their own frequency dependence, and then additionally a second dependence that represents the delay in crossing the SCR. This delay is identical to the case of strong absorption at one side of the SCR, therefore the dependence is that of equation 2.18, with the appropriate hole or electron velocity depending on the carrier type. In general, the velocity may be treated as approximately the saturation velocity.

An important characteristic of the full model is that the frequency behaviour of the current sources is dependent only on the one-dimensional (in the simple vertical junction case) or two-dimensional (in the case of a lateral device) geometry and material properties of the junction. However, the junction capacitance and series resistance varies with the device area as seen by incoming light.

2.4 Receivers

As shown schematically in Figure 2-3, a photodiode signal may be measured as a current shift relative to the case of a dark diode or as a shift of the voltage across the diode for a given current condition. It is obvious from the current-voltage relationship of the diode that the latter method is highly non-linear and depending on the receiver configuration may introduce an undesired large DC offset due to the flow of amplifier bias current through the diode. Current-mode monitoring avoids these problems. An excellent discussion of receiver design may be found in [60], and

some of the significant points are repeated here. A very simple example of a current-mode (or transimpedance) approach is shown in Figure 2-14.

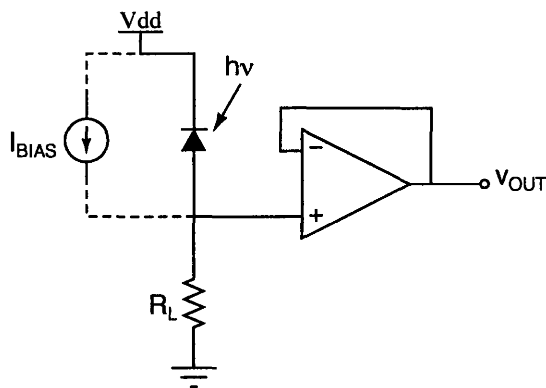


Figure 2-14: Basic receiver consisting of a load resistor and a voltage buffer. A bias current may be operating in parallel to adjust the DC offset of the output signal.

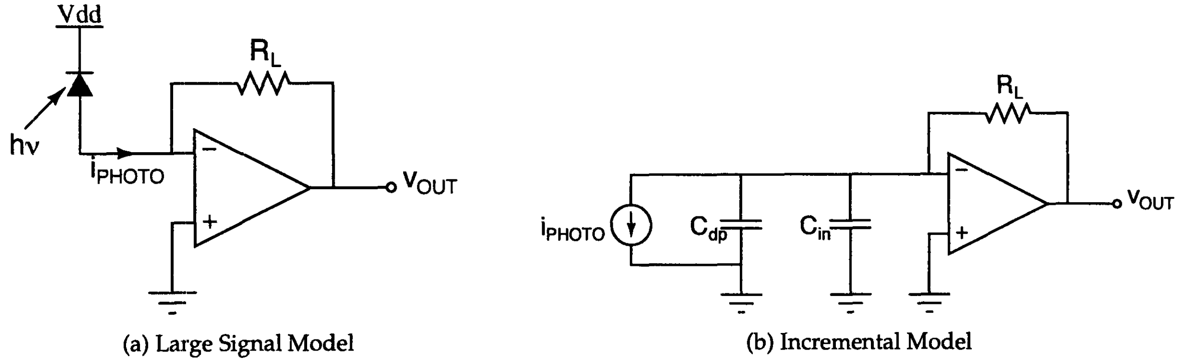
For the discussion of receivers, the simple small signal circuit of Figure 2-5 (i.e. no parasitic elements) will be used in order to simplify the derived expressions. The input resistances of the amplifiers will be ignored, but the input capacitances may be significant in determining the frequency response. In the case of the receiver of Figure 2-14, the signal transimpedance and bandwidth may written by inspection, where C_{in} is the amplifier input capacitance:

$$g_{TIA} \equiv \frac{v_{OUT}}{i_{PHOTO}} = R_L \tag{2.44}$$

$$\omega_b = \frac{1}{R_L(C_{dp} + C_{in})} \tag{2.45}$$

The transimpedance-bandwidth product is therefore fixed at $1/(C_{dp} + C_{in})$, and this limit is a consequence of the signal voltage being developed over the diode equivalent capacitance. Other topologies exist that decouple the photocurrent from the signal voltage at the photodiode.

The receiver of Figure 2-15 has the same transimpendance as that of 2-14, under the assumption of an ideal operational amplifier. The effective load resistance seen by the photodiode is the feedback resistance divided by the open-loop gain of the amplifier, which is infinite in the ideal



(a) Large Signal Model (b) Incremental Model

Figure 2-15: Transimpendance amplifier with a feedback resistor, decoupling the voltage across the diode and the photocurrent.

case.

$$g_{TIA} = R_L \tag{2.46}$$

$$\omega_b = \frac{1}{R'_L(C_{dp}+C_{in})} \quad ; \quad R'_L = \frac{R_L}{A_{OL}} \to 0 \tag{2.47}$$

$$g_{TIA}\omega_b = \frac{A_{OL}}{C_{dp}+C_{in}} \to \infty \tag{2.48}$$

Therefore for this topology in the ideal case, the gain and bandwidth are decoupled, and the bandwidth is determined by parasitic elements. However, for a real feedback configuration, the amplifier is generally designed with a dominant pole to ensure stability [61]). For most practical cases, the operational amplifier rolls off as approximately a single-pole response until at least the unity-gain crossover frequency ω_c. Therefore it is convenient to make an approximation $A_{OL}(\omega) = \omega_c/\omega$. Therefore at the system bandwidth ω_b, one finds $A_{OL} \approx \omega_c/\omega_b$. Substituting this expression into equation 2.47, the actual bandwidth may be found to be the geometric mean of the amplifier unity-gain crossover frequency and the simple TIA bandwidth of equation 2.45.

$$\omega_b = \sqrt{\frac{\omega_c}{R_L(C_{dp}+C_{in})}} \tag{2.49}$$

As long as $\omega_c > \omega_b$ from equation 2.45, the bandwidth is increased by the feedback topology, typically by a factor of 10 to 100. There are many variations on this general topology, but the principle of decoupling of gain and bandwidth is similar in all cases. The main differences are

in the feedback system (for example to reduce DC offset), the operational amplifier design, and the connection of the photodiode to the amplifier (for example, fully differential techniques are interesting for the ability to reject external interference [62]).

2.4.1 Amplifier Stability

There are two important parasitic effects that may have a detrimental effect on the stability of the transimpedance amplifier. First, the feedback resistor R_L in fact appears as an impedance $R'_L = \frac{R_L}{A_{OL}(\omega)}$. Since R'_L increases with frequency, it is effectively an impedance consisting of a resistor in series with an inductor. Under the proper set of values, this may form an LC tank at the input of the amplifier that causes oscillations. If the characteristic frequency is not above the unity-gain crossover of the amplifier, the oscillations may be sustained around the feedback loop.

The second effect is caused by the photodiode capacitance. In the feedback loop from amplifier output to inverting input, the feedback resistor R_L sees the capacitance $C_{dp} + C_{in}$, thereby forming a low-pass filter. If this filter's characteristic frequency is near the unity-gain crossover frequency of the op amp, it severely degrades the phase margin of the loop gain causing ringing and potentially oscillation, as shown in Figure 2-16.

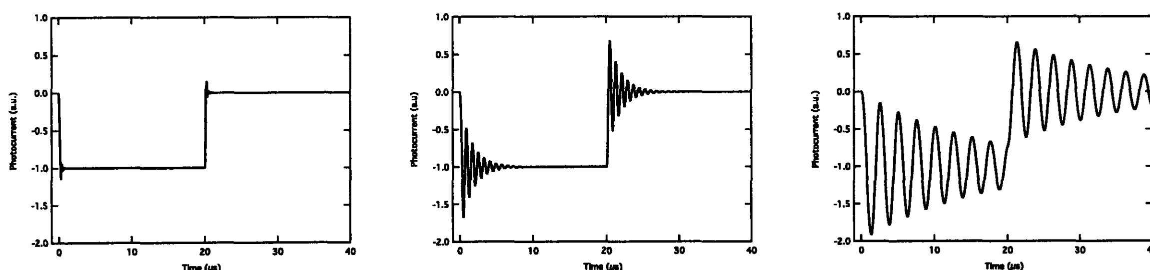


Figure 2-16: SPICE simulation example of a TIA circuit with a single-pole operational amplifier. As the photodiode capacitance is raised, the phase margin is decreased causing overshoot/ringing of the output voltage.

There are several strategies for compensating the TIA under these conditions. In fact, parasitic capacitance in parallel with the feedback resistor often provides sufficient compensation, but at the cost of decreased bandwidth. A full discussion of this issue is outside the scope of this thesis, however this result further motivates the need to decrease photodiode capacitance for high bandwidth applications.

2.5 Noise

The only source of noise in photodiodes in shot noise, caused by the fact that current flowing across a junction is not smooth, but is comprised of individual electrons arriving at a random time. This gives rise to broadband white noise that increases with average current. The spectral density of shot noise associated with a current I is given by

$$I_{sh}^2 = 2qI \tag{2.50}$$

This density is an RMS spectral density, therefore the actual noise current is obtained by integrating over the bandwidth of interest and taking the square root. One mA of current has about 20 pA/$\sqrt{\text{Hz}}$ of shot noise, modelled as a current source in parallel with the junction in the case of a photodiode. It is desirable for a photodiode to have low dark (reverse) current to minimise shot noise.

Other sources of noise are associated with the receiver/amplifier circuit and a full discussion of these effects is beyond the scope of this work. However, it is useful to compare the two amplifier topologies of the previous section in reference to their noise performance. Assuming that the operational amplifier in both circuits (Figures 2-14 and 2-15) is the same, the only difference in noise is the placement of the load resistor. The resistor exhibits a thermal noise density

$$V_{th}^2 = 4kTR_L \tag{2.51}$$

In the case of Figure 2-14, the voltage noise is added to the photocurrent at the input to the amplifier/buffer. In the case of Figure 2-15, the voltage noise is added at the output node, but is reduced by a factor of the gain at the input node. The full behaviour is more complicated and the reader is advised to consult [60] for a thorough derivation. As a gross approximation, it is clear that for circuits with the same bandwidth, the noise in the second case is much lower than that of the first case due to the larger load resistor. An example specification of input-referred noise current for a modern discrete transimpedance amplifier is 1.1 μA for a 10 Gbps bandwidth [63]. Such a value is roughly equivalent to the shot noise generated by 0.6 mA of current over the same bandwidth.

Chapter 3

Photodetector Frequency Response Measurement

Accurate measurement of photodetector frequency response requires a robust measurement apparatus due to the many sources of noise, interference, and parasitic elements generally present, as well as the low levels of the photocurrent signal. Measurement techniques may be separated into time-domain and frequency-domain classifications, and the relative advantages and disadvantages of the two approaches will be discussed. For this research, a frequency-domain apparatus was designed, built, and used for all device measurements presented in later chapters. This apparatus will be discussed in detail. Finally, it is apparent upon scanning the literature of photodetector and modulator measurements that many publications include data that, upon further analysis, may be found to be incomplete or dominated by a parasitic or interference effect. Select cases of this sort will be shown in the appropriate sections.

Time-domain measurements may be further separated into two distinct types of measurement: impulse response and eye diagram measurement. The former is useful in measuring very high bandwidth devices, by applying a short excitation pulse and directly extracting the system function in the time domain. The latter (more precisely a repeated step response measurement with infinite measured trace persistence) is generally used to measure the time-domain characteristics of a transmission system rather than an individual device. Both of these measurements are closely linked to the geometry and parasitic elements of the devices and apparatus itself, which may lead

to an incorrect interpretation of results.

3.1 Impulse Response

The impulse response of a system is directly related to its system function in the frequency domain. The system function is written as $h(t)$ or $H(j\omega)$ where the two functions are related by the Fourier transform

$$h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(j\omega) e^{j\omega t} d\omega \tag{3.1}$$

If the system is excited by an impulse excitation at time zero, $x(t) = A\delta(t)$, the output is a convolution in the time domain $y(t) = h(t) \otimes x(t)$ or a multiplication in the frequency domain. The transform of the impulse in the time domain is a constant in the frequency domain, therefore the output $y(t)$ is a scaled version of the system impulse response $h(t)$.

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(j\omega) X(j\omega) e^{j\omega t} d\omega \quad ; \quad x(t) = A\delta(t) \xrightarrow{\mathcal{F}} A \tag{3.2}$$

$$y(t) = \frac{A}{2\pi} \int_{-\infty}^{\infty} H(j\omega) e^{j\omega t} d\omega = Ah(t) \tag{3.3}$$

For a real signal, the impulse excitation will not be infinitely sharp and will therefore have some frequency components, more precisely it will be approximately a sinc in the frequency domain with a characteristic frequency related to the inverse of the pulse length. Equation 3.3 is valid as long as the bandwidth of the pulse is much greater than the bandwidth of the system function (i.e. the width of the pulse is much shorter than the width of the system impulse response). In practice, for photodetectors this is accomplished by using a pulse from a mode-locked laser with a duration in the ps or sub-ps regime, corresponding to a characteristic frequency in the hundreds of GHz whereas the device under measurement may be expected to have a bandwidth at most in the tens of GHz in the majority cases. A schematic diagram of a typical apparatus is shown in Figure 3-1.

The electrical pulse measured from the photodetector is Fourier transformed to produce a frequency response characteristic at different reverse biases as set through the DC arm of the bias T. The operation of this apparatus is relatively simple, and the difficulty of the measurement lies

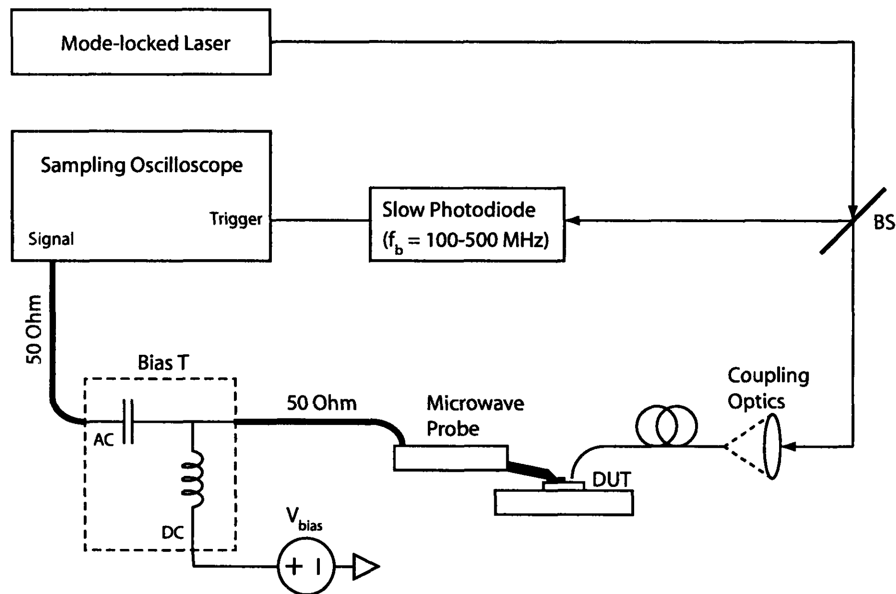


Figure 3-1: Schematic representation of an impulse response measurement system. The figure as drawn includes free-space coupling of a mode-locked laser beam, such as Ti:Sapphire operating around 850 nm. The same general principle applies to fibre coupling at the laser output, including then a fibre splitter rather than the beam splitter.

in the details of the connections.

The mode-locked laser should produce short pulses (ps length or shorter, depending on the laser design) at a relatively slow repetition rate (typically on the order of 10-100 MHz). The slow photodiode is used to trigger the sampling oscilloscope, synchronising the laser repetition rate to the captured waveform. It is often desirable to amplify the slow photodiode signal in order to ensure a dependable trigger lock, depending on the beam splitter ratio and the beam intensity.

The mode-locked laser beam may be coupled to the DUT by a number of means. The simplest is a free-space method using a lens to focus the beam, with the DUT placed at the focal point. Such a configuration may in practice result in a relatively large spot size, depending on the geometry of the apparatus. An alternative free-space method involves the use of a microscope objective to reduce the spot size of the beam. Such a system may be accompanied by a CCD camera and a beamsplitter to allow viewing of the DUT through the objective, making alignment more reliable. The beam may also be coupled into a fibre as suggested in Figure 3-1. This requires more precise optical alignment than the aforementioned methods in order to couple into the fibre with high efficiency. Because of the very high bandwidth of the optical pulse, multi-mode fibre may not be used

due to the high modal dispersion that would drastically broaden the pulse, making the alignment more difficult. This arrangement offers the greatest flexibility in terms of the arrangement of the rest of the apparatus. In particular the electrical connection between the probe and oscilloscope should be well-controlled as will be discussed below. If the mechanics of optical and electrical connections may be decoupled on the lab bench by the use of fibre, there is a great advantage not only in aesthetics of the apparatus, but also in the signal integrity.

The electrical connection to the DUT is the source of the majority of error in the system. In general, it is difficult to calibrate for the parasitic elements of the probe, cable, and bias T. There are generally three approaches to the calibration problem. The first is an approximate method: if the probe, bias T, and cables are chosen such that have a much higher bandwidth (at least twice as high) than the expected bandwidth of the DUT, then the time constant of the pulse response should be approximately an RMS combination of the time constants of the device and the other elements. The time constant of the device alone may thus be extracted based on the manufacturer's specifications of the other elements. However, the performance of a given element is generally different (and superior) to its specifications, therefore this method is only useful as a rough guide for further analysis. It may be equally valid to ignore the contribution of the probe and bias T, and to concentrate solely on the 50 ohm coaxial cable. Cables have a loss that increases with frequency and may be summarised in units of dB/length/GHz. Some reports specify the losses that have been de-embedded or assumed for the electrical side of the impulse measurement, for example 1 dB/10 GHz for a 12-inch cable [42]. However, such statements may be viewed with some scepticism without a full listing of the part numbers of the elements in the measuring apparatus, in the particular case of the citation in light of the fact that low-cost cables with a loss of 0.5 dB/10 GHz are readily commercially available (e.g. Pasternack Enterprises 50 GHz flexible assemblies, model no. PE35611-X).

One method for actual measurement and compensation of the electrical side of the measurement apparatus is time-domain reflection (TDR). In this method, a pulse is generated in the head of the sampling oscilloscope and reflections from the line are monitored by the sample head for a set of known terminations (i.e. termination at the reference plane, or the connection to the DUT). Attenuation and reflections due to impedance mismatch connectors are monitored and transformed to create a frequency-domain model of the system connected to the oscilloscope. Measured traces

are also transformed, corrected by the system model, and inverse-transformed to produce a corrected time-domain trace. This method has improved greatly in recent years; an Agilent 86100C with software option 202 and a 54754A TDR plugin is now capable of producing an S11 of a given network identical to that produced by a vector network analyser for frequencies up to approximately 10 GHz [64]. This calibration method currently has two limiting factors. First, most TDR heads (e.g. Agilent 54754A, Tektronix 80E04) produce reliable calibrations to only approximately 20 GHz. This frequency is limited by the width of the generated TDR pulse. This area has also seen recent progress, as Picosecond Pulse Labs has developed the 4020 TDR source enhancement module that decreases the pulse width from 35 to 9 ps, allowing for greater calibrated bandwidth [65]. Second, TDR is most often used with connectorised systems, so calibration substrates must in the future become part of the known standards base for use with on-chip probing. This has begun with the introduction of trimmed resistor standard substrates by Agilent Technologies.

The third, and most reliable, method of electrical connection calibration is the use of a Vector Network Analyser (VNA). The VNA has a higher dynamic range than TDR measurements and is considered the standard for S11 or S12 calibrations. There is a long history of on-wafer probing in this field, and thus a full calibration is possible from instrument port to substrate. The method for correction is the same as in TDR, but must be carried out off-line from the measurement: the frequency response of the connections is measured by the VNA for several known terminations, building an equivalent system function model. The impulse response measurement is transformed to the frequency domain and corrected using this measured connection system function.

3.1.1 Impulse Response Examples

The following three figures are real impulse response measurements on different detectors with different experimental apparatus. They illustrate several different bandwidths and modes of operation.

Figure 3-2 shows the impulse response of a Ge-on-Si vertical photodiode. The optical pulse was generated by a 1040 nm fibre mode-locked laser and free-space coupled to the detector through a variable attenuator and a focusing lens. The pulse was recorded by a Tektronix 11801C sampling

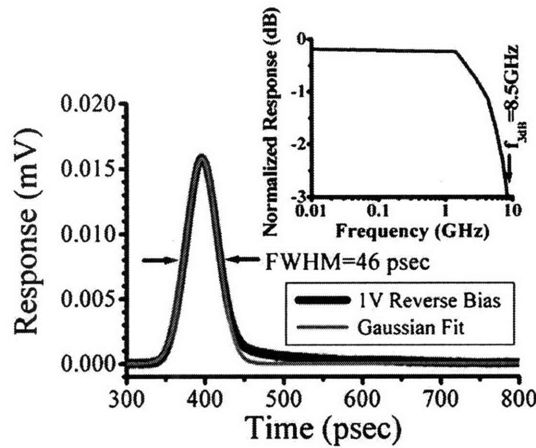


Figure 3-2: Vertical Ge-on-Si photodiode illuminated by 1040 nm 1 ps-pulse. Measured by the author and Jifeng Liu, reported in [59].

oscilloscope with a 50 GHz SD-32 sampling head. The remaining elements of the apparatus: GSG probe - 50 GHz Cascade Microelectronics I50-A-GSG-100 with 2.4 mm connector, bias T - 50 GHz Picosecond Pulse Labs 5542-202 with 2.4 mm connector, 2.4 mm to 3.5 mm female-female precision adapter, and Pasternack Enterprises cable with 3.5 mm connectors. The trace after the primary pulse shows no additional pulses, indicating no reflections from the electrical connections. The pulse is also entirely positive, indicating no dominant second-order behaviour (i.e. negligible parasitic inductance). The rise and fall of the pulse are approximately symmetrical, meaning that the device response is dominated by drift delay through the SCR. A capacitance or diffusion-limited response would be asymmetrical, with a slower decay on the falling side of the pulse.

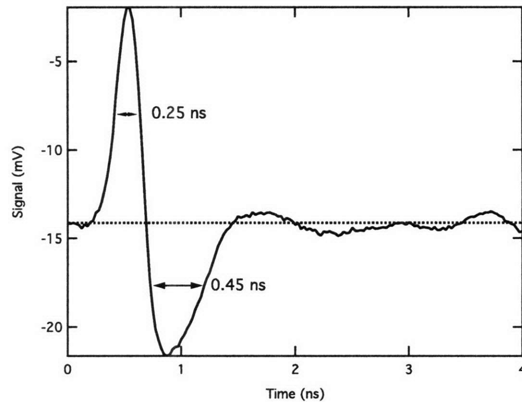


Figure 3-3: Lateral GaAs photodiode fabricated in a standard VLSI process, illuminated by 850 nm sub-ps pulse. Measured by the author, reported in [66]. The device bandwidth was measured independently to be 2 GHz, and the pulse response is dominated by series inductance.

Figure 3-3 shows the impulse response of a GaAs lateral p-i-n photodiode, discussed in detail in section 5.1. The optical pulse was generated by a Ti:Sapphire mode-locked laser and free-space coupled to the detector through a variable attenuator and a microscope objective lens. The pulse was recorded by a Tektronix 11801C sampling oscilloscope with a 20 GHz SD-24 sampling head. The remaining elements of the apparatus: GSG probe - 20 GHz Picoprobe GSG probe with 3.5 mm connector, bias T - 12 GHz Picosecond Pulse Labs 5575A with SMA connector, and Pasternack Enterprises cable with SMA connectors. The device bandwidth was measured independently by the frequency-domain method as discussed in section 5.1.4. The device was packaged for measurement by die-attach and wire-bonding to a glass substrate which had been patterned with gold probe pads. The trace clearly shows a second-order behaviour. Device capacitance was approximately 1 pF and series inductance was estimated at 3-4 nH, therefore corresponding to the situation simulated in section 2.3.4. The trace following the pulse exhibits oscillations indicating electrical reflections in the cable/connectors. This is an example of an unusable impulse response, as the result is dominated by packaging and instrument parasitics.

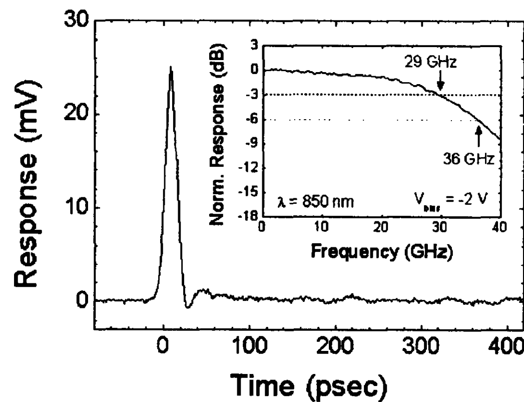


Figure 3-4: Lateral Ge-on-SOI photodiode reported by Dehlinger et al. [40]

Figure 3-4 shows the impulse response of a Ge lateral p-i-n photodiode grown on ultrathin SOI (also called Ge-on-I) [40, 41, 42]. The optical pulse was generated by a Ti:Sapphire mode-locked laser, coupled from free-space into SMF, and coupled to the detector from the SMF. The beam was attenuated through a 10:90 beamsplitter, with the weaker beam going to the trigger photodetector. Further elements of the apparatus were not specified, but from the description it may be deduced that an Agilent 86100 series oscilloscope with a 86117A 70 GHz remote sampling

head was used to capture the impulse signal. This particular sampling head is very interesting as it is connected to the oscilloscope frame by an extension cable, and may therefore by directly attached to the output of the bias T with the oscilloscope standing some distance away. This is helpful in reducing cable and connector mismatch related errors. The pulse is approximately symmetrical indicating a drift delay time dominated frequency response. The falling edge of the pulse has a slight negative excursion. This may be an indicator of series inductance, though the authors do not acknowledge this possibility. If this is the case, then the bandwidth determined by FFT is slightly overestimated. There is some ripple following the pulse, particularly around 210 and 320 ps, likely indicating some reflection in the electrical connections. These reflections are clearly visible as a ripple overlaying the FFT response seen in the figure inlay. In sum, this is a rather clean impulse response and corresponding FFT that includes many minor features of the intrinsic and parasitic response. The general problem of electrical parasitic calibration is highlighted by the ambiguous negative excursion following the pulse.

3.1.2 Impulse Response Summary

Aside from the cost or availability of the required equipment, the impulse response measurement has both intrinsic positive and negative characteristics, as summarised below.

Positive:

- The bandwidth of the pulse is determined by the optical process of mode-locking. The pulses have much higher bandwidth than an electrical pulse followed by optoelectronic conversion (e.g. laser modulation). This allows for the measurement of detectors with very high bandwidths up to nearly 200 GHz [67] using advanced sampling techniques. Other methods may generally measure to approximately 40-50 GHz, such that the increased bandwidth is the most important advantage of the impulse measurement method.
- One calibration step is eliminated from the measurement. Since the length of the optical pulse is much shorter than the length of the resultant photocurrent pulse, it may be treated mathematically as a delta function and no further calibration is required. Other methods require a calibration of the electrical to optical conversion at the signal generator.

Negative:

- Calibration of the electrical connections on the receiver side of the apparatus may be difficult depending on the equipment available and the relative time constants of parasitic elements. Additionally, calibration is generally performed offline and with different pieces of apparatus, whereas parasitics may often be variable depending on the details of mounting and contacting of the sample.
- To minimise calibration errors, it is most desirable for the electrical connection components to have a higher bandwidth (as an approximate rule, at least twice as high).
- Dynamic range. The sampling oscilloscope appears as a 50 Ω load in the circuit, therefore the measured voltage signal is the instantaneous current multiplied by 50 Ω. The newest oscilloscopes (e.g. Agilent 86100C) allow for 1 mV/division resolution on screen, and therefore the optical pulse must be strong enough to generate at least a 4-5 mV pulse for clean data collection. However, depending on the geometry or material properties of the photodiode, such an optical pulse intensity may introduce higher-order effects, in particular a lowering of the electric field in the SCR as discussed in section 2.3.2. Therefore, there is a maximum optical intensity to maintain linear operation of the photodiode, which in some cases may be so low that it is not possible to capture a low-noise trace.

3.2 Eye Diagram Measurement

The eye diagram technique is a simple but powerful measurement system for assessing the data-handling ability of a digital transmission system. It is used to identify bandwidth limits, non-linearities, and mechanical errors in the design of an entire transmission system, rather than one particular element. In fact, the eye diagram presents a picture that incorporates the properties of all elements in the system, and it is very difficult to extract the properties of a particular element. Essentially the measurement presents a graphical representation of the signal to noise ratio (SNR) of the system, and not directly a representation of the system bandwidth. However, many reports of detectors, modulators, or combinations of an optoelectronic element with a driver or receiver present eye diagram measurements or the related (simpler) step response measurement as a proof

of device bandwidth. It is an important technique to understand, especially in light of its often inappropriate application. An example of an eye diagram is shown in Figure 3-5.

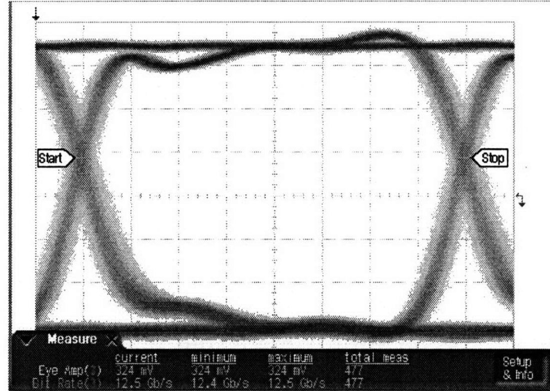


Figure 3-5: Eye diagram of a 12.5 Gbps NRZ 2^{31} PRBS signal generated by an Agilent N4906B BERT and received by an Agilent 86100C sampling oscilloscope with a 50 GHz 86117A electrical plugin. The generated and receiver were connected by a 50 cm cable with 2.4 mm connectors.

The eye diagram is an oscilloscope trace synchronised to the received signal by a trigger signal proportional to a multiple of the signal clock rate, where each triggered trace has infinite persistence. Therefore over the time of the measurement the constant levels, as well as the rising and falling edges of the received data are sampled and superimposed to form the eye diagram. The levels and transitions appear as qualitatively fuzzy lines due to the noise and jitter of the data stream. The generator produces a clock signal at a requested data rate, then encodes pseudo-random data onto the clock to simulate the transmission of arbitrary data through the digital link.

3.2.1 Bit Error Rate

The most common figure of merit for digital links is the bit error rate (BER)[68] . This is defined as the number of bit errors detected divided by the total number of bits sent over a specific time interval. Typical error rates for optical fibre telecommunication systems range from 10^{-9} to 10^{-15}. The SNR is related to the BER through the following expression.

$$\text{BER} = \frac{1}{2\pi}\int_0^\infty \exp\left(-\frac{x^2}{2}\right)dx \approx \frac{1}{2\pi}\frac{e^{-Q^2/2}}{Q} \tag{3.4}$$

The symbol Q represents the SNR for simplicity of notation. The approximate expression holds for BER less than about 10^{-3}. For reference, common values are $Q = 6$ for BER $= 10^{-9}$ and $Q = 7$

for BER = 10^{-12}. The value Q may be determined from the eye diagram.

$$Q = \frac{I_1 - I_0}{\sigma_1 + \sigma_0} \tag{3.5}$$

In this expression I_1 and I_0 are the average detected signal levels for 1 and 0 bits respectively, and σ_1 and σ_0 are the respective RMS widths of the 1 and 0 levels. In a digital transmission system, the decision threshold may be set at any given time within the eye opening, therefore Q is calculated for the position in time where the eye opening is most distinct. Modern BER testing equipment (abbreviated as BERT) includes a pattern generator and an error detector, and may thereby directly measure bit errors at a defined decision point within the eye over an extended period of time. This is helpful in certain systems where a particular combination or stream of symbols may lead to non-linearity in the receiver which may not be immediately perceptible on the eye diagram.

3.2.2 Data encoding

For transmission over a single optical fibre, there are essentially only two data encoding schemes: amplitude modulation using non-return to zero (NRZ) and return to zero (RZ). NRZ is the simpler form of encoding: the data 1 or 0 is latched by the rising edge of the reference clock at each cycle. Therefore a 1 is encoding as a high level for the duration of a single clock cycle, and conversely for a 0. A string of 1's or 0's are encoded as a quasi-DC high or low level until a data transition occurs. RZ encoding is a logical AND operation of the clock signal and NRZ data stream. Data is latched on the rising clock edge, and returns to the 0 level on the falling edge of the clock. A simulated data stream encoded in both formats is shown in Figure 3-6.

Optical transmission systems up to 10 Gb/s most commonly use NRZ encoding. However for higher data rates over long distances, researchers have found significant problems associated with fibre dispersion and improvements may be achieved by using the RZ format. [69]. NRZ has two other important disadvantages. First, power consumption is much higher than RZ since a 1 is transmitted as a light source "on" pulse for the entire clock period rather than only half. Second, a long string of 1's or 0's contains no timing data since it appears to be a DC level. To overcome this problem so-called block codes are used that introduce redundant bits into the bitstream to enable

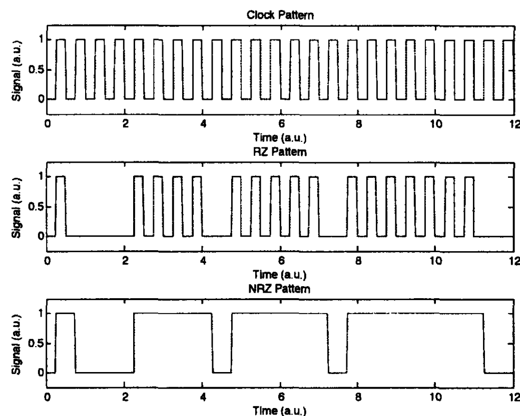
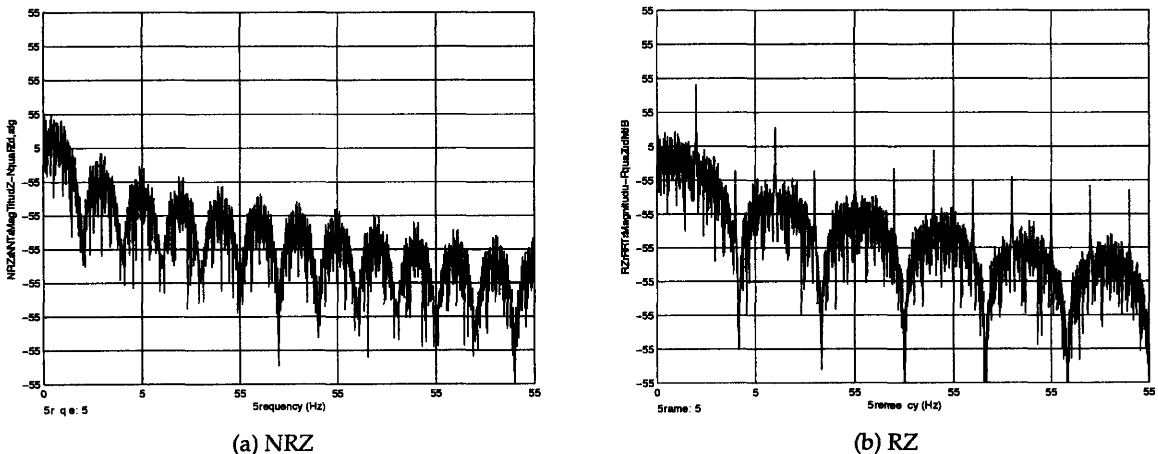


Figure 3-6: 24 bit sequence encoded in RZ and NRZ format. The sequence is 100011110111110111111100.

clock recovery and also provide error monitoring. NRZ encoding with block codes is used for example in various versions of Ethernet and in the Fibre Channel standard.[68, 70] Alternatively, RZ encoding provides a strong timing signal. This may be readily seen by comparing the spectra of NRZ and RZ data in Figure 3-7.

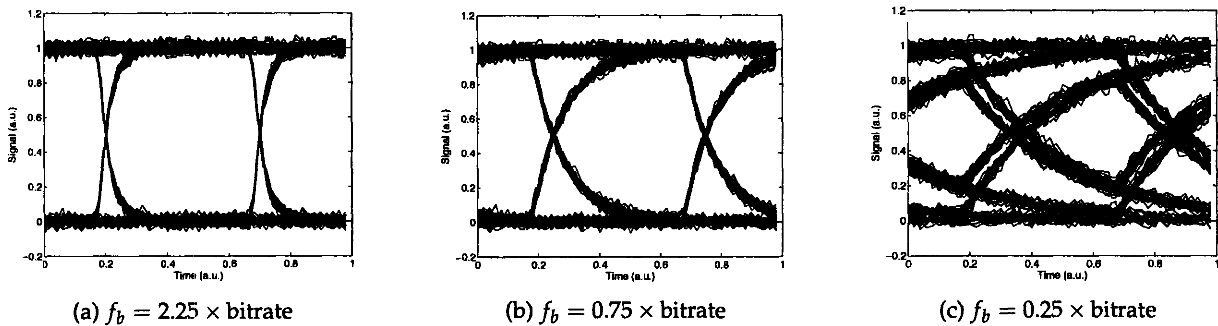


(a) NRZ (b) RZ

Figure 3-7: FFT spectra of a data stream encoded in NRZ and RZ formats. The clock rate is 2 bps in this calculated example.

The NRZ data feature minima at harmonics of the fundamental clock frequency, whereas the RZ data feature maxima at those points. In fact, most of the power of the RZ bitstream is dominated by the component at the fundamental clock frequency, whereas the NRZ bitstream power is mostly in the range below the fundamental. Therefore, it is apparent that the RZ bitstream is de-

sirable from the point of view of clock recovery at the receiver. However, if the perfect bitstreams are subjected to a bandwidth-limited channel (the simplest example being a single-pole low pass filter), the effect of a given filter will be stronger on the RZ data because of the strong frequency component in the range where the NRZ data has a minimum. This effect was simulated in Matlab, and the results of the bitstreams, also with the addition of some noise for a more realistic picture, is shown in Figures 3-8 and 3-9.



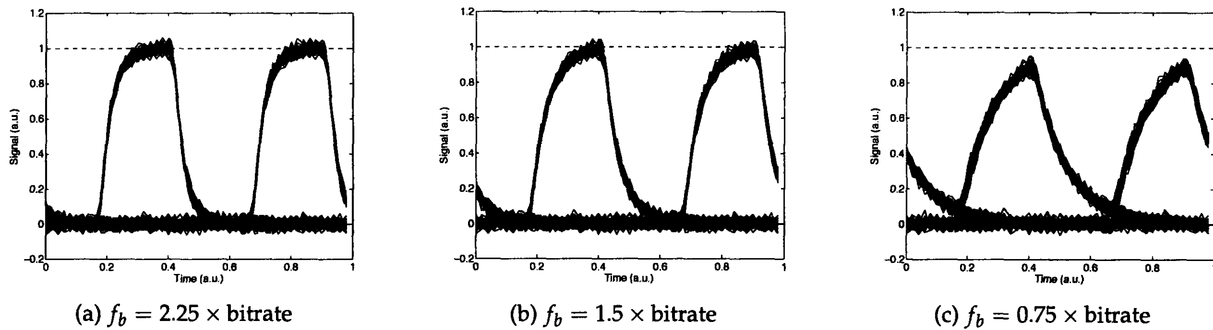
(a) $f_b = 2.25 \times$ bitrate (b) $f_b = 0.75 \times$ bitrate (c) $f_b = 0.25 \times$ bitrate

Figure 3-8: Simulated eye diagrams for a NRZ bitstream under increasingly severe filter corner frequencies.

For filter bandwidths much above the clock rate, the NRZ eye diagram appears perfect with a wide opening. For a bandwidth of 0.75 times the bitrate, the eye is still completely open with the BER essentially unchanged from the perfect case. Below the frequency the eye degrades quickly, as seen in the case of filter frequency of 0.25 times the bitrate whose Q is very low.

The distinction between bandwidth and bitrate in the specification of an optical element is very important as seen in Figure 3-8. The bitrate achievable for a given frequency is 4/3 times the frequency to achieve and undegraded eye diagram. As a rule of thumb, the general conversion of frequency to bitrate is 1.5, and for some communication systems (for example electrical backplanes), a factor 2 is accepted [71, 72]. This consideration is of utmost importance when choosing components for a test system - the frequency response or the eye diagram of the component should be examined in the specification sheet in order to determine which definition is used by the vendor. Confusion in these specifications will certainly lead to calibration errors.

For filter bandwidths much above the clock rate, the RZ eye diagram also appears perfect with a wide opening. However, in the RZ case a bandwidth of 1.5 times the bitrate is required to



(a) $f_b = 2.25 \times$ bitrate (b) $f_b = 1.5 \times$ bitrate (c) $f_b = 0.75 \times$ bitrate

Figure 3-9: Simulated eye diagrams for a RZ bitstream under increasingly severe filter corner frequencies.

prevent eye degradation as shown in Figure 3-9. This is therefore the disadvantage of RZ encoding compared with NRZ encoding, i.e. twice as much bandwidth required for the same bitrate. Cost considerations therefore have limited RZ to areas where it is required (long distances, high bit rates requiring precise clock recovery), though interest has been increasing in recent years as the cost-performance metric of drivers and receivers improves with advances in IC fabrication processes.

3.2.3 Eye Diagram Apparatus

An apparatus to measure eye diagrams was designed and assembled as part of this work. To account for different applications, two sources in wavelength range of interest were used: an 850 nm VCSEL in an LC fibre-coupled package, and a 1550 nm laser with an external $LiNbO_3$ Mach-Zehnder modulator. A schematic of the apparatus is shown in Figure 3-10. The VCSEL is driven directly to predetermined low and high levels (both above lasing threshold, approximately 10 dB extinction ratio), while the modulator package includes a driving circuit and is driven at low and high levels as listed in the component specification sheet.

The connection between the DUT and the sampling oscilloscope is similar to that of the impulse measurement described in section 3.1, and the same comments regarding calibration of the apparatus elements (probe, bias T, cables, connections) apply in the case of the eye diagram measurement. There are two important differences as compared with the impulse measurement. First, the eye diagram trace is not transformed into the frequency domain, therefore calibration is unnecessary. Of course this statement implies that any frequency-dependence of the elements and

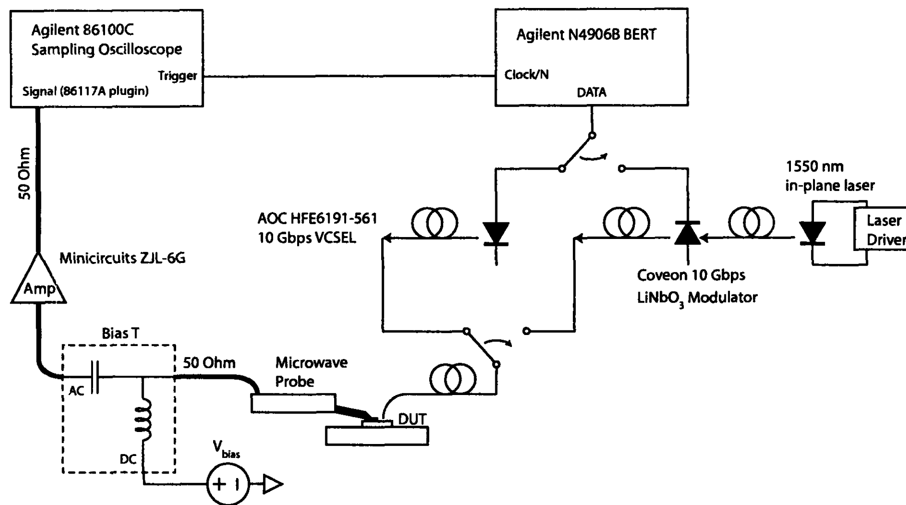


Figure 3-10: Schematic representation of the eye diagram measurement system designed and assembled by the author. The switches represent manual replacement of a unit, cable, or fibre. Certain details have been omitted, in particular the polarisation controller between the 1550 nm laser and the external modulator, and the DC electrical connections to the external modulator.

connections may appear as artefacts in the eye diagram. Second, an amplifier is generally required between the photodiode and the sampling oscilloscope in some cases. For example, the VCSEL source is able to produce approximately 200 μW as a "high" signal. For a device with ideal external quantum efficiency (0.675 A/W), the photodiode would produce 135 μA of current, which in turn produces 6.75 mV at the oscilloscope load. The minimum resolution of the oscilloscope is 1 mV/division, and at that level there is approximately 1-1.5 mV of noise. Therefore an unamplified bitstream would be heavily dominated by instrument noise. In the case of the 1550 nm laser with external modulator, higher intensities are possible. However in such a case one may approach the saturation limit of section 2.3.2. The amplifier used in this application is a Minicircuits ZJL-6G RF amplifier. The specifications are: bandwidth 20-6000 MHz, minimum gain 10 dB, typical gain 13 dB, noise figure 4.5 dB.

3.2.4 Eye Diagram Examples

A near-perfect eye diagram was presented above demonstrating the pattern generator signal. In this section, three imperfect diagrams will be shown.

Figure 3-11 shows the eye diagram in the case of a detector with bandwidth well below the NRZ limit. In addition, the amplifer in this case (Minicircuits ZKL-2R7, 20 dB gain) contributes

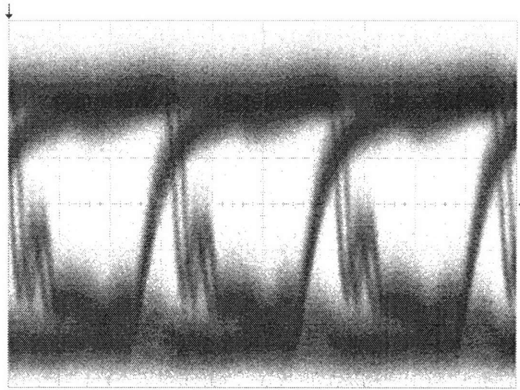


Figure 3-11: Vertical Ge-on-Si detector illuminated by an 850 nm VCSEL driven with a 660 Mb/s NRZ PRBS, the minimum bitrate of the N4906B BERT. The detector bandwidth was independently measured to be below 400 MHz. The eye diagram exhibits a low Q with substantial amplifier noise. Some electrical reflection may be seen on the falling data edge.

substantial noise as seen in the low and high levels. The falling data edge appears to exhibit some electrical reflection from the interconnect. In sum, all of the imperfect factors in this measurement contribute to a poor eye, but it is difficult to extract the main cause (or bandwidth of the slowest element) due to the additional noise and non-idealities of the system.

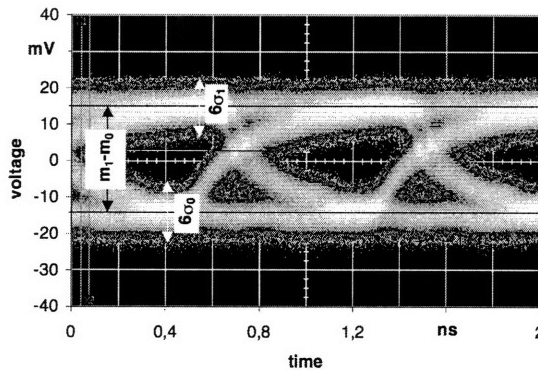


Figure 3-12: Lateral Si detector illuminated by an 850 nm VCSEL driven with a 1.25 Gb/s NRZ PRBS, from [14]

Figure 3-12 shows the eye diagram of a lateral silicon photodetector at 20 V reverse bias. The eye is slightly superior to that of Figure 3-11, but also exhibits a relatively low Q and substantial jitter at the transitions. The authors estimated a Q of 5.7 resulting in a BER of 6×10^{-9} [14]. This is rather high for use in a modern system where a BER of less than 10^{-9} is required, and typically 10^{-12} is typically desired. Additionally, the points of the eye where σ_0 and σ_1 are measured are

indicated in the figure. However, these values were defined at different times rather than at one decision point. Therefore, the actual Q is lower than stated by the authors.

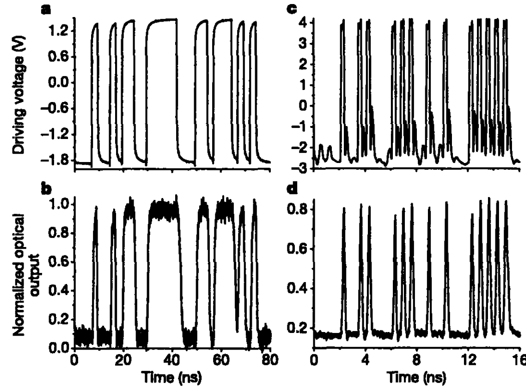


Figure 3-13: Silicon ring modulator driven by 400 Mb/s NRZ and 1.5 Gb/s RZ PRBS signals with light output measured by a high speed photodetector and amplifier, from [73]

Figure 3-13 shows a string of step responses in time rather than an eye diagram. The device in question is a silicon ring modulator. This example illustrates two important aspects of eye diagram or step response measurement. First, if the step response measurement of the modulator were redrawn as an eye diagram, the eye opening would be small in both cases. Particularly in the RZ case there is the question of why the optical pulses reach only 0.8 on the normalised scale, indicating a poor eye opening if the data were to be redrawn. It is difficult to make any judgement of device bandwidth based on this data. Second, regardless of this last consideration it is clear that the NRZ and RZ performances of the device are very different, in other words the device operates in a very non-linear regime. An eye diagram is more likely to expose non-linear behaviour than other measurement methods.

3.2.5 Eye Diagram Summary

The properties of the eye diagram measurement make it more appropriate for certain devices/systems than others. In particular, the eye diagram is most appropriate when measuring an entire system from transmitter to receiver and least appropriate when attempting to measure the linear properties of a single device.

- Pulse generator shape and optical source properties are embedded in the transmitted signal.

- Receiver noise is embedded in the eye diagram, meaning a relatively high signal level is required in order to avoid misinterpreting the RMS width of the low and high levels, as well as the jitter.
- Generator and receiver noise are part of the measurement and may not be filtered.
- Non-linearities in the devices and electrical connections are exposed as the signal extents are larger than the small signal regime. This is precisely both the power of the eye diagram method (the ability to capture all non-linearities in the system) and its weakness (the difficulty of judging the performance of a particular component of the system).

3.3 Frequency Domain Measurement

The frequency domain measurement technique produces the system function of the DUT as its output by measuring the small signal response of the DUT over a wide range of frequencies. On the transmitter side, a light source is biased at a convenient DC operating point and a small sinusoidal variation is added to the DC signal at the input port of the light source. The light, whose intensity is some DC value slightly varying at the frequency of the additive small signal, impinges on the DUT resulting in a photocurrent that consists of a DC value and a small sinusoidal variation. The DC and AC components are separated. The DC component may be collected to determine the DUT DC sensitivity (and the DC arm may also be used to fix the DUT at an operating point, i.e. a reverse bias), while the AC component of the photocurrent is compared to the driving small signal thereby giving the magnitude of the system function at the small signal frequency. The small signal frequency is swept, typically over 2 or 3 orders of magnitude, to produce the system function. This configuration is shown schematically in Figure 3-14.

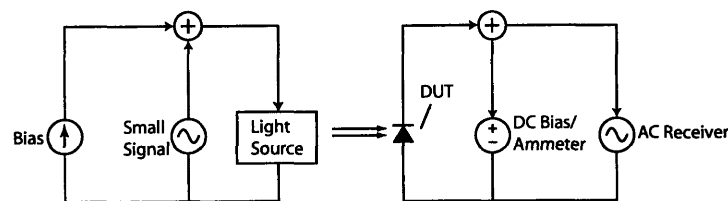


Figure 3-14: Schematic representation of the design of the frequency domain measurement apparatus

With the right choice of measurement equipment, it is possible to calibrate for the majority of parasitic effects in the apparatus setup. Compared to the impulse response technique, the frequency domain method allows for much greater control of the frequency points included in the measured system function, as well for a more precise control over the bias and illumination conditions.

The data in chapter 5 of this work was measured on a frequency-domain system designed and assembled by the author. This system compares extremely favourably to other systems used to measure data in the literature (for example [14, 54]), and is comparable or superior to commercial products such as the discontinued Agilent 86030A and the preliminary (at time of writing) Agilent N4373A.

3.3.1 Frequency Domain Measurement Apparatus

A detailed schematic diagram of the frequency domain measurement apparatus designed and assembled by the author is shown in Figure 3-15. The system may be separated into two halves: the optical sources and the optical receivers. The sources include two wavelength regimes of interest: 850 nm (VCSEL) and 1550 nm (DFB laser with external modulator). Sources in the 780 nm (VCSEL) and 660 nm (RCLED) regime have also been successfully used with the apparatus with no significant modifications other than an appropriate electrical mounting and fibre-coupling scheme depending on the packaging of the source. The receiver side includes the DUT as well as a DC power meter and a wideband reference receiver.

An important and novel aspect of the apparatus is the interface between transmitter and receiver sides: a fibre U-bench with a variable free-space attenuator. This scheme allows for over 30dB adjustment of light intensity transmitted to the receiver side of the apparatus with no change in optical source operating parameters. Typically, intensity is controlled by changing the bias point of the optical source or defocusing the beam incident on the DUT in order to decrease the modal overlap. The former method requires extensive recalibration of the sources and modulation depth at each operating point and may generally yield at best a 20dB adjustment range. The latter is a simpler method but is difficult to control accurately, will generally not yield more than a 20 dB adjustment range, and depending on the device structure may result in carrier generation away

from the DUT that distorts or contributes an undesired component to the measurement result. The variable attenuator is therefore very useful in achieving a wide range of light intensity, and more importantly ensuring the consistency and repeatability of intensity-dependent measurements.

It should also be apparent that with a slight rearrangement of components (in particular, moving the probe to the generator side of the apparatus) and with the addition of micro-positioning equipment, this apparatus may be used to measure the frequency domain response of a laser, LED, or modulator as the DUT.

3.3.2 Measurement Procedure

The basic steps of the measurement procedure are establishing a DC operating point, calibrating the frequency response of the optical source, and obtaining DUT data. These steps are described in detail below, and various aspects of the procedures are discussed in further sections.

1. DC operating point. On the receiver side, the optical multimeter is selected from the three available receivers of Figure 3-15. The desired emitter is chosen, and the variable attenuator is turned to its most transparent position. In the case of the 850 nm source, the current is raised until the predetermined operating point is reached. The DC intensity should be below the point at which thermal effects cause an intensity roll-off. In the case of the 1550 nm source, it is operated at DC at a predetermined intensity and the modulator is adjusted such that it is biased at quadrature. This is done using a pattern generator and sampling oscilloscope and the appropriate offset inputs are generally stable over a longer time period.

2. Source calibration. On the receiver side, the reference detector is selected. The VNA is set to perform an S12 measurement (with appropriate power, sweep range, measurement bandwidth, and trace averaging). The resulting trace is the frequency response of the optical source.

3. Attenuator setting. On the receiver side, the optical multimeter is selected. The attenuator is adjusted until the desired DC power level is reached as measured by the multimeter.

4. Device measurement. On the receiver side, the DUT is selected. The DUT is contacted by the RF probe (optionally with an amplifier between the probe and network analyser) and the

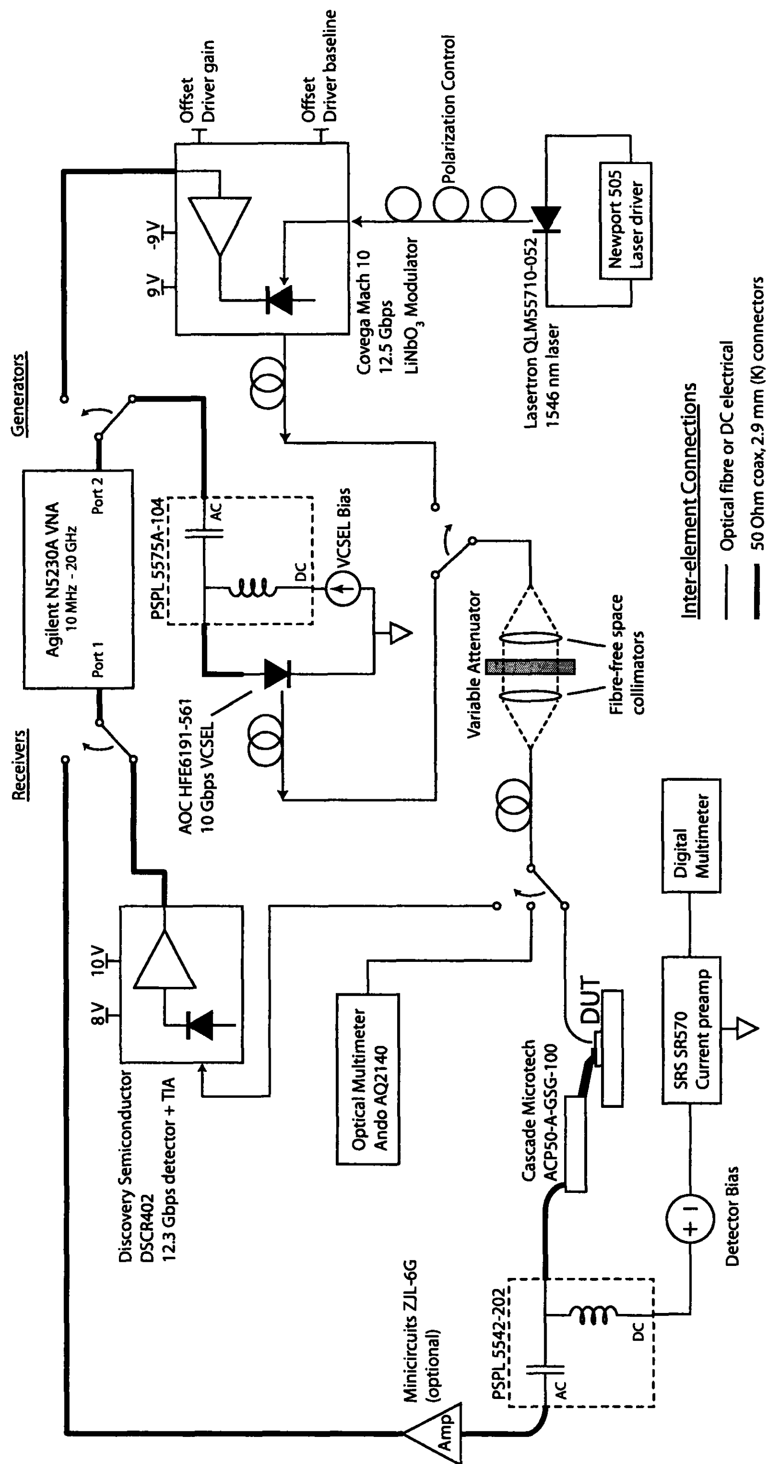


Figure 3-15: Detailed schematic diagram of the frequency domain measurement apparatus. The switches as shown in the diagram are implemented as manual fibre or cable connections, but may also be constructed as semi and fully automated electrical and optical switches/splitters.

fibre probe is positioned over the DUT and adjusted to maximise coupling efficiency. The VNA is set to perform an S12 measurement with the same measurement settings as used in the source calibration sweep, optionally with different trace averaging settings which do not affect the calibration. The resulting trace is the product in the frequency domain of the frequency response of the DUT and the frequency response of the optical source.

5. Steps 3 and 4 may be repeated for different devices attenuator settings and devices. The calibration of step 2 is valid for all resultant traces as long as the electrical excitation settings of the source are not changed.

The magnitude of the DUT frequency response is equal to the measured DUT trace divided by the optical source frequency response. Since the data produced by the VNA is given in dB ($\propto \log_{10} |I|$), the calibration data in dB is subtracted from the measured DUT data in dB to give the magnitude of the DUT system function in dB.

3.3.3 Calibration and Operating Point Considerations

A very important aspect of the frequency domain apparatus is the ability to calibrate the sources and connections in order to compensate for system non-idealities. This is done in the second step of the the measurement procedure, and takes into account not only the frequency-dependance of the source(s), but also the loss in bias T's, adapters, and cables. The latter factor is of particular importance: cable loss consists of connector bandwidth and dielectric loss. The connector bandwidth typically results in a rapid decay magnitude decay or a sharp increase in impedance mismatch at or beyond the specified connector bandwidth. The dielectric loss increases with frequency and will result in an apparent decay of transmission. For the cables used in this work, this loss contributes approximately a -1.5 dB decay from DC to 10 GHz for a 36" cable with 2.9mm (K type) connnectors (bandwidth $\approx$ 40 GHz). This loss may be compensated by calibration with a reference detector. The reference planes for the calibration are shown in Figure 3-16.

The same cable is used to connect the network analyser to the reference photodetector and TIA, and to the output of the DUT bias T (or optionally the amplifier). The calibration captures the frequency behaviour of the network analyser generator, bias T or modulator driving network, the VCSEL or modulator itself, the connections between the network analyser and the source(s),

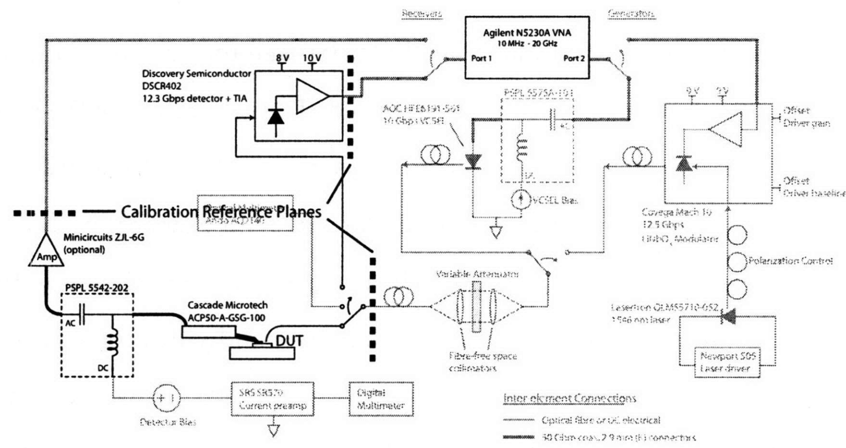
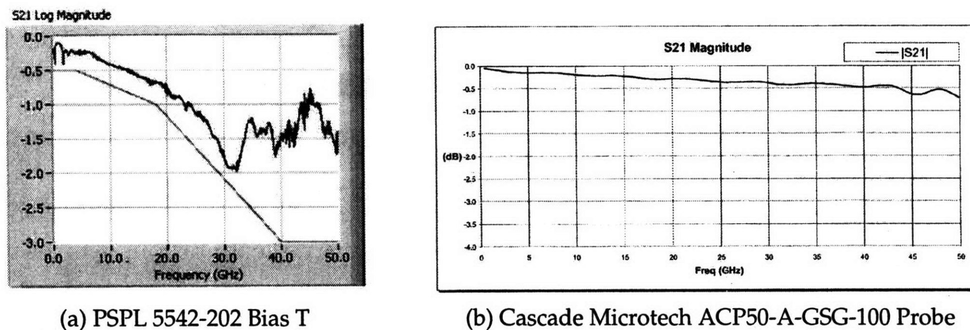


Figure 3-16: Frequency domain measurement apparatus (Figure 3-15) showing calibration reference planes. The elements of the apparatus that are corrected for in the calibration are in faint grey. The cable between the network analyser and either the reference photodetector or the output of the DUT (or amplifier) is physically the same cable.

and the cable between the receiver output and the network analyser receiver port. The reference photodetector and TIA are listed to perform at 12.3 Gbps, and specified with a minimum bandwidth of 9.5 GHz (typical 10 GHz) limited by the TIA, with a typical (minimum is not specified) gain flatness of ± 0.75 dB from DC to 7 GHz. For the purposes of calibration, it is assumed that the total detector and TIA response is simply flat until near the bandwidth. It is worth noting that more precise and calibrated reference detectors are available commercially, though most models are pigtailed with SMF. This is problematic since the VCSEL emits into a MMF, and therefore the TIA in the package is helpful in maintaining a useful signal level after the MMF-SMF coupling ($>$ 20 dB loss).

The calibration reference planes compare the response of the reference photodetector and TIA to the response of the DUT, microwave probe, bias T, and amplifier. The amplifier may be calibrated separately by obtaining its S21 curve using the network analyser and an appropriate connectorised calibration kit. This curve may be subtracted (in dB) from the measured data to apply the correction. Alternatively, it may be possible to connect the amplifier to the output of the reference photodetector during that stage of calibration in order to embed the amplifier response in the reference curve. The applicability of this method depends on the component choice of the

reference photodetector and amplifier; in the case of the apparatus described here, the reference TIA produces a DC value at its output that is likely to damage the amplifier during the connection/power up stage, therefore this method was not used. The calibration of the bias T and probe may be carried out by using a reference substrate to measure the S11 (reflection) and thereby build an equivalent model of resistor, capacitor, and inductor. Such an exact calibration was however not required for the measurements in this work. The components were chosen with 2.4 mm connectors and bandwidths in excess of 50 GHz (to maximise bandwidth for an accurate impulse measurement system), and the response decay in the range DC-10 GHz is not significant as seen in Figure 3-17. The total flatness from DC to 10 GHz is approximately -0.4 - +0.1 dB, which is a small enough range that calibration is not required. The connections between the probe, bias T, and amplifier/cable is implemented by 2.4-2.4mm, 2.4-3.5mm, and 3.5-3.5mm adapters whose properties are dominated by connector bandwidth (of the slower connector) or 26.5 GHz.



(a) PSPL 5542-202 Bias T

(b) Cascade Microtech ACP50-A-GSG-100 Probe

Figure 3-17: Manufacturer calibration/verification data for the bias T and microwave probe used in the measurement apparatus

VCSEL Operating Point, Coupling, and Packaging

Proper VCSEL operation in the test environment is affected by several factors that can contribute to misleading or inconsistent test results. For a well-calibrated and repeatable measurement system these factors must be understood and the negative effects minimised.

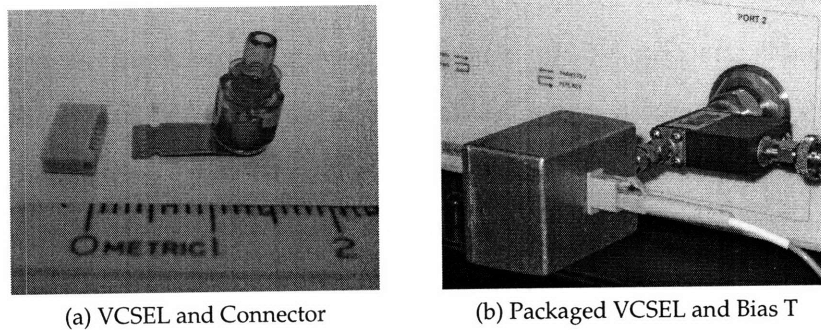
VCSEL bandwidth is determined by two factors: parasitic device/circuit elements (resistance and capacitance), and the photon relaxation resonance in the laser cavity. For high-speed VCSELs, the latter is the dominant bandwidth limit. The device is generally empirically modelled as an LRC

circuit decoupled from parasitic elements based on the measurement of resonant frequency and response overshoot height. The VCSEL bandwidth may be approximately related to the output power $f_{3dB} \propto \sqrt{P_0}$ [74]. Therefore, the relaxation oscillation limit of VCSEL bandwidth increases with increasing DC operating power. In practice, there is increased damping at higher powers that establishes an intrinsic modulation bandwidth capability of the laser that depends on size, carrier lifetime, and gain in the laser cavity. This consideration is the primary reason behind the use of the variable attenuator in the measurement apparatus: the power incident on the DUT is decoupled from the operating point, such that the laser may be biased in a regime where the relaxation oscillation bandwidth is optimised.

Fibre-coupling of the VCSEL light may also affect measured bandwidth. Generally VCSELs for communications applications emit in multiple transverse modes. Poor VCSEL coupling arising from improper geometry (i.e. fibre offset) can degrade optical link performance due to selective mode coupling. This always results in reduced coupling efficiency, and depending on the precise geometry may result in an oddly-shaped frequency response as different modes (with different coupling efficiencies) are active in different parts of the frequency sweep. In principle this effect may be tolerable since the laser calibration curve may account for such effects. However, in practice this may affect the dynamic range or noisiness of the measurement due to lower signal level in parts of the trace, and may generally lead to unexpected behaviours due to mode hopping around connectors or bends in other areas of the apparatus.

To ensure consistent and optical coupling, a 10 Gbps VCSEL was chosen that is in a moulded packaged with integrated LC connector coupling. The VCSEL was then further packaged in a connectorised metal box to ensure mechanical stability and to reduce the problem of signal leakage discussed below. A small circuit board was designed to serve as the mechanical support, and included a 50 Ω microstrip line to an SMA connector. A connector was soldered to the board, and the VCSEL was inserted into the connector. This is shown in Figure 3-18.

The connector was used due to the worry of damaging the devices which were not easily obtained. However, the connector resulted in an impedance mismatch to the strip line of the PCB, creating resonant peaks in the laser response where a reflection from the mismatch resulted in destructive interference with the driving signal, as seen in Figure 3-19. It is generally possible to compensate for this effect with source calibration, though this technique is not always effective in



(a) VCSEL and Connector (b) Packaged VCSEL and Bias T

Figure 3-18: Photograph of the AOC 10 Gbps VCSEL and connector, as well as of the completed source package connected to a bias T and network analyser.

low-photocurrent regimes.

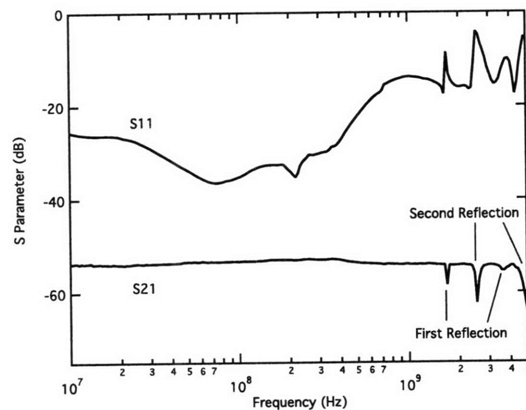


Figure 3-19: Measured S11 (reflection) and S12 (transmission) of the packaged VCSEL. There are two reflections clearly visible in both traces at multiples of approximately 1.8 GHz and 2.5 GHz. These interference peaks correspond to a reflection at the cable-PCB junction (smaller mismatch/reflection) and the PCB-VCSEL connector/jack junction (larger mismatch/reflection). The S21 trace is very low due to the coupling loss between VCSEL-coupled MMF and reference detector-coupled SMF which results in approximately 40-50 dB of photocurrent loss).

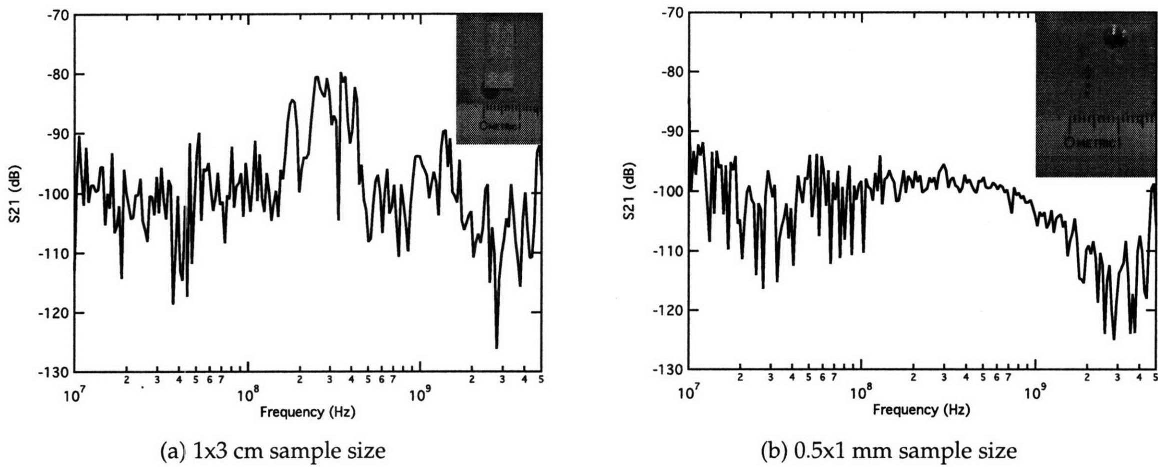
Interference, Leakage

At high measurement frequencies of greater than 1 GHz, the wavelength in free space falls on the order of millimetres, and even wire, metal trace, or substrate may begin to act as an antenna emitting or receiving radiation. This effect may have a disastrous effect on measurements as the receiver (in this case the network analyser) will detect and display a real signal with significant

magnitude, but the signal may or may not be a result of photodetection.

Interference may be divided into two categories: external and internal, also known as leakage. External interference is due to other devices in proximity to the measurement apparatus such as cellular phone (1.9 GHz), wireless ethernet or bluetooth (2.4 GHz), and wireless phones (900 MHz). This effect may be recognised as it is intermittent during a measurement session and independent of the source configuration. For example, a change in the source excitation power will result in a corresponding change in received signal power, but not a change in the measured quantity (e.g. S11, S21) since the instrument always displays the ratio of received to transmitted power. A signal caused by external interference would therefore appear to be gaining or losing strength as the source power is changed. If external interference is affecting the measurement, the only solution is to eliminate the interference source if possible, or to isolate the receiver (i.e. in this case the DUT) by placing it in a grounded metal enclosure.

Leakage is real signal emanating from the generator and sensed by the receiver, but that travels by an undesired path (i.e. not generator-optical source-fibre-attenuator-DUT-receiver). The undesired path may be inside the network analyser itself, which may be verified as part of the instrument calibration process (for a modern instrument, an isolation of at least 100 dB is usually achieved). The path may also be outside of the analyser, that is by radiation at the source. This effect is very troublesome when operating at high attentuation and/or with low-sensitivity detectors. In these cases the leakage signal may be 30-40 dB higher than the DUT signal. There are several possible solutions to the leakage problem. Similar to external interference, the receiver side of the measurement apparatus may be isolated. DC connections to the amplifier and bias T should be isolated (i.e. using coaxial cable with short connections to the components), as the amplifier power supply coupling and bias T DC leakage may result in a significant amount of the signal. On the generator side, the mounting or packaging of the optical source may be improved, as shown in Figure 3-18. Finally, sample size plays a significant role in capturing the leakage signal. A smaller sample is a less effective antenna, and this effect alone may reduce leakage by 10-20 dB as shown in Figure 3-20.



(a) 1x3 cm sample size

(b) 0.5x1 mm sample size

Figure 3-20: Leakage signal for a large and small sample size, as seen in the inset photos. The smaller sample reduces leakage by up to 20 dB in the 150-400 MHz range, and by 10-15 dB in the 1-3 GHz range. In the measurement setup, the DUT is probed while the VCSEL is modulated and coupled to the optical multimeter.

Electrical reflections

Imperfect cable and connectors may result in some slight impedance mismatches, causing a small frequency response ripple. In the case of perfect matching at only one end of the transmission line, no ripple generally occurs. In a case of very large mismatch at one end of the transmission line and perfect matching at the other end, sharp resonances occur as previously seen in Figure 3-19. For slight mismatches at both ends of the line, a full ripple may be observed and is dependent on the length of the transmission line as shown schematically in Figure 3-21.

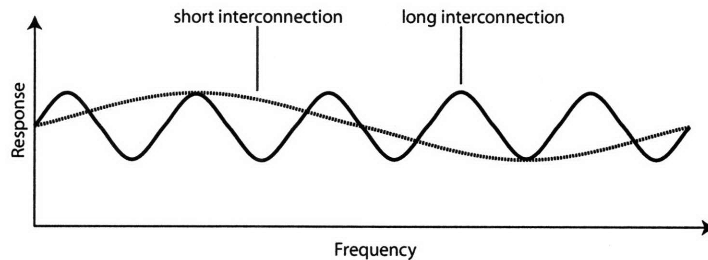


Figure 3-21: Frequency response due to standing waves in the electrical transmission line.

Strategies for minimising transmission response ripple are shortening the transmission line, improving line termination, or buffering the signal. The former method is used on the generator

side with the VCSEL, as seen in Figure 3-18. The bias T is connected directly to the network analyser port, and the VCSEL is attached to the bias T by means of a 3.5-3.5mm adapter. Additionally, the network analyser port and the bias T transmission line are well-matched, resulting in negligible reflections. In the case of the external modulator, the network analyser provides a signal to the (terminated) modulator driver which in turn buffers the signals from the modulator itself.

The situation is somewhat more complicated on the receiver side. First, the DUT is effectively a capacitor, or an open circuit until sufficiently high frequencies are reached such that its impedance begins to fall. Therefore there is significant reflection of signals propagating to the DUT. Secondly, it is not possible to connect the DUT directly to the network analyser because the probing mechanism is physically situated some distance away. In the case of the apparatus described in this work, a 36 inch cable connected the DUT to the network analyser, a length sufficient for the appearance of a large number of ripples. The problem is not seen with the reference photodetector as both the TIA output and network analyser are terminated to 50 Ω, therefore the mismatch at both ends of the line is small for an appropriate cable. The reflection at the DUT may be solved in two ways. First, a 50 Ω termination resistor may be fabricated beside the photodiode, and the voltage signal across the termination may be detected rather than the current signal directly entering the 50 Ohm cable. This technique is used for example in [55] and in commercial photodiodes sold by Discovery Semiconductors. The disadvantages are a 3 dB loss of signal as the photodiode is split between the termination resistor and the transmission line, and a more complicated fabrication procedure particularly in the case of devices where the substrate is not isolated from the substrate (i.e. where the substrate is one of the device contacts). The second technique is to insert a termination into the transmission line, preferably as close as possible to the DUT. This may be done using an amplifier. In the apparatus described here the probe, bias T, and amplifier are connected by adapters forming a very short transmission line. The terminated output of the amplifier then drives the long transmission line to the network analyser. The collected data may also be post-processed to smooth away the ripples.

Amplifier

RF amplifiers offering high gain over a wide bandwidth are almost exclusive constructed in a feedback architecture; for example, the Minicircuits ZJL and ZKL series used with this apparatus. The Minicircuits models in particular illustrate a possible behaviour that may lead to large errors, generally resulting in an overestimation of bandwidth. The ZJL and ZKL series of amplifiers use a Darlington pair gain block based on InGaP HBT technology that has become very widespread in RF amplifiers in recent years. A schematic diagram is shown in Figure 3-22.

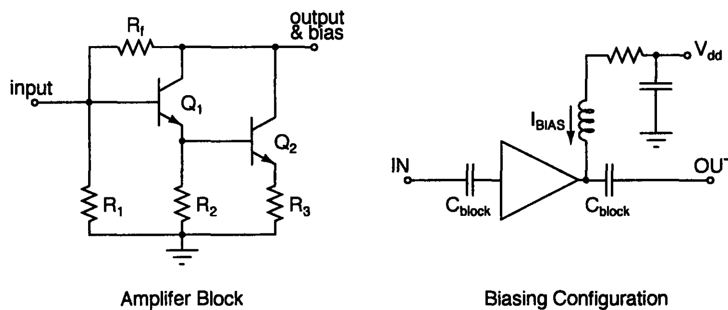


Figure 3-22: Schematic diagram of the Darlington gain block and the bias configuration of the class A amplifier including blocking capacitors that determine the low-frequency bandwidth limit of the amplifier.

The ZJL-6G (minimum 10 dB gain, bandwidth 20 MHz-6 GHz) uses a single gain block, while the ZKL-2R7 (minimum 24 dB gain, bandwidth 10 MHz-2.7 GHz) uses two cascaded gain blocks. It is also apparent from the biasing diagram why the power supply connection to the amplifier is an efficient interference and leakage coupling mechanism. Resistors R_1-R_3 serve to bias transistors Q_1 and Q_2. Resistor R_f is the feedback resistor that determines the gain of the block. The amplifiers are designed to operate in a 50 Ω terminated system, with an appropriate phase margin. However, by placing a capacitor (DUT) at the input, the feedback loop forms a low-pass filter of the feedback resistor and DUT capacitance, as seen in section 2.4.1. This reduces the phase margin of the amplifier, leading to ringing in the time domain or gain peaking the frequency domain. The latter may lead to an incorrect evaluation of DUT bandwidth. For example, a device with 9 pF capacitance (RC time constant approximately 350 MHz) was measured to have a bandwidth of 300 MHz with no amplifier and 650 MHz with the ZKL-2R7 amplifier. Similarly devices with RC time constant of 2.5-3 GHz were measured to have a bandwidth of over 5 GHz with a 2-3 dB peak around 3.5-4 GHz. The solution to this problem is either an amplifier with larger phase margin, or

an amplifier with a lower gain which extends the parasitic filter bandwidth to higher frequencies.

3.3.4 Frequency Domain Measurement Examples

This section will show examples of frequency domain measurements conducted by the author and from the literature. Figure 3-23 shows one of the first frequency domain traces measured with the apparatus, demonstrating many of the sources of error discussed in the previous section.

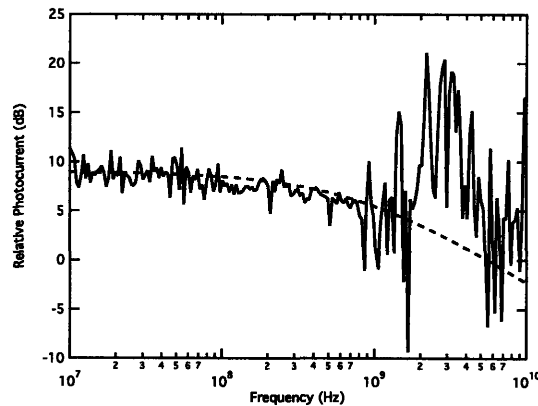


Figure 3-23: Frequency domain trace of a Ge-on-Si vertical photodiode. The trace exhibits significant noise and interference/leakage.

The data was captured using a poorly packaged VCSEL (with no isolation), poorly connected power supply pins (several cm long supply and ground wires), a too large measurement bandwidth setting, and large substrate. The trace at low frequencies is dominated by noise with peaks of +/- 2 dB coming from noisy power supply connections and inappropriate trace averaging and measurement bandwidth. Above 800 MHz, the trace is dominated by leakage of the RF signal from VCSEL pins to DUT substrate (and subsequent coupling to the RF probe). The device fit shown by the dashed line is poor and not valid beyond several hundred MHz.

The same device was measured after improvements to the measurement apparatus, and the very different traces are shown in Figure 3-24. The data was captured following improved packaging of the VCSEL as discussed above, isolation of the power supply lines to the DUT bias T and amplifier, and reduction of sample size. The trace ripple is reduced to +/- 0.5 dB. Additionally, the two DUTs may be accurately fitted to decay functions from 10 MHz to 8 GHz. The upper frequency is the measurement limit of the apparatus, set primarily by the bandwidth of the VCSEL which is specified to 10 Gbps operation, and also to some extent by the reference photodetector

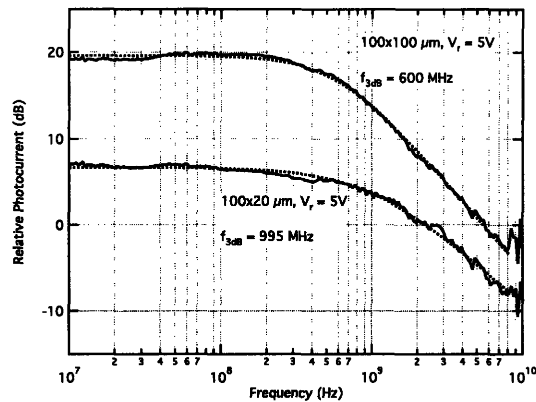


Figure 3-24: Frequency domain trace of the same Ge-on-Si vertical photodiode as in Figure 3-23 (bottom trace). Noise and leakage have been eliminated. The traces follow curve fits to approximately the 8 GHz, the measurement limit of the apparatus with the 850 nm VCSEL source.

TIA whose gain ripple is specified to 7 GHz. Compared to the majority of frequency domain curves presented in the literature, Figure 3-24 shows remarkably clean traces. As an extreme example, one may consider the data of Figure 3-25 taken from [75].

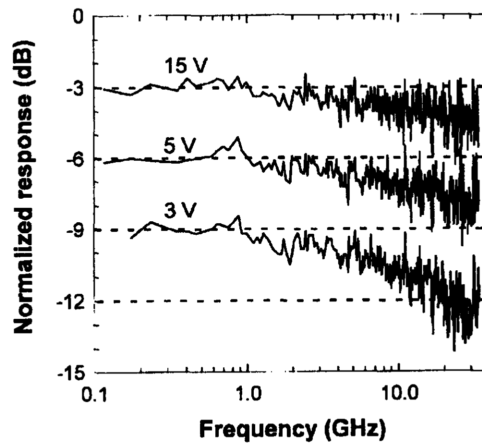


Figure 3-25: Frequency domain trace of a SiGe resonant cavity photodetector taken from [75]

The data of Figure 3-25 may be very easy to misinterpret. The frequency roll-off of the three traces is very slow, and does not correspond to any of the bandwidth-limiting mechanisms discussion in section 2.3. Furthermore, the three traces of different biases are very similar in their behaviour. This would indicate a p-i-n type structure where the capacitance does not significantly change with applied bias. In that case, the bandwidth limit may not be RC or diffusion time, but

is more likely drift transit time. This sort of decay would be significantly more abrupt than the shown traces. Also, the noisy behaviour beyond 1 GHz is identical for all traces, and there is a gain peak around 800 MHz in all of the traces. This suggests that in fact the trace decay is not a strong function of the device, but is in fact cased by interference or (more likely) leakage. The derived bandwidths of the measurement are therefore likely not the real device bandwidths. It is difficult to determine the reason for which the low-frequency parts of the traces are flat and the high frequency parts are very noisy. The trace is apparently recorded with linear frequency spacing, meaning that there are very few data points from 100 MHz to 1 GHz. It may then be a simple coincidence that the few data points are flat.

3.3.5 Frequency Domain Measurement Summary

The properties of frequency domain measurement make it most appropriate for deriving a linear model of a particular device. It is not appropriate for the measurement of an entire transmission system.

- Optical source properties may be measured by a reference detector and calibrated out of measured traces. The attenuation properties of the electrical connections may be corrected by the same procedure.
- Instrument noise may be eliminated by the use of narrow measurement point bandwidths and trace averaging. These two steps can yield a very clean and reproducible signal.
- The VNA as the central instrument offers the large dynamic range and lowest noise of the three measurement techniques.
- A free-space attenuator allows for the decoupling of light intensity from DC device operating point, and allows for a very wide adjustment range.
- The measurement technique is a true small-signal measurement, and therefore is only moderately useful for determining non-linear behaviour (by repeated traces at different DC intensities for example).

Chapter 4

Lateral Photodetector Design and Fabrication

4.1 Device Design

A large device is desired for coupling with polymer optical fibre in order to maximise modal overlap, thereby increasing the effective external device responsivity. Current commercial applications of polymer fibre make use of readily available detectors that are significantly smaller than a typical POF core. The largest commercially-available devices are GaAs MSM detectors designed for glass multimode fibre, with diameters approaching 100 μm as discussed in section 5.1.5. The advantage of interdigitated detectors as compared with vertical junction detectors is a much lower capacitance per unit device area. This may be seen with reference to Figure 4-1. In the vertical device, the junction area is the device area. In the lateral device, the junction area is a proportion of the total area, therefore to a first order approximation the capacitance for a given area device is lower by a factor of $0.5w/(w+s)$ where w is the finger width and s is the inter-finger spacing.

The disadvantage of an interdigitated structure is reduced responsivity due to metal contact shadowing. The ohmic contacts in a vertical structure are designed to minimise their area, therefore due to the metal fill factor alone (other effects ignored), the lateral structure responsivity is lower by a factor of $s/(s+w)$. However, assuming that the difference in bandwidth between the two types of structures is determined by RC time constant and the difference in sensitivity is de-

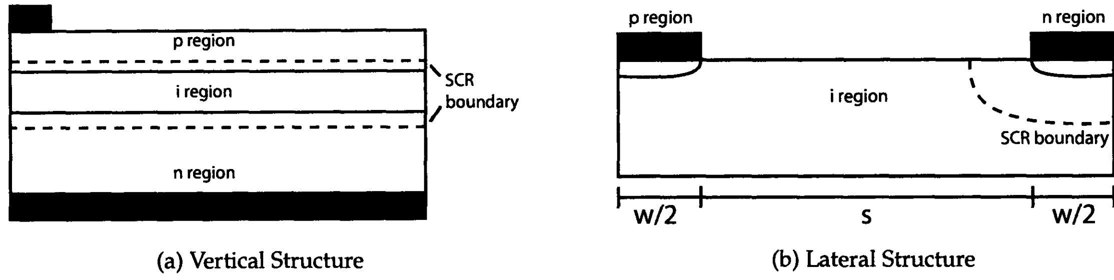


Figure 4-1: Schematic representation of a simple vertical and lateral detector structures. The dimensions indicated on the lateral structure unit cell are inter-finger spacing s and finger width w.

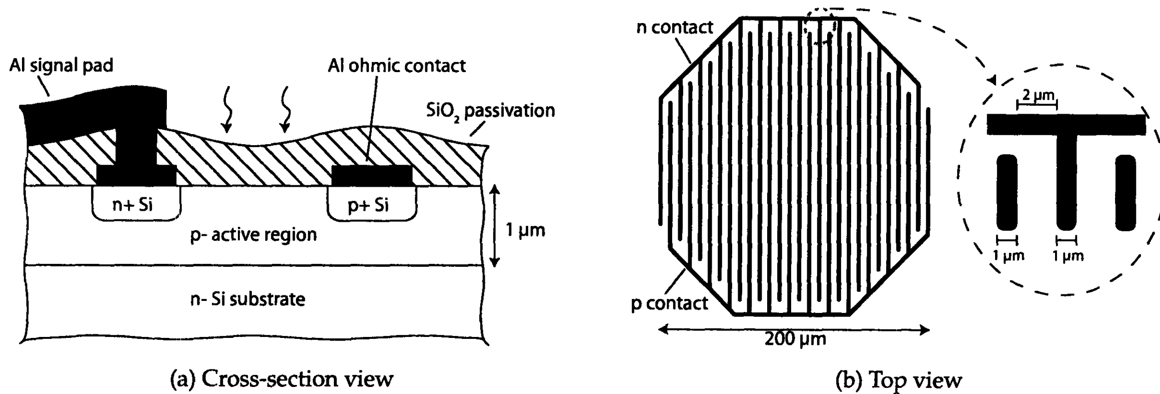
termined by metal shadowing one may write the difference in the bandwidth-sensitivity product figure of merit.

$$\text{BW} \times \text{sense} = \frac{s}{s+w} \times \frac{w+s}{0.5w} = 2\frac{s}{w} \tag{4.1}$$

This expression is greater than one for $s > 0.5w$, therefore there should exist a wide device design space where bandwidth-sensitivity product is superior to a vertical device, and where the particular bandwidth and sensitivity may be optimised for a particular application.

The device structure investigated in this work is shown schematically in Figure 4-2. Variations on this theme may include a germanium active region, or a completely different semiconductor device material, for example GaAs. The important difference between this structure and an MSM device is the lack of Schottky junction, i.e. all junctions are implanted and have ohmic contacts. This allows for the use of a standard CMOS process to fabricate this device. It is apparent that the active region could be formed by a well implant in a double or triple-well process, with the n+ and p+ fingers formed by source/drain implant regions. A device fabricated in a standard commercial VLSI process (in the case of this work, a GaAs process) is discussed in section 5.1. The planar device structure also results in a device that is easier to integrate into a processing flow, as opposed to a vertical device which may present problems of topology that limit lithographic resolution and yield.

A further important aspect of this device is the junction between the active region and substrate. The purpose of this junction is to block carriers generated in the substrate from being collected by the external contacts. This is particularly important in silicon devices, where absorption is weak at wavelengths of interest (780, 850 nm), therefore carriers generated in the substrate



(a) Cross-section view

(b) Top view

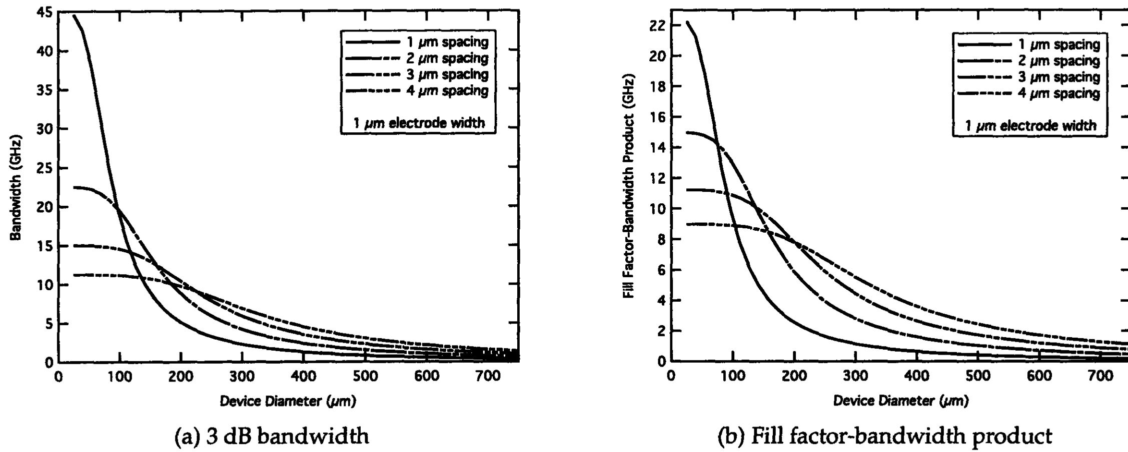
Figure 4-2: Side and top schematic views of the silicon lateral p-i-n device structure, including some typical or expected width, spacings, and thickness measurements.

must travel a long distance to be collected. The bottom junction is a design decision that trades sensitivity for bandwidth. Some sort of barrier is required for high-speed operation of silicon photodiodes, for example a buried oxide as presented in other work [21].

4.1.1 Maximum possible bandwidth

Maximum device bandwidth is achieved when the entire semiconductor material between the n and p fingers is depleted, and therefore carriers move only by drift through the SCR. Furthermore it should be assumed that the electric field is sufficiently strong that carriers move at their saturation velocity. In this case, the capacitance may be calculated by a conformal mapping method [76] and an RC time constant calculated assuming a 50 Ω load. The transit time bandwidth may be estimated by equation 2.18 assuming a saturation velocity of 10^7 cm/s [77] . As both of these effects have a monotonic step response, the total time constant may be estimated as the square root of the sum of squares $\tau_{total}^2 = \sum_i \tau_i^2$ [78]. This result as well as the metal electrode fill factor-bandwidth product are shown in Figure 4-3.

For small diameters and larger electrode spacing, the bandwidth is determined by the transit time effect as seen from the flat bandwidth vs. diameter characteristic. For larger diameters, the RC time constant is dominant. Different electrode spacings are optimal for different device diameters. This is particularly apparent with reference to sensitivity-bandwidth product since the high bandwidths of tightly-spaced electrode devices is achieved at the cost of reduced photoresponse. The calculation suggests that for the target device size in this work (200 μm diameter), a



(a) 3 dB bandwidth

(b) Fill factor-bandwidth product

Figure 4-3: Estimated bandwidth and metal fill factor-bandwidth product for a silicon lateral p-i-n device assuming complete depletion of the semiconductor between the detector fingers, assuming 1 μm wide fingers.

bandwidth above 10 GHz may be achieved in silicon. In this extreme case, the bandwidth has a weak dependence on electrode width. Interestingly, the maximum predicted bandwidth of a bulk germanium lateral p-i-n device is lower than that of a silicon device due to the higher dielectric constant and lower saturation velocity of germanium (6×10^6 cm/s) [79] as seen in Figure 4-4. The bandwidth of a 200 μm diameter device in this extreme limit is 7.4 GHz (achieved for a 3 μm electrode spacing). A device made of epitaxial germanium on silicon would be dominated by the slower saturation velocity, and an effective dielectric constant between that of germanium and silicon, depending on the precise germanium thickness.

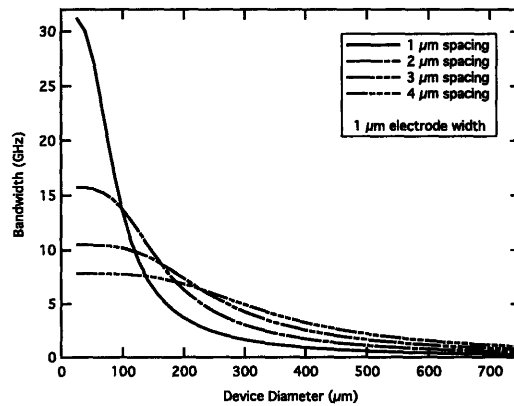


Figure 4-4: Estimated bandwidth a bulk germanium lateral p-i-n device assuming complete depletion of the semiconductor between the detector fingers, assuming 1 μm wide fingers.

4.1.2 Realistic first-order bandwidth estimate

The previous section assumed a fully depleted semiconductor region with carriers travelling solely by drift. For an integrated detector, a more realistic picture would assume some doping of the active region, and therefore some QNR remaining between the contact fingers for typical finger spacings and bias voltages with current flowing in the QNR by diffusion. In fact, it is proposed here that it may be desirable to design a device allowing for some carrier diffusion (which is generally the opposite of what is recommended, e.g. [45, 80]) in order to minimise metal shadowing for a given bandwidth specification. Such a case would appear as in Figure 4-5, with a depletion region of width d and a QNR of width $s - d$ where s is the inter-finger spacing.

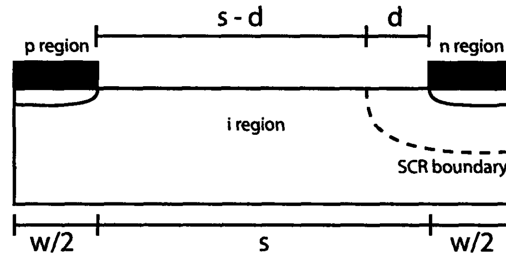
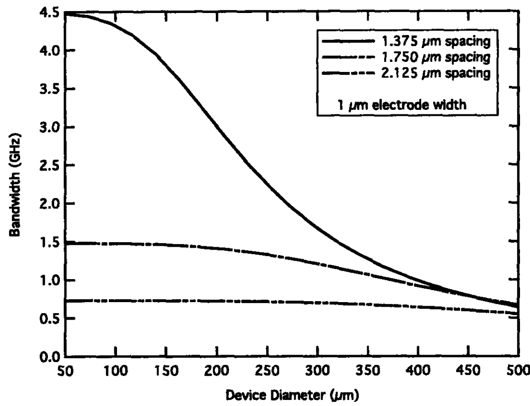


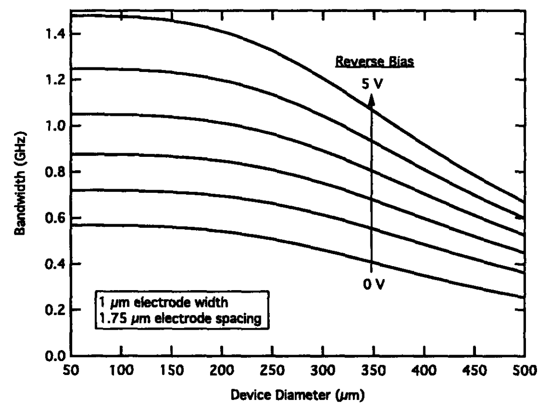
Figure 4-5: Schematic diagram of the device structure used to estimate the lateral p-i-n device bandwidth based on carrier diffusion and RC effect.

Similarly as in the previous discussion, bandwidth is separated into two effects: RC and diffusion. Both effects have a monotonic step response and therefore a total time constant may be easily estimated. The diffusion time constant is calculated according to the discussion of section 2.3.3, assuming a mobility corresponding to a doping of 10^{16} cm^{-3} [81]. The distance d is assumed to be the depletion region width, calculated according to equation 2.10 assuming an asymmetrically doped junction (i.e. the active region has lower doping). The RC time constant is calculated assuming a 50 Ω load and capacitance area corresponding only to the n finger region of the device. This is an underestimation of capacitance as the actual area is higher due to the implant sidewall and lateral expansion of the depletion region under bias. Under these assumptions, the bandwidth as a function of finger spacing, active region doping, and reverse bias may be calculated, as shown in Figure 4-6.

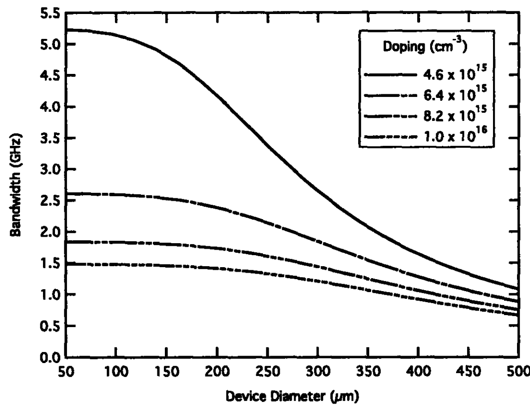
The calculated results indicate two distinct bandwidth-limited mechanisms as shown in Figure



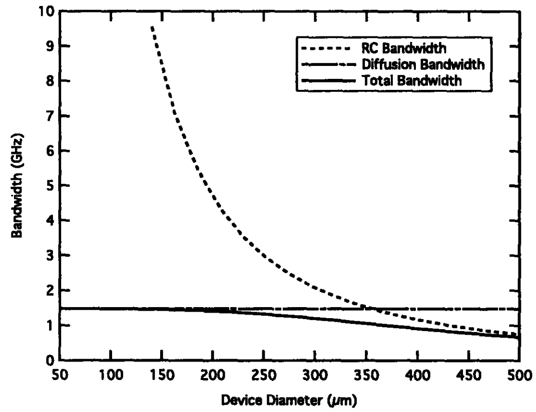
(a) Constant doping (10^{16} cm^{-3}) and reverse bias (5 V), varied inter-finger spacing



(b) Constant doping (10^{16} cm^{-3}) and inter-finger spacing (1.75 μm), varied reverse bias



(c) Constant inter-finger spacing (1.75 μm) and reverse bias (5 V), varied doping



(d) Baseline device, showing both bandwidth components

Figure 4-6: Estimated bandwidth of silicon lateral p-i-n devices with only partial active region depletion and carrier flow by diffusion in the QNR. All subfigures include a baseline curve corresponding to 10^{16} cm^{-3} doping, 1.75 μm inter-finger spacing, 1 μm finger width, and 5 V reverse bias.

4-6d. The first is RC effect, as seen in the previous section, with bandwidth scaling as $1/C$ and therefore as $1/r^2$ where r is the device radius. Furthermore the total device capacitance for a given area rises as inter-finger spacing is decreased, as seen in the case of a large diameter in Figure 4-6a and in the previous section. The RC bandwidth varies proportionally to $\sqrt{N_A}$ in the active region and is approximately proportional to $\sqrt{V_r}$. The second mechanism is diffusion time, where the relevant time constant is independent of device size and varies solely with inter-finger spacing. The diffusion bandwidth is proportional to $1/(s-d)^2$, where d has the same dependance on doping and reverse bias as the RC time constant, and is therefore more sensitive to doping and

bias than RC bandwidth. This may be seen graphically in subfigures b and c where, as conditions change towards higher bandwidth, the diffusion-dominated (flat) part of the curves for small diameters becomes shorter and RC bandwidth dominates over a wider diameter range.

For an integrated device, the doping level may be fixed by the requirements of other devices in the process, for example MOSFET body doping requirements to achieve a desired threshold voltage. Additionally, even if additional mask steps are accepted, in an implanted process the minimum doping must be higher than the substrate doping. Therefore, though this dependence is significant, the photodiode device designer may have little control over this value. Likewise the bias voltage is typically set to the high power rail and an additional diode bias supply is generally not economically desired. There have however been efforts to generate an additional bias on-chip [29]. The main design tool is the electrode spacing. Figure 4-6a indicates that for a 200 μm device diameter, several bandwidths are achievable depending on spacing, but with different metal fill factors. The designer should therefore optimise the electrode pattern to achieve a desired bandwidth specification but with the minimum fill factor. The initial design for this work concentrated on a 200 μm device with a desired bandwidth of 1 GHz. A 2 μm finger spacing was chosen as the nominal design, with 1-4 μm spacings also included to confirm bandwidth behaviour. A doping of approximately 10^{16} cm^{-3} was chosen for reliable manufacturing on readily-available substrates. From the figure, this set of parameters appears to be the interface between diffusion and RC-dominated behaviour for 5 V reverse bias.

A germanium device is expected to behave similarly, with bandwidths scaled appropriately by material parameters. Diffusion bandwidths should increase by a factor of approximately 2.5 for p-type active regions due to the much higher electron mobility. RC bandwidths should decrease by a factor of approximately 0.7 due to the higher permittivity. In general, this signifies higher bandwidths for small to moderate sized devices ($\lesssim$ 300 μm diameter), but lower bandwidths for very large devices. For most useful wavelengths near IR or visible wavelengths, the germanium devices are expected to have significantly higher sensitivity than silicon devices (up for a factor of 15-20 for 850 nm light for example) due to the much higher absorption coefficient.

The above calculations and design considerations are only approximately valid. In reality, the diffusion region edge interacts with the substrate-active region junction resulting in electric field line bending and expansion of the depletion region laterally rather than vertically or isotropically.

This has an effect on the actual device capacitance and QNR width. However, this is a difficult two-dimensional problem and must be solved numerically, as discussed in section 6.

4.2 Silicon Photodetector Fabrication Procedure

A total of five fabrication lots were carried out for the silicon lateral photodetectors. The first and second lots revealed several flaws in the process that were corrected for the subsequent lots. The most important corrections were: thickness of the thermal (passivation) oxide, choice of wet/dry process to etch the implant finger patterns, choice of contact metal thickness/composition, pad-to-ohmic contact via etching (wet/dry), and ion implant conditions (dose and energy). The process used for lot 3 and subsequently is presented in section 4.2.1, and particular decisions and/or processing techniques are discussed in section 4.2.2.

4.2.1 Detailed Description of Process

A schematic view of the fabrication process is shown in Figure 4-7. The description that follows refers to the step numbers in the figure. A precise description of the entire process flow including relevant laboratory instrument names/designations is presented in Appendix B.

1. Begin with prime silicon wafer, either p-type (resistivity $1 - 2\ \Omega\cdot$cm) or n-type (resistivity $2 - 4\ \Omega\cdot$cm) with a boron implant to form the active region in the case of the n-type substrates. The junction depth is approximately 1 μm after all subsequent annealing steps.

2. RCA clean, wet thermal oxidation of silicon surface (950 °C, 30 minutes), resulting in 250 nm of SiO_2 serving as surface passivation, and also resulting in boron implant anneal and dopant diffusion.

3. Mask step 1 - photolithographic pattern and oxide etch to define stepper alignment marks. The oxide etch may be a wet etch (2.5 minutes BOE - 7:1 NH_4F:HF) or a dry etch (CH_3F, CF_4 plasma).

4. Mask step 2 - photolithographic pattern and dry oxide etch (CH_3F, CF_4 plasma) to define finger openings for subsequent boron and phosphorus implants.

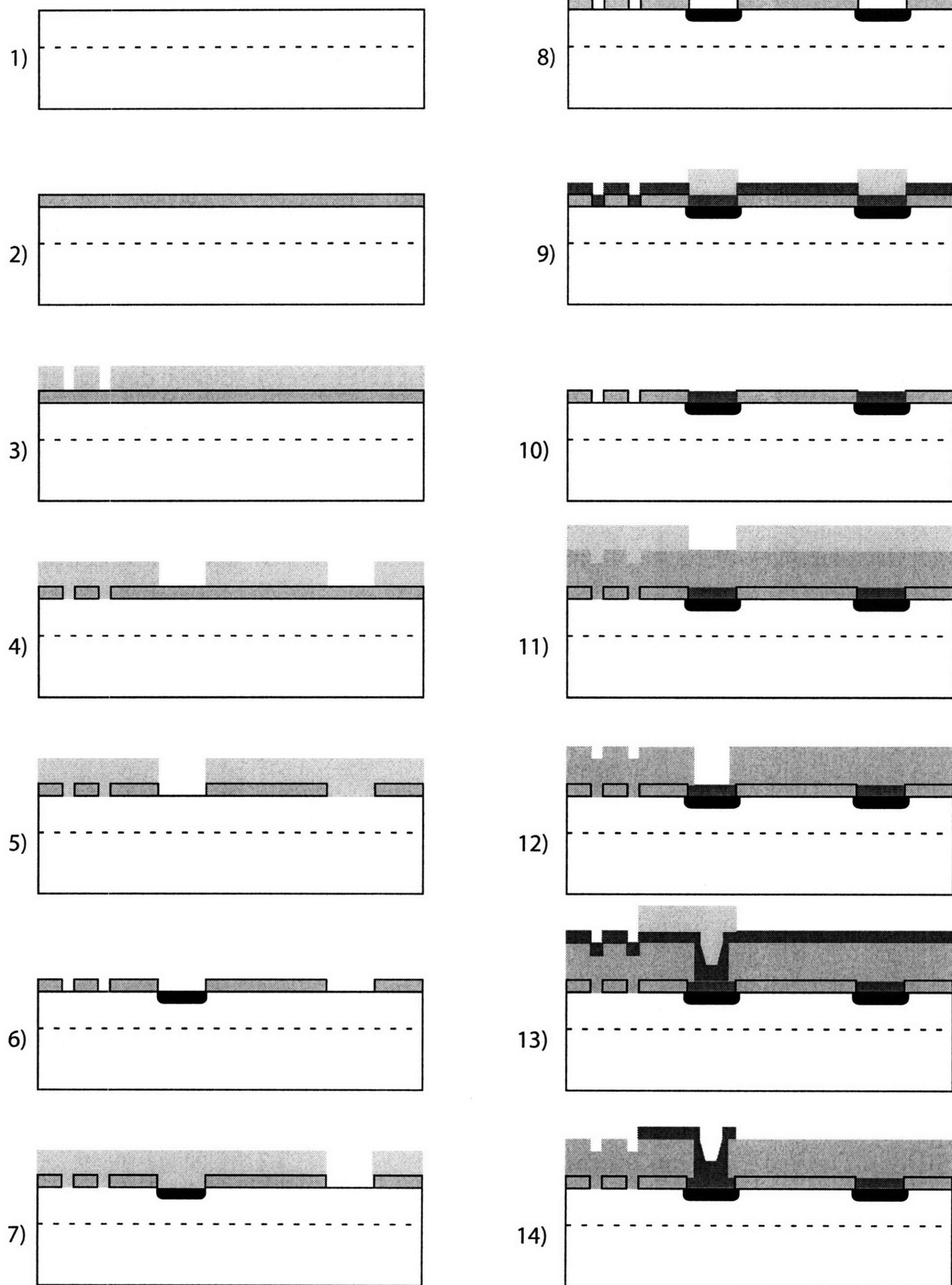


Figure 4-7: Silicon lateral detector process flow schematic diagram indicating selected important steps.

5. Mask step 3 - photolithographic pattern to define implant mask for boron implant.

6. Boron implantation, post-implantation clean: 3 minutes oxygen plasma ash and two piranha cleans (10 min 3:1 H_2SO_4:H_2O_2 each).

7. Mask step 4 - photolithographic pattern to define implant mask for phosphorus implant.

8. Phosphorus implantation, post-implantation clean: 3 minutes oxygen plasma ash and two piranha cleans (10 min 3:1 H_2SO_4:H_2O_2 each). RCA clean, post-implantation anneal (950 °C, 30 minutes in nitrogen ambient).

9. Pre-metal clean (piranha, 15 seconds 50:1 DI H_2O:HF), metal sputter deposition (250-500 nm AlSi(2%)). Mask step 5 - photolithographic pattern to define ohmic contact fingers over implanted regions.

10. Metal dry etch (Cl_2, BCl_3, Ar plasma - duration determined by instrument endpoint monitor). Metal sinter (400 °C, 30 minutes in 95:5 N_2:H_2 ambient).

11. PECVD SiO_2 deposition - 500 nm. Mask step 6 - photolithographic pattern to define via opening to ohmic contact metal.

12. Dry oxide etch (C_2F_6 plasma) to etch via opening, stopping on underlying metal.

13. Metal sputter deposition (750 nm AlSi(2%)). Mask step 7 - photolithographic pattern to define contact pads.

14. Metal dry etch (Cl_2, BCl_3, Ar plasma - duration determined by instrument endpoint monitor).

4.2.2 Silicon Device Process Comments

Metallisation and Implantation

After deposition of an ohmic contact metal to a semiconductor surface, the sample is sintered to encourage alloying and repair at the interface, thereby improving the electrical contact. Aluminum and silicon form a solid solution with up to nearly 1.5% concentration just below the eutectic temperature. Physically, this entails diffusion of silicon at the interface into the aluminum,

leaving voids in the silicon wafer. The aluminum may fill these voids and potentially short-circuit a junction if it is sufficiently shallow. This problem has a long history in semiconductor manufacturing, and may be solved in multiple ways.

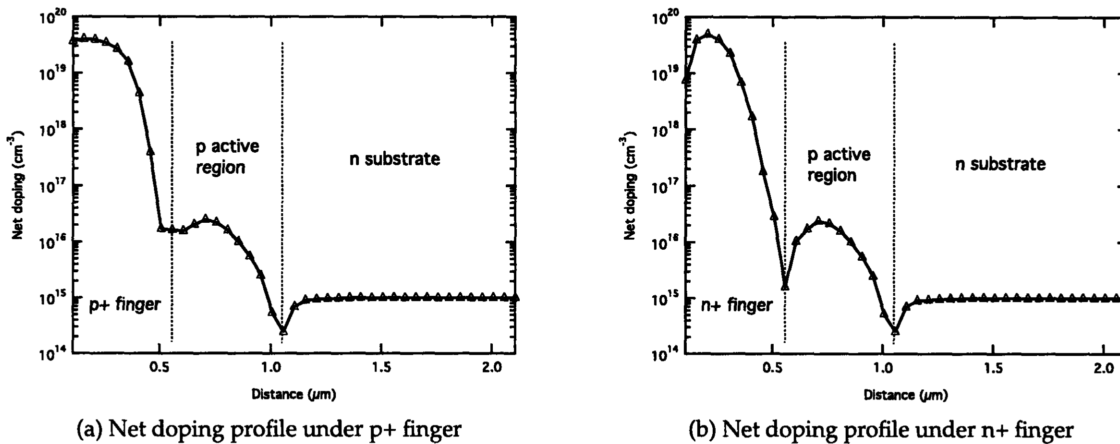
In this work, it was chosen to directly deposit an alloy of AlSi(2%). Additionally, implantation conditions were chosen to form relatively deep junctions (0.4 μm from the surface), such that any eventual spiking would not reach the junction. In a small number of devices, the metal fingers did not adhere to the silicon after etching, more precisely in the wafer spin-rinse procedure following etching to remove unreacted chlorine. In the lot 5 devices, the annealing and etching steps were reversed in order to determine whether such a sequence would improve adhesion. After the anneal, the metal on the oxide field was visually rougher than directly after deposition, and after the metal etch a large number of small grains remained on the surface. Further investigation revealed anecdotal evidence that the dissolved silicon may precipitate from the deposited alloy upon cooling, forming silicon islands on silicon dioxide [82].

An alternative method more commonly used in current fabrication procedures involves the deposition of a diffusion-barrier metal, typically Ti. An ohmic contact stack consisting of a thin (25-50 nm) layer of Ti followed by Al to a desired thickness would be deposited. This method was not chosen, primarily because the dry metal etch apparatus endpoint monitor does not include a signal for Ti etching. Therefore, the Ti etch would have to be timed which could result in non-uniformity or over-etching that might damage an unintentionally expose Si surface below.

Metals were deposited by sputtering. An advantage of this method is that a multi-chamber system may include, without a large increase in cost or complexity, an *in situ* cleaning step, for example to remove native oxide passivation from an Al surface before the deposition of further Al, as in the case of a multi-level metal process. It was believed that the system used in this work was configured with such a pre-deposition clean, however it was discovered only after the completion of Lot 5 that this step was omitted in the standard sequences prepared by the instrument engineer. The lack of this clean is believed to be responsible for the majority of series resistance seen in lots 3-5 devices.

Implantation conditions for the p- active region were chose to provide a reasonably flat carrier concentration and deep junction. For lots 2 and 3, a two-step implant sequence was used, and for lots 4 and 5, a three-step sequence was used. The thermal oxidation and later dopant anneal step

served to flatten the boron active region profile. The n+ and p+ finger implantation conditions were chosen to provide a high surface doping concentration ($> 10^{18}$ n-type, $> 10^{19}$ p-type) for good ohmic contact formation, and a relatively deep junction to avoid possible spiking effects as discussed above. Devices in lot 2 had a mid-1e17 surface n-type concentration which was too low and resulted in poor contact behaviour. The dose was raised by an order of magnitude for subsequent lots. Implantation of the fingers was done directly into exposed silicon. The consequence of this decision is that there was some dopant diffusion out of the wafer during the annealing step. Alternatively, a thin layer of oxide may have been left during the etch, and implantation may have occurred through it. The oxide acts as a dopant diffusion barrier during annealing. However, this sequence was not chosen due to concerns about etching uniformity of the remaining oxide (prior to metal deposition), and the possible requirement of an additional mask for that step. Implantation was simulated in TSUPREM4, and a sample result is shown in Figure 4-8.



(a) Net doping profile under p+ finger

(b) Net doping profile under n+ finger

Figure 4-8: TSUPREM4 simulation of doping profiles following implantation and annealing

Via etch

The via etch through the passivation oxide to expose the underlying ohmic metal (step 12) was performed in two ways. For lots 1-3 (half of the wafers in lot 3), a BOE wet etch was used. The disadvantages of this etch are its isotropic behaviour (widening of the patterned feature during the etch) and poor selectivity of oxide and aluminum etch rates. The later problem was exacerbated by the slower than expected and non-uniform etch rate of the PECVD oxide. The etch rate as well as the colour progression of the oxide during etching (during the final approximately minute

the oxide appears dark grey-brown and rough) suggest a silicon-rich condition, making a clear judgement of required etch time difficult. Additionally, the aluminum etch rate is faster than the oxide etch rate in the final stages, often resulting in significant undesired ohmic metal etching. One solution to this problem, used on some wafers in lots 2 and 3, is the deposition of a thin Ti layer over the Al ohmic metal. Though Ti is etched by HF, the etch rate is significantly lower in the buffered solution compared to an aqueous solution. This does not however address the issue of feature widening. An alternative wet chemistry is the commercial etchant Silox Vapox manufactured by Transene Company Inc., based on aqueous HF and containing acetic acid and a corrosion inhibitor to enhance the oxide-aluminum selectivity.

It is preferable to perform the via etch using only plasma etching. However, an appropriate chemistry is required to maintain oxide-aluminum selectivity. For lots 3-5 processing, permission was granted for use of a high density plasma etcher based on C_2F_6 etchant. The process is operated on the boundary between etching and deposition, where exposed oxide is etched and remaining chip area is covered with a passivating polymer. The disadvantage of this tool is the apparent incompatibility of the standard photoresist with the processing temperatures. The result was a thin remaining polymer layer on the passivation oxide surface that could not be removed, even after several oxygen plasma and piranha cleans. This resulted in a rougher than expected pad metal layer deposition. This remaining polymer film is also the reason for not using this etcher in the front-end process (finger etch through oxide), since it was found that the film affected subsequent photolithography conditions.

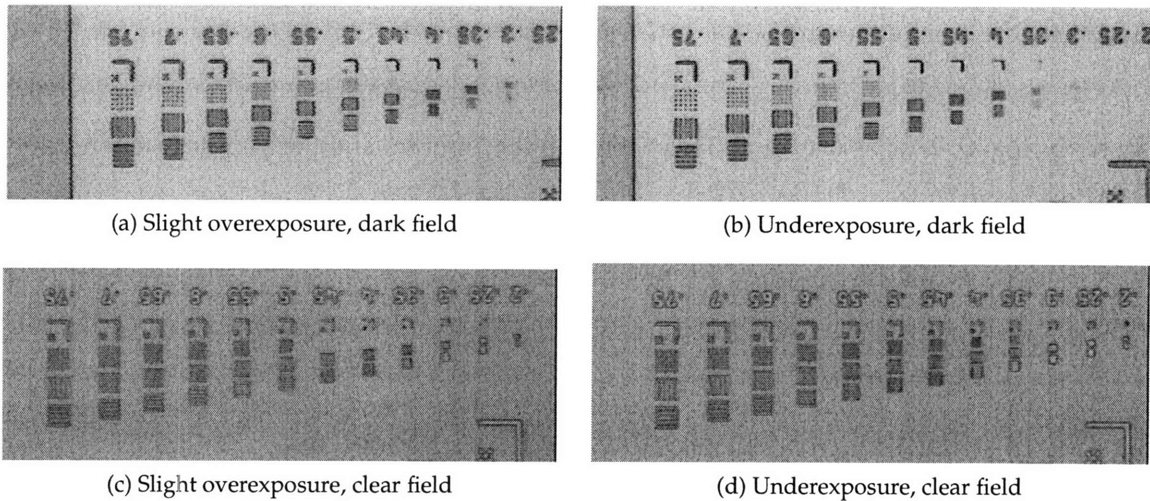
Photolithography

The lateral p-i-n device layout was designed with some variation/misalignment tolerance in mind, however it is important for high yield that feature dimensions are faithfully transferred from the lithography mask to the wafer surface. For a clear field pattern, overexposure may result in smaller than planned features and underexposure may result in larger than planned features. For a dark field pattern, the relationship is of course reversed. Additionally, auto-focus errors may result in inexact reproduction of features. It was confirmed after lot 1 processing that the exposure conditions originally suggested by the instrument engineer resulted in clear-field features

that were 0.2 μm narrower than expected and dark-field features that were 0.2 μm wider than expected. This is a significant error considering that the p and n finger (and ohmic contact metal) widths are 1 μm as drawn. The accepted operating procedure of the stepper was to confirm features dimensions by SEM after processing to determine a correct exposure condition. However, optimum exposure conditions may vary depending on the reflectivity of the underlying material, resist application, environment conditions, and instrument-related variation. For example, in March 2006 after a repair and recalibration by a Nikon service engineer, the optimum exposure time for the definition of the oxide finger etch changed from 135 to 165 ms.

Using a standard test reticle supplied by the manufacturer, a new procedure was employed to find an optimum condition for each reticle prior to each lithography step. The test reticle includes a set of clear field and dark field patterns of lines and dots with dimensions labelled in 0.1 or 0.05 μm steps. The stepper also includes a standard program ("TPR") that exposes a matrix of different exposure times and focus offsets of the same test pattern. It is therefore possible to quickly evaluate the matrix to find the optimum conditions. Optimum exposure time is determined by finding the time for which the minimum visible feature size is the same for the clear and dark field patterns. Optimum focus offset is determined by finding the condition under which the smallest features are reproduced. It is also useful to examine exposure times above and below the optimum in order to change the actual on-wafer feature without changing the mask. An over or underexposure of 20 ms may reduce or enlarge features by 50-100 nm on each side. This was used, for example, when exposing the implant mask patterns. In those cases, an overexposure is helpful to move the mask pattern away from the oxide opening edges in order to minimise implant shadowing by the resist. The use of the test reticle in this way prior to each photolithography step increased device yield significantly with only a minor increase in processing time. Examples of the test reticle for different exposure conditions are shown in Figure 4-9.

The stepper may be programmed to exposure an entire reticle or an arbitrary rectangular area within it (or more precisely up to four areas in a physical reticle for a given logical reticle definition). It is also possible to define the logical reticles as a part of the physical reticle with an arbitrary shift in the x and y directions. This property may be useful in many applications. In this work it was used for two reasons: first to isolate alignment marks and print them in desired locations on the wafer (not where originally drawn), and second to reduce reticle cost by drawing differ-



(a) Slight overexposure, dark field

(b) Underexposure, dark field

(c) Slight overexposure, clear field

(d) Underexposure, clear field

Figure 4-9: Exposure test patterns showing the results of slight overexposure (distinguishable features: 0.45 μm in dark field, 0.5 μm in clear field) and significant underexposed (distinguishable features: 0.5 μm in dark field, 0.4 μm in clear field).

ent patterns in each quadrant of the reticle (in effect each physical reticle contained four different die patterns). It was found during Si lot 4 and Ge lot 1 processing that a systematic offset was introduced by the stepper when aligning one quadrant to another. This offset is not an alignment error, but a very precisely repeatable systematic offset that does not vary with wafer material or topology. The effect of this offset may be significant and is discussed in section 5.2.1. A set of test wafers was prepared to evaluate the offset problem in the following way.

1. Starting with a test grade Si wafer, deposition of 250 nm of PECVD silicon dioxide.
2. Photolithography, exposure using a reference reticle (top right quadrant, dark field pattern including alignment marks).
3. Dry etching of silicon dioxide through the photoresist pattern.
4. Metal deposition - 250 nm Al.
5. Photolithography, exposure using one of the three remaining reticles of interest.
6. Dry etching of aluminum through the photoresist pattern.

Evaluation of the resulting sample using SEM, compared with the designed patterns, revealed a systematic offset that varied among reticles. The offset was in the x direction in all cases, and

also in the y direction in the case of two reticles. Additionally, the offset varied in some cases with the y-axis position of the pattern within the reticle. Using the measured offsets, the systematic offset of the logical reticles was modified in the stepper program and the samples were repeated. The observed offset was reduced to below 50 nm, the alignment resolution of the stepper. The cause of the offset is either a mask manufacturing error or an error in the optical alignment of the stepper.

4.3 Germanium Photodetector Fabrication Procedure

Germanium lateral photodetectors were fabricated based on the same design as the silicon devices. The fabrication procedure is inspired by that used in vertical Ge-on-Si p-i-n structures, where the growth substrate acts as the p-contact and an implanted polysilicon acts as the n-contact [83, 34]. In the case of the lateral devices, the vertical structure is "folded" and appropriately implanted polysilicon acts as the n and p contacts respectively. The procedure is different than the only other report of Ge-on-Si lateral devices [41] since the germanium is not directly metallised, and metallisation takes place by standard deposition and etch techniques rather than by lift-off. The germanium surface in this work is covered by a passivation dielectric that may also act as an anti-reflection coating. Additionally, the other report concentrated on small devices patterned by e-beam lithography, with the germanium grown on very thin SOI. Several steps in the fabrication procedure refer to RCA and/or Ge RCA cleaning steps. A discussion of these cleaning steps may be found in appendix A.

4.3.1 Germanium growth and preparation

The epitaxial germanium is grown by UHVCVD in a two step process that involves a slow low-temperature thin buffer growth (< 400 °C, ≈ 60 nm) and a faster high-temperature step (> 600 °C, $\approx 1\mu$m). The growth procedure is discussed in detail elsewhere [59, 32, 33]. A key aspect of the epitaxy is a high-temperature anneal following the thick layer growth which reduces point defect and threading dislocation density. The anneal may take place *in situ* following the thick layer growth, or may be carried out by rapid thermal annealing or cyclic anneal after the growth. In the latter case, the germanium surface must be protected during the anneal, particularily if the

vacuum conditions are not UHV. Three possible annealing scenarios are shown in Figure 4-10. The *in situ* case is shown in the third column, whereby device processing is carried out on the wafer as it exists the growth system with no further preparation. For a RTA (or cyclic process), further steps must be taken.

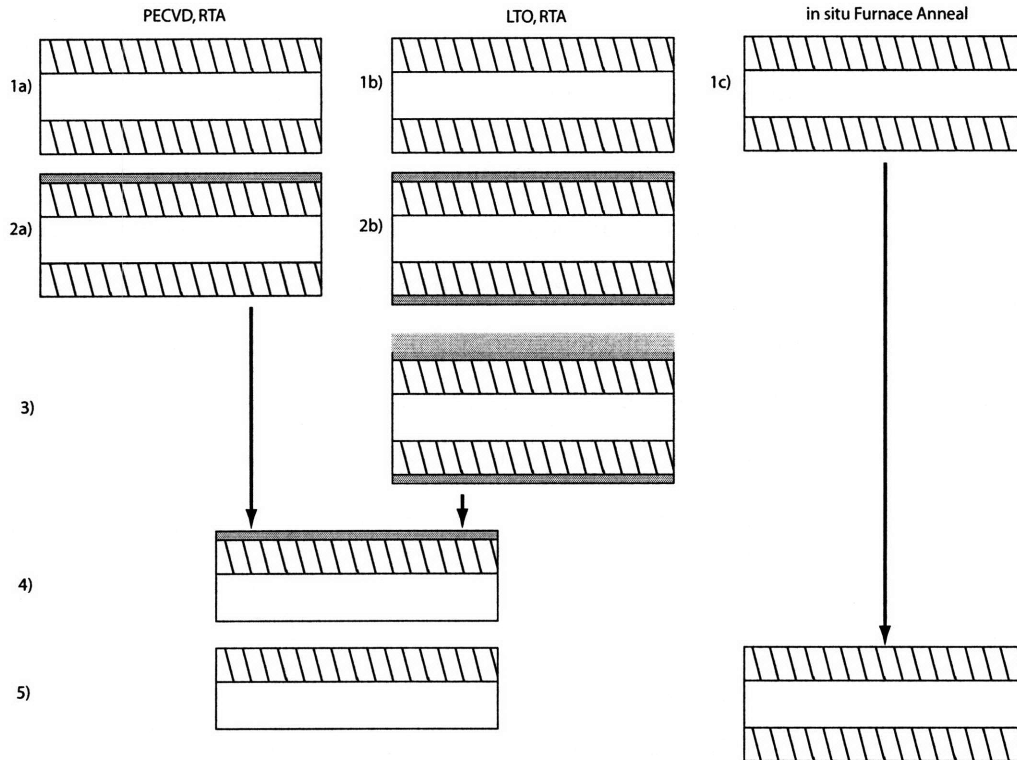


Figure 4-10: Germanium lateral detector process flow schematic diagram from growth to after the annealing step for different annealing methods.

1. Germanium growth by UHVCVD. Germanium is present on both front and back sides of substrate.

2. Ge RCA clean, protective oxide deposition. The oxide may be PECVD SiO_2 (case a) or low temperature furnace-deposited oxide (LTO, case b).

3. In the case of LTO, the front surface is protected by photoresist and the back side oxide is etched by BOE.

4. Backside germanium is removed in a 1:3:4 H_2O_2:HCl:H_2O solution (etch rate > 100 nm/min).

This step is necessary because the annealing instrument uses a pyrometer to measure substrate temperature, and the pyrometer is calibrated for an unpolished Si surface. Next, rapid cyclic annealing takes place - 10 cycles of 30 seconds at 650 °C followed by 30 seconds at 850 °C.

5. Protective oxide is removed by BOE. If necessary, germanium film thickness is reduced to a desired level by etching in 1:6 H_2O_2:H_2O solution (etch rate ≈ 60 nm/min).

Anecdotal evidence exists that a silicon-germanium alloy may be formed at the substrate-epitaxial film interface during growth, confirmed by Rutherford backscattering spectroscopy (RBS), therefore leading some research groups to use a polysilicon dry etch in step 4 to be certain of the removal of all germanium [84]. This phenomenon has not been observed in as-grown, not annealed, UHVCVD samples by X-ray diffraction rocking curve analysis as shown in Figure 4-11, where only Si substrate and Ge epitaxial layer peaks are distinguishable.

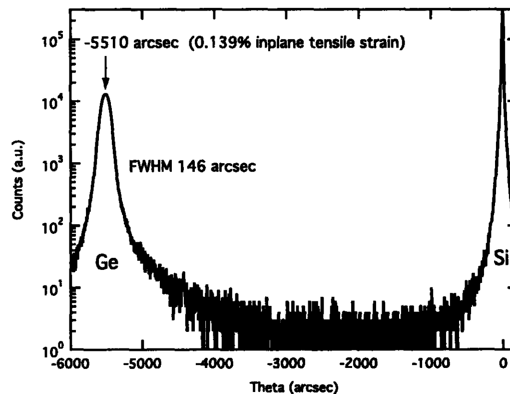


Figure 4-11: (004) rocking curve of 700 nm as-grown epitaxial UHVCVD germanium on silicon

It has also been reported that substantial silicon-germanium alloying may take place during annealing, as observed by X-ray diffraction. The presence of SiGe in the film may decrease the observed photodiode responsivity. This is the motivation behind epitaxial growth on ultra-thin SOI, so as to minimise the amount of Si available for alloying [40, 42] This effect was previously not observed in furnace-annealed germanium films on bulk substrates grown within the author's research group, and it was thought that the alloying effect may be enhanced by the strain in the ultra-thin SOI substrate. However, in this work an X-ray rocking curve of epitaxial Ge annealed by rapid cyclic annealing did exhibit some Ge-rich alloying, as seen in Figure 4-12. The shape of the

curve would appear to support the claim of an approximately 5% Si layer as stated in the citations above. More investigation is required on the annealing conditions that result in SiGe alloying.

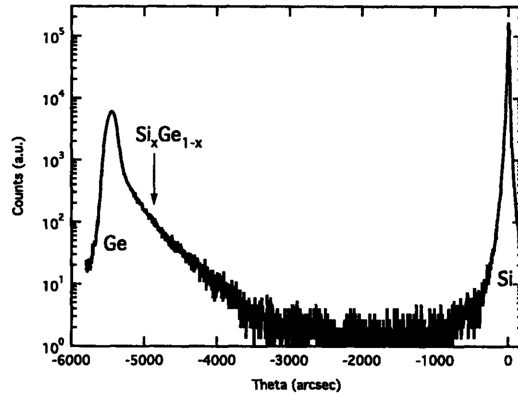


Figure 4-12: (004) rocking curve of 700 nm cyclic annealed epitaxial UHVCVD germanium on silicon

4.3.2 Detailed Description of Process

A schematic view of the fabrication process is shown in Figure 4-13. The description that follows refers to the step numbers in the figure. The figure assumes an annealed Ge-on-Si film as the starting material, regardless of the annealing method as discussed in the previous section. A precise description of the entire process flow including relevant laboratory instrument names/designations is presented in Appendix B.

1. Begin with epitaxial Ge-on-Si wafer, annealed using one of three procedures as discussed in section 4.3.1.

2. Ge RCA clean, deposition of passivation dielectric. The dielectric may be LTO, PECVD silicon dioxide or oxinitride.

3. Mask step 1 - photolithographic pattern and oxide etch to define stepper alignment marks and contact finger openings. Ge RCA clean and undoped LPCVD polycrystalline silicon deposition (200 nm).

4. Mask step 2 - photolithographic pattern to define implant mask for boron implant.

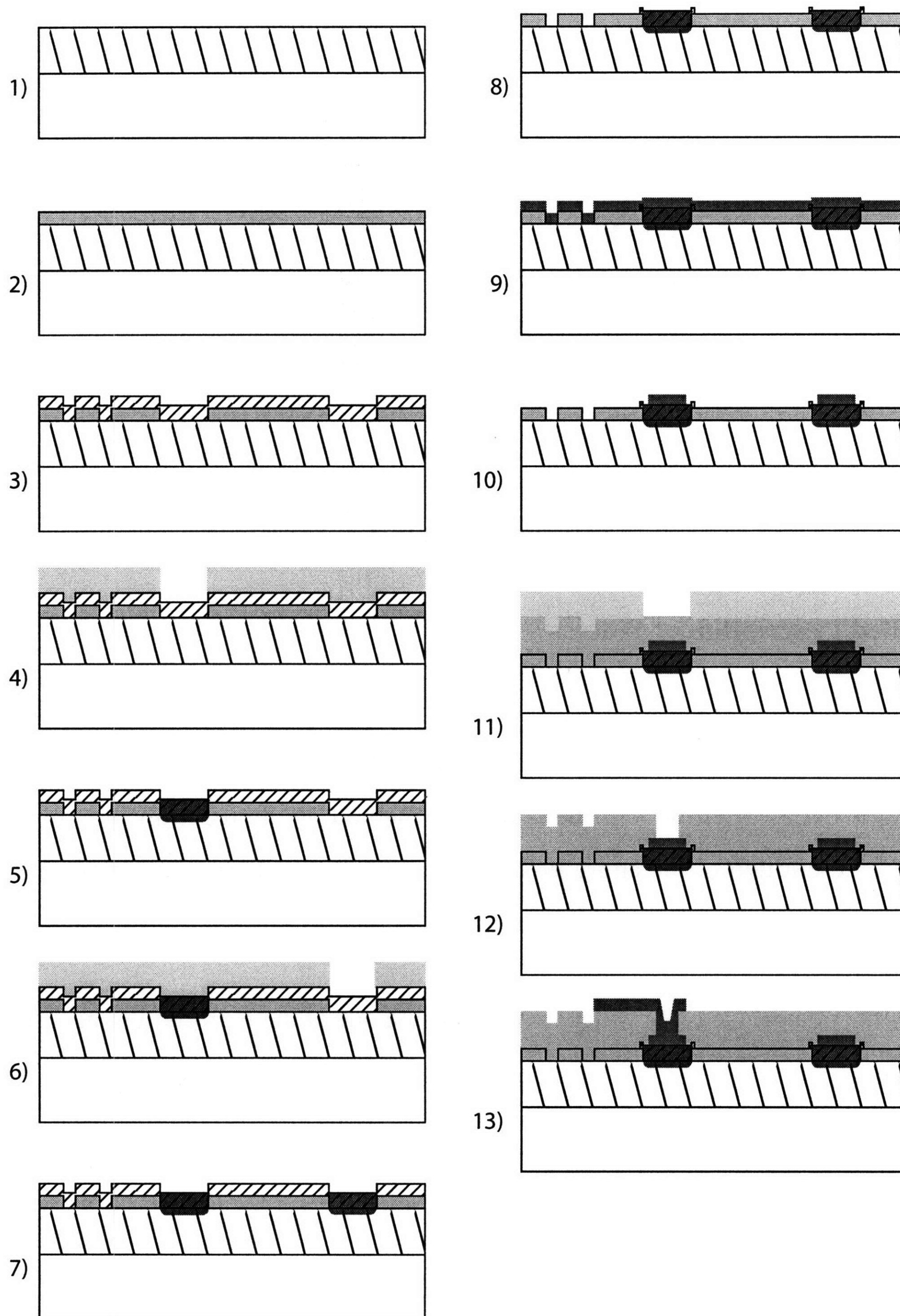


Figure 4-13: Germanium lateral detector process flow schematic diagram starting with annealed wafers.

5. Boron implantation, post-implantation clean: 3 minutes oxygen plasma ash and two piranha cleans (10 min 3:1 H_2SO_4:H_2O_2 each).
6. Mask step 3 - photolithographic pattern to define implant mask for phosphorus implant.
7. Phosphorus implantation, post-implantation clean: 3 minutes oxygen plasma ash and two piranha cleans (10 min 3:1 H_2SO_4:H_2O_2 each). RCA clean, post-implantation anneal (650 °C, 30 minutes in nitrogen ambient).
8. Mask step 4 - photolithographic pattern to define polysilicon contact fingers (pattern slightly wider than dielectric openings such that germanium is not exposed). Polysilicon dry etch (Cl_2, HBr - duration determined by instrument endpoint monitor).
9. Pre-metal clean (piranha, 15 seconds 50:1 DI H_2O:HF), metal sputter deposition (250-500 nm AlSi(2%)).
10. Mask step 5 - photolithographic pattern to define ohmic contact fingers over implanted regions. Metal dry etch (Cl_2, BCl_3, Ar plasma - duration determined by instrument endpoint monitor). Metal sinter (400 °C, 30 minutes in 95:5 N_2:H_2 ambient).
11. PECVD SiO_2 deposition - 500 nm. Mask step 6 - photolithographic pattern to define via opening to ohmic contact metal.
12. Dry oxide etch (C_2F_6 plasma) to etch via opening, stopping on underlying metal.
13. Metal sputter deposition (750 nm AlSi(2%)). Mask step 7 - photolithographic pattern to define contact pads. Metal dry etch (Cl_2, BCl_3, Ar plasma - duration determined by instrument endpoint monitor).

4.3.3 Germanium Device Process Comments

Material Quality

A full discussion of epitaxial material quality is beyond the scope of this thesis. Collaborators were responsible for the epitaxial growth and general materials characterisation. However, several important points will be mentioned in this section.

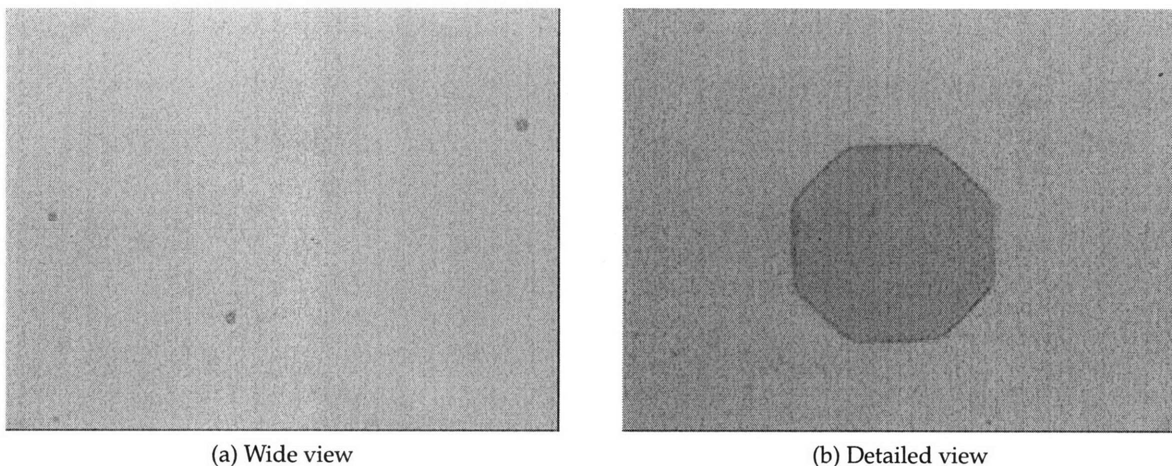
The epitaxial material for Ge lot 1 devices was taken from several growths on p-type Si substrates, with nominal thicknesses in the range of 400-800 nm. The surface roughness was above 3nm RMS and the material was believed to be undoped. Visually, the material was shiny with isolated areas of polycrystalline growth on two of the wafers due to cleaning problems. During processing, particularly the protective PECVD deposition and cyclic annealing, there was no visual change in material quality. Four point probe measurement was not performed by the author.

The epitaxial material for Ge lot 2 devices was taken from 3 dedicated growths. All growth thicknesses were nominally 1.1 μm, epitaxial layers were optionally thinned by 1:6 H_2O_2:H_2O etchant (approximately 60 nm/min etch rate) either after protective oxide removal (annealing step 4) or by extending the second solution dwell time in the first Ge RCA procedure (step 2). Compared to the lot 1 devices, the GeH_4 gas source was changed from 15% GeH_4 in He to 100 % GeH_4, the growth temperature was changed, and the cleaning procedure was altered. The first two changes resulted in a much higher growth rate (approximately 500 nm/hour) than prior (approximately 200 nm/hour). The cleaning procedure was altered from a piranha clean followed by dilute HF dip to an RCA clean followed by baking *in situ* at 750 °C under hydrogen flow. The oxide desorption temperature is typically stated as approximately 800 °C [85], but this method also resulted in visually good growth. The roughness was below 1 nm RMS, suggesting improved material quality.

Four point probe measurement indicated a sheet resistance of 35 $\Omega/\square$ for furnace-annealed material (consistent over several wafers). For confirmation, one wafer was etched back and the measured sheet resistance rose proportionally to the smaller Ge thickness. The sheet resistance corresponds to a p-type doping of 6×10^{18} cm^{-3} or an n-type doping of 2×10^{18} cm^{-3} [86]. This is much higher than previous epitaxial germanium growths in an older 100mm UHVCVD system used to develop the growth process where the unintentional doping was calculated from C-V data to be under 1×10^{16} cm^{-3}. Such a high apparent doping is expected to result in a significantly higher than designed capacitance and shorter depletion region extent.

During cyclic annealing, the lot 2 epitaxial Ge material exhibited the formation of large octagonal defects, with the sides of the octagon following crystallographic directions. The width of the largest defects approached 30 μm, and the depth was the entire thickness of the epitaxial layer. The number of defects varied between wafers, generally with more present in the bottom half of

the wafers (i.e. closer to the flat which faces upwards in the growth chamber). Examples of these defects are shown in Figure 4-14. The spacing was quite large, with approximately 100-150 defects per wafer. Wafers protected by LTO had fewer octagonal defects, but a larger number of small dots that appeared to be pinholes. Furnace-annealed wafers exhibited a slightly smaller number of pinholes and no octagons. The origin of these defects was not fully explained, but is likely related to a combination of factors: incomplete pre-epi cleaning, porosity of the protective oxide, and growth defects.



(a) Wide view

(b) Detailed view

Figure 4-14: Optical microscope pictures of defects formed during the cyclic annealing process.

Photolithographic alignment

As discussed in section 4.2.2, the problem of a systematic unexpected alignment offset was discovered during the course of Si lot 4 and Ge lot 1 device processing. The consequences for the germanium devices are more severe than those for silicon devices for two reasons. First, the diffusion coefficient of boron and phosphorus dopants in germanium is significantly higher than in silicon. For this reason the post-implantation anneal is carried out at 650 °C [83]. This temperature is high enough for dopant activation and annealing of the polysilicon following the damage due to implantation, and results in a reasonably short diffusion distance in germanium. In polysilicon, for practical purposes, there is no boron or phosphorus diffusion at that temperature, and therefore no significant lateral dopant diffusion in the polysilicon fingers. This is different than in the case of the silicon dopant anneal where significant (100-150 nm) lateral diffusion occurs. There-

fore, the tolerance for implantation-ohmic metal alignment is more demanding in the germanium devices.

The second consequence has to do with the polysilicon etch to dielectric opening alignment. In this case, a misalignment causes the polysilicon protecting the underlying germanium to be etched. If this occurs, the subsequent piranha pre-metal clean causes rapid etching of the germanium and undercutting of the passivation dielectric. An example of this effect is shown in Figure 4-15.

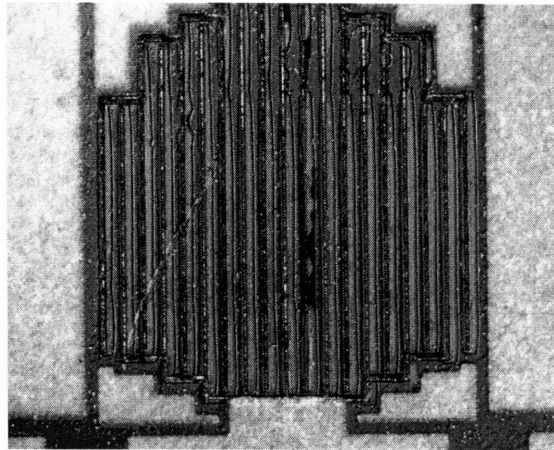


Figure 4-15: Optical microscope photo of undercutting of germanium epilayer due to polysilicon etch misalignment. The undercut area is to the right of each finger.

Metallisation

The germanium device design called for a 0.5 μm wide ohmic metal line on top of the 1 μm wide doped polysilicon electrodes. During Ge lot 1 processing, adhesion problems were observed with metal electrodes after the etch process and before annealing. After the dry etch, the wafers are spin-rinsed and dried to clean away unreacted chlorine from the wafer surface. During this process, a large number of the ohmic metal lines detach, particularly towards the edge of the wafer. An example of this effect is shown in Figure 4-16.

To correct for this effect in Ge lot 2, it was decided to reverse the annealing and etching steps. It was hoped that the alloying taking place during the annealing would enhance metal adhesion. Unexpectedly, the annealing of the blanket metal resulted in a very poor wafer surface. The wafers appeared clean and as expected following premetal clean and metal deposition. However, during

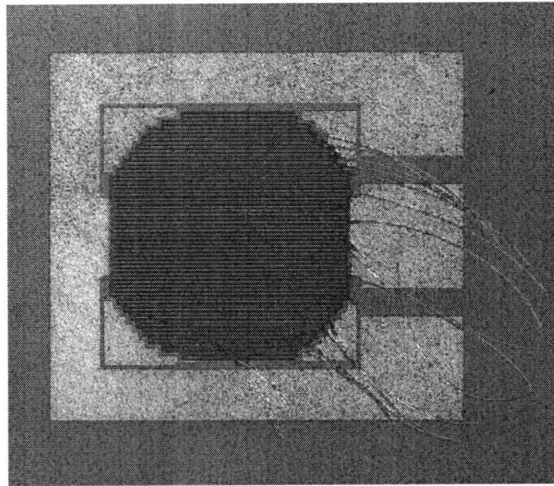
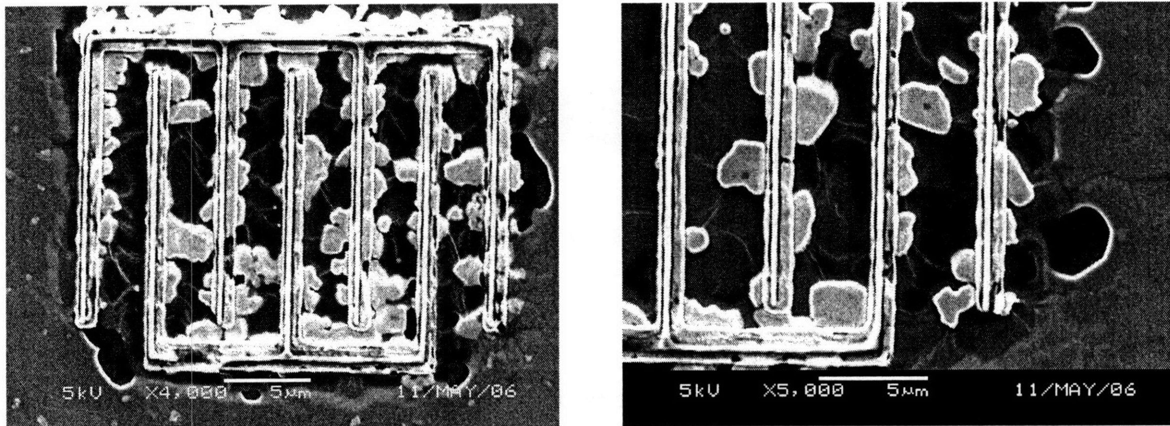


Figure 4-16: Optical microscope photo of metal fingers that did not adhere to the polysilicon contacts, but remained fixed to the substrate by the wider metal bus.

the annealing a large number of features (large cavities or hills) were formed apparently in the polysilicon and germanium films.



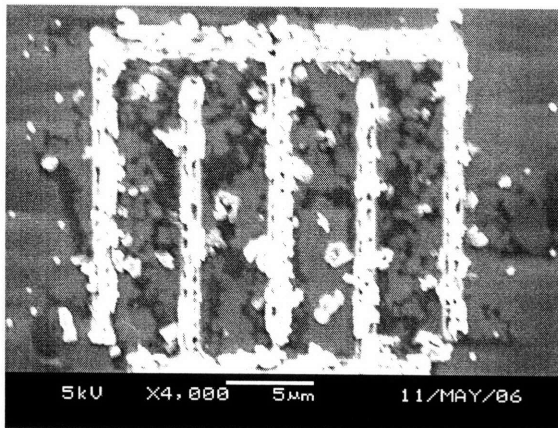
(a) Wide view

(b) Detailed view

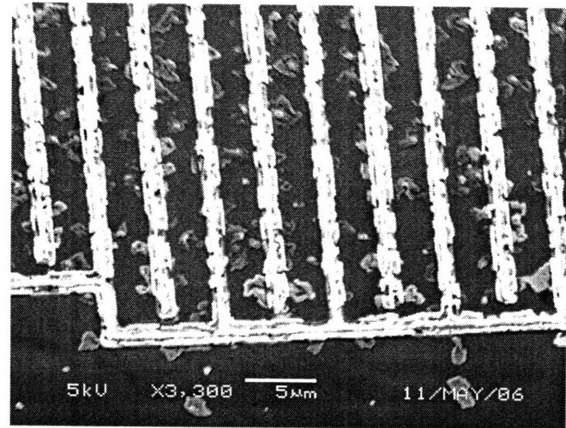
Figure 4-17: SEM Images of Ge Lot 2 devices following blanket AlSi(2%) deposition and anneal (400 °C, 30 minutes). The images suggest the formation of large craters in, or general disappearance of, the metal film in and around the devices.

Figure 4-17 shows the wafer surface after blanket metal deposition and anneal. There appear to be craters in the metal between and beside the polysilicon electrodes. The polysilicon itself is in tact, however with indentations on some edges. The thin pattern on top of the polysilicon is the ohmic metal etch pattern (i.e. photoresist), exhibiting non-ideal alignment or offset.

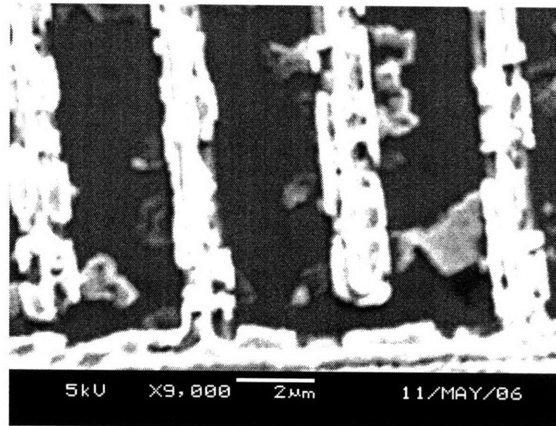
Figure 4-18 shows the germanium device process after metal deposition, anneal, and metal



(a) Top-down view



(b) 40 degree tilt, wide view



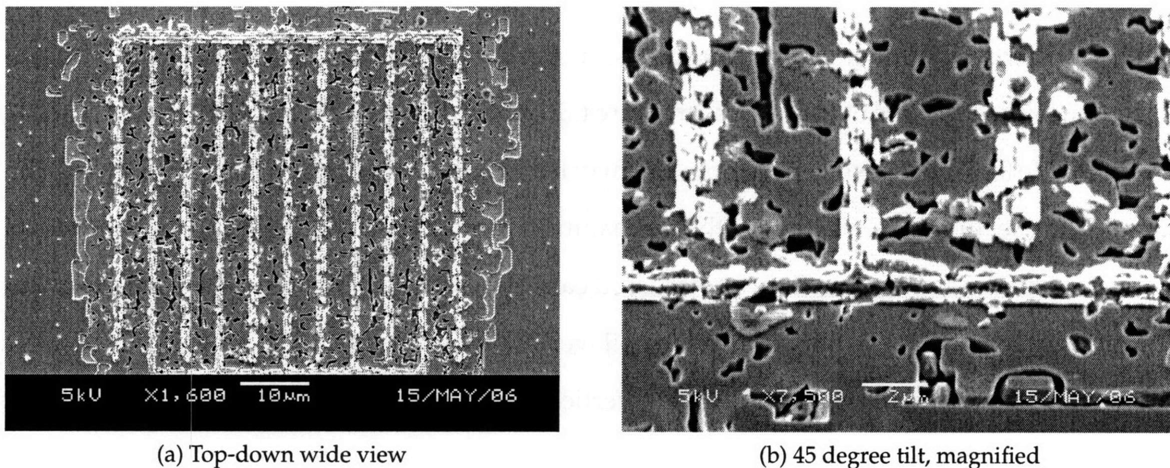
(c) 40 degree tilt, detailed view

Figure 4-18: SEM Images of Ge Lot 2 devices following blanket AlSi(2%) deposition, anneal (400 °C, 30 minutes), and metal etch.

etch. From the tilted images, it appears that the polysilicon fingers are generally in tact. However, in all of the images it is apparent that not all material was etched from the finger surfaces. The images are ambiguous, because it appears that some of the thin metal pattern in also present on the fingers, but it is difficult to determine whether the polysilicon fingers are covered in additional jagged metal or whether the fingers themselves are substantially degraded. The top-down device view shows a sort of shadow pattern between the fingers which is also apparent under an optical microscope as a discolouration. Some of these shadows extend several microns away from the polysilicon fingers. From the two images taken with a 40 degree tilt, it appears that many of these shadow features are craters or hills below the surface of the passivation dielectric. The features were present only around the device area and not in the field, indicating that the effect was due

to an interaction between the AlSi(2%), the implanted polysilicon fingers, and the underlying germanium. It is thought unlikely to be a simple problem of polysilicon porosity or gap between the polysilicon and passivation dielectric because the wafers appeared normal during post-piranha clean inspection (as opposed to Figure 4-15 for example).

As a final analysis step, the passivation SiON/LTO and metal was etched by a long BOE dip, with results show in Figure 4-19. The images reveal severe destruction of the Ge material, showing that the shadowed features present in Figure 4-18 are craters in the Ge film covered by the overhanging passivation dielectric. The irregular shape of the pits suggests that they are not caused by penetration of the pre-metal clean or solely by crystal defects. It should be noted that these pits exist only within and around the devices, and not in the field of the die. The polycrystalline silicon fingers are also severely damaged, with sections missing generally around a pit location.



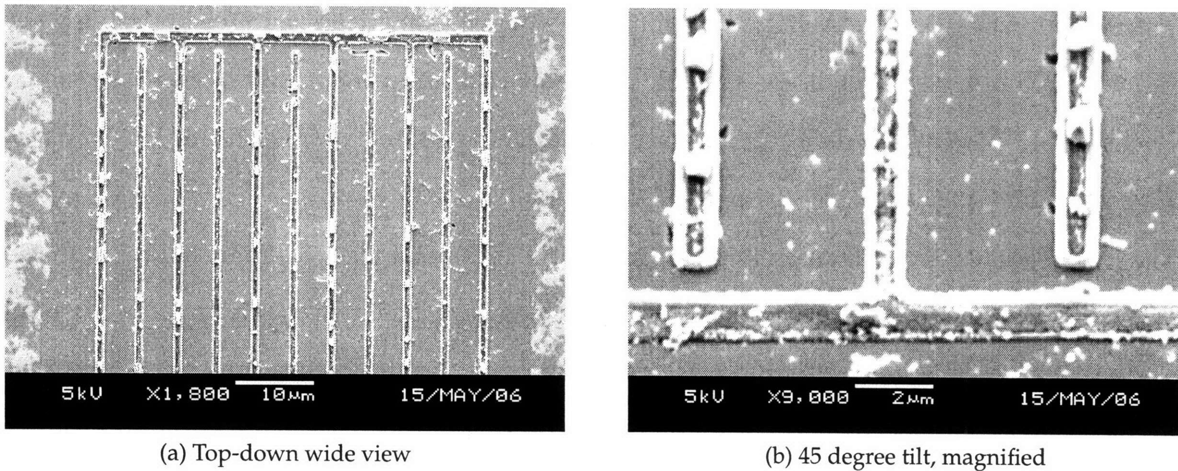
(a) Top-down wide view

(b) 45 degree tilt, magnified

Figure 4-19: SEM Images of Ge Lot 2 devices following blanket AlSi(2%) deposition, anneal (400 °C, 30 minutes), metal etch, and BOE dip exposing the Ge surface.

Two wafers were processed with the same process flow as Ge lot 1 (etch before anneal). Wafers were inspected after every step, and a small number of features appeared after the metal anneal step. Following processing completion, a long BOE dip was used to etch away all metal and passivation layers, with the result shown in Figure 4-20. Some small holes may be seen, primarily near the polysilicon fingers, however the damage is insignificant compared to the images of Figure 4-19. Aside from the order of the metal etch and anneal steps, the processing of the two samples was identical starting from the epitaxial growth.

It is therefore believed that this process is related to the metal annealing step itself, and might



(a) Top-down wide view

(b) 45 degree tilt, magnified

Figure 4-20: SEM Images of Ge Lot 2 devices processed according to the original process flow, and following a BOE dip exposing the Ge surface. Small sections of metal remained on the polysilicon electrodes and in the field.

be due to wafer stress under heating that accelerates metal interdiffusion. The difference in thermal expansion coefficients of the device structure (aluminum: 23 ppm/K, germanium 5.9 ppm/K, silicon dioxide 0.6 ppm/K, silicon 2.6 ppm/K) may lead to the formation of cracks or pinholes during annealing leading to a reaction of the germanium during the 400 °C anneal. The pinholes or other defects in the germanium film, as discussed earlier, may contribute to such a mechanism due to the non-uniformity of the films above the epilayer. Anecdotal evidence exists that this problem was noticed in the past during fabrication of vertical Ge-on-Si detectors (using the same polysilicon deposition system and phosphorus implant dose and energy as this work). The researcher changed from a standard furnace sinter to a rapid thermal anneal sinter, and the problem was not investigated further.

A possible mechanism worthy of further examination is aluminum-induced crystallisation (AIC) of amorphous silicon. It has previously been reported that amorphous silicon layers deposited on aluminum crystallise at temperatures far below the eutectic or crystallisation temperatures (577 °C and 620 °C respectively) [87]. Additionally, the aluminum is for the most part displaced to the surface of the crystallised silicon (therefore the behaviour is also called aluminum-induced layer exchange, ALILE), a behaviour that cannot be explained by diffusion alone at the temperatures in question. No widely accepted model of this process has been developed, but it is believed that the driving force is the difference in free energy between the amorphous and

crystalline phases and that AIC below the eutectic temperature is a solid phase process. Recently work has appeared confirming this mechanism in amorphous silicon-germanium films deposited by co-evaporation of Si and Ge sources onto aluminum on a quartz substrate [88]. In this case, crystallisation occurs below the Al-Ge eutectic temperature of 420 °C (phase diagram shown in Figure 4.3.3, with a similar behaviour as in the amorphous silicon case. Material from 10% to 90% Ge was crystallised, with the significant finding that irregularly-shaped pinholes approximately 0.5-2 μm in dimension were also formed with density increasing with Ge composition. This circumstance is not yet understood, but was believed to be related to increasing nanocrystallinity of as-deposited material with Ge content.

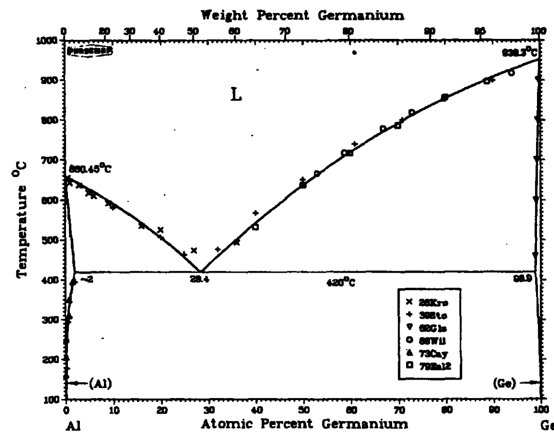


Figure 4-21: Al-Ge phase diagram from [89]

The Ge lot 1 and 2 devices were annealed at 650 °C for 30 minutes after implantation in order to activate the implanted species and recrystallise the polysilicon which is damaged, possibly amorphised, during the implantation. It is possible that this anneal is insufficient to completely repair the material, and additionally it is possible that the edges of the polysilicon fingers are amorphised by the polysilicon dry etch. In such a case, it is hypothesised that an anneal of the blanket AlSi(2%) film may result in a reaction as described above. The silicon fingers and/or their sidewalls (i.e. along the interface between the fingers and passivation dielectric) may serve as a reaction path for aluminum transfer and further reaction with the underlying Ge epitaxial layer. Anecdotal evidence of different polysilicon behaviour following implantation and annealing may be seen in Figure 4-22. The pattern shown in the figure is polysilicon (over passivation dielectric, with no exposure of underlying germanium), with an interior square region subject to boron im-

plantation. The deposited metal was annealed, etched, and passivation dielectric was etched by an extended BOE dip. It is apparent from the sizes and distribution of the grains over the polysilicon region that some inter-diffusion or reaction of polysilicon and AlSi(2%) has occurred, and that the behaviour is different for implanted versus non-implanted regions.

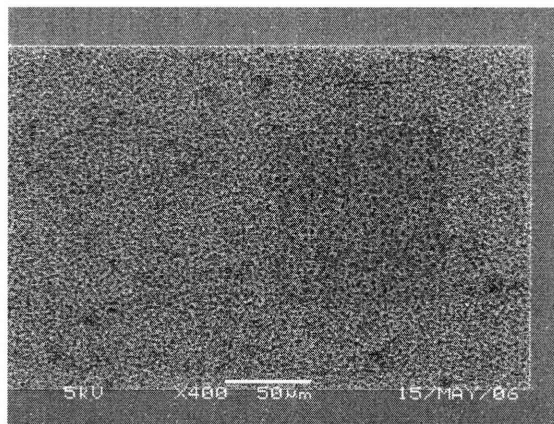


Figure 4-22: Polysilicon test pattern with square boron-implanted region following metal deposition, anneal, metal etch, and surrounding dielectric etch. Grains of different size and distribution are visible on the implanted and non-implanted areas of the polysilicon.

Once the aluminum reaches the germanium layer, a number of reactions are possible including the formation of a eutectic or liquid phase (where subsequent cooling may result in the formation of amorphous Al-Ge metastable regions [89]). This hypothesis would explain the apparently missing areas of Al after annealing, the destruction of the Ge epilayer and parts of the polysilicon fingers, and the fact that the Ge epilayer is only attacked around the polysilicon electrodes. On the other hand, very little damage was seen in the Ge lot 1 samples and lot 2 samples where the metal pattern was etched prior to annealing. In these cases the metal is localised on top of the polysilicon fingers without access to the finger sidewalls. It is therefore possible that the finger sidewalls participate most strongly in the reaction. However, a small number pinholes are also visible near the fingers in the "good" device case. This would indicate a limited effect of migration through or along the sides of the fingers. It is unclear if the process may be initiated during the metal deposition itself where the wafer is heated significantly by the impinging sputtered material. Further work is required to establish a full understanding of the behaviour of the metal-polysilicon-germanium stack, and such an understanding will help increase yield in a large variety of Ge-on-Si detector structures.

Chapter 5

Photodetector Measurement Results

Fabricated photodetectors were measured in several different ways in order to extract relevant properties, and most of the same measurements were performed in simulation to confirm that measurement and simulation were consistent.

- I-V Measurement. Determination of reverse current (i.e. dark/leakage current), diode ideality factor, series resistance.
- C-V Measurement. Determination of device capacitance as a function of reverse bias. The data is also transformed for analysis of junction doping and area, as discussed in section 2.3.1.
- Sensitivity Measurement. Determination of photocurrent vs. incident intensity and photocurrent vs. reverse bias relationships.
- Bandwidth Measurement. Determination of device small signal bandwidth, additionally quantitatively allowing for the identification of diffusion and RC components of the bandwidth, and determination of the dominant mechanism.

The first two measurements are standard device characterisation techniques, carried out in this case using a probe station, HP4145A semiconductor parameter analyser, and an HP4284A LCR meter. The last two measurements were carried out using the apparatus described in section 3.3.1.

I-V characteristics

The diode may be modelled by the following equation, representing an ideal diode and some higher order effects (generation/recombination in the SCR, series resistance).

$$I_d = I_s\left(\exp\frac{q(V - I_d R_s)}{nkT} - 1\right) \tag{5.1}$$

The ideality factor n has a value between 1 and 2 and represents the relative sizes of diffusion and recombination currents in the device. The parameters are extracted by plotting the log of the current vs. the voltage, in which case the slope of the forward current is proportional to $1/n$ assuming no series resistance, and undergoes a smooth roll-off at higher voltages (from which R_S may be calculated).

C-V characteristics

The C-V data provides two important pieces of information. First, it determines bandwidth based on RC delay (generally assuming a 50 Ω load for consistent measurement purposes). Second, it provides information regarding the doping and shape of the SCR by the relationship $1/C^2$ vs. V. The junction in the lateral structures of this work is between the n finger region and p- active region, therefore for an n+ – p- junction the appropriate expression is the following.

$$\frac{1}{C^2} = \frac{2}{\epsilon q A^2 N_{Ap}}(\phi_b - V_D) \tag{5.2}$$

Therefore for a one-dimensional uniformly doped junction, the slope of the line with depends on the junction area and doping level, as well as material constants. For a real junction, the doping level may not be constant and particularly in the case of lateral devices the SCR area changes with reverse bias, as will be discussed in the upcoming sections.

Bandwidth characteristics

As discussed in section 4.1.2, the lateral device designs of this work are dominated by carrier diffusion and device capacitance. The carrier drift time constant will appear at frequencies above those of interest. As discussed in section 2.3.3, the decay characteristics of lateral carrier diffu-

sion (approximately 10 dB/decade) are different than those of RC time constant (20 dB/decade). Therefore if these two bandwidths are spaced widely apart, the dominant mechanism is directly determined by observation of the decay slope. However, as shall be seen in the measured results, in many useful geometries the time constants are closely spaced (for example a factor of 2 in frequency apart), and it is difficult to determine their precise positions, particularly with the lack of phase information and in the presence of measurement noise or non-idealities. In this case it is convenient to define a curve fit that describes an empirical breakpoint with an exponent that defines a decay slope between 10 and 30 dB/decade (i.e. the decay for frequencies above the high break frequency).

$$|H| = \frac{1}{(1+(\omega/\omega_{RC})^2)^{\frac{1}{2}}}, \frac{1}{(1+(\omega/\omega_{diff})^2)^{\frac{1}{4}}}$$

$$\text{empirical form, in dB} \quad \longrightarrow \quad 10\log\left(\frac{1}{(1+(\omega/\omega_b)^2)^{\frac{c}{2}}}\right) \quad ; \quad 1 \leq c \leq 3 \quad (5.3)$$

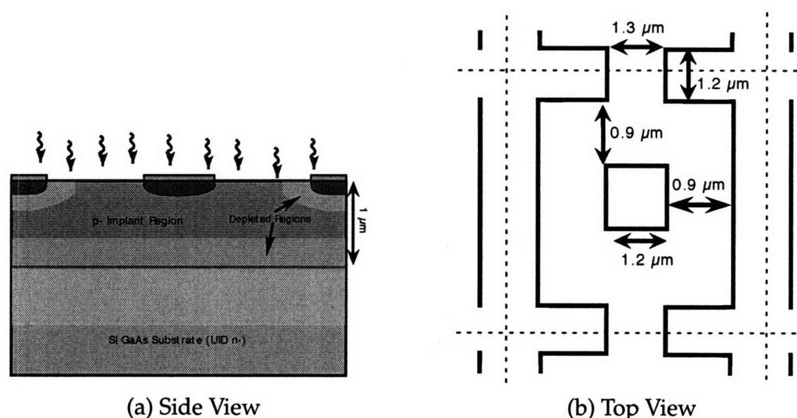
Using this form of empirical relationship, it is possible with one curve fit to determine a corner frequency and the dominant frequency decay mechanism. For example $c = 1$ indicates pure diffusion effect, $c = 2$ indicates pure RC effect, and $c = 3$ indicates coincident diffusion and RC time constants.

5.1 Lateral p-i-n devices in a standard commercial GaAs VLSI process

The detector of choice for many applications has traditionally been the metal-semiconductor-metal (MSM) photodetector because it is an intrinsically simple device that is planar, has good responsivity, and has low parasitic capacitance. It would seem that the MSM detector is an obvious candidate for integration with GaAs integrated circuits (ICs) using MESFETs. However, in practice it is necessary to make substantial additions to commercial integrated circuit processes to realize high performance MSM photodetectors. In particular, additional masks steps are required to form special (implanted) lightly-doped n regions for the MSM active regions. These changes are complicated and costly to implement, and there remains a great need for an integrated photodetector that can be fabricated without making changes to commercial GaAs MESFET IC processes.

The solution was a lateral p-i-n diode photodetector fabricated in the standard Vitesse Semiconductor Corporation H-GaAs-IV and H-GaAs-V processes.[90, 91, 92]

The lateral device structure used in the current work is shown in side view in Figure 5-1a and in overhead view in Figure 5-1b. The pattern is tiled to create a large area device with a width of 75 μm. This lateral structure can be fabricated in the standard Vitesse H-GaAs-IV and H-GaAs-V processes in which MESFETs are formed in lightly-doped p-type wells created by ion implantation into a nominally semi-insulating (physically lightly n-type) GaAs wafer. The H-GaAs-IV and V processes provide for isolation implants around the p-wells, and p+ ohmic contacts to these wells (for use as a MESFET body contact). The n-type region of the photodetector is made using the MESFET source/drain implant step. The main differences between the H-GaAs-IV and V processes were improved photolithography allowing for more compact design rules, as well as a change in the doping level and depth of the p- MESFET wells and source/drain regions.



(a) Side View (b) Top View

Figure 5-1: Schematic of the side and tops view of the lateral p-i-n device structure, including ion implant and ohmic contact structures but excluding interconnect metal. The width of the i region is defined photolithographically as the separation between the p and n implant regions.

In the device pictured in Figure 5-1, the p-well forms the i-region and the p+ and n implants are used to form the p and n-regions, respectively. The width of the i-region is determined photolithographically by the separation between the two later implants, 0.9 μm in the best devices. The n implant is patterned around the p+ implant in a mirrored "C" shape. The intent of this layout is to achieve a compromise between the proportions of device area covered by metal contacts, and the exposed area between the n and p+ regions.

5.1.1 DC Electrical Characteristics

The DC electrical characteristics were measured using a HP 4145B semiconductor parameter analyser, and are summarised in Table 5.1.1. The devices exhibit low dark current at bias points of interest and a high breakdown voltage. An ideality factor greater than 1 is expected in a p-i-n device.

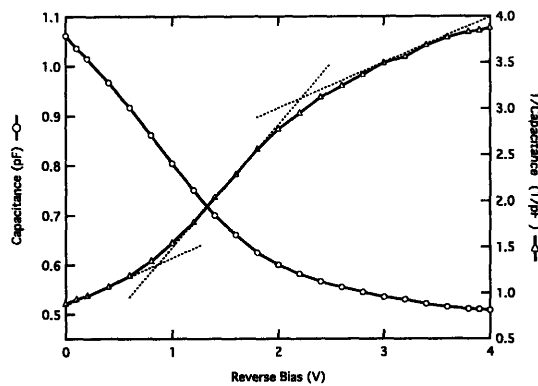
Dark Current (V_{Bias} = 10 V)	< 90 pA
Dark Current (V_{Bias} = 5 V)	< 30 pA
Breakdown voltage	> 30 V
Ideality Factor	1.9

Table 5.1: Electrical characteristics of the 0.9 μm lateral p-i-n devices.

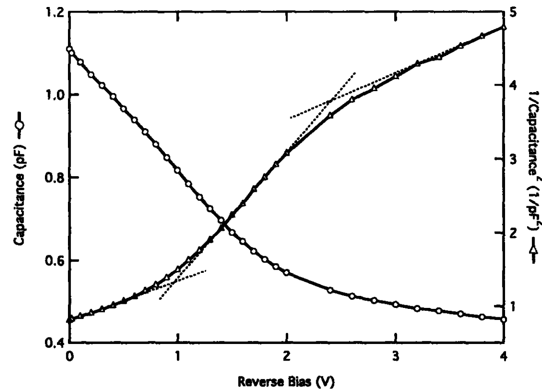
5.1.2 Capacitance

Figure 5-2 shows C-V and $1/C^2$ plots for different device i-region widths and two different IC processes, as noted below each subfigure. The data indicate that for all of the devices, assuming a 50 Ω load, bandwidth of over 1 GHz is possible assuming that they are RC-dominated. The dashed lines in the figure indicate regions of linear inverse capacitance behaviour, and the intersection of these lines indicate reverse biases at which the capacitance behaviour of the devices undergoes a significant change.

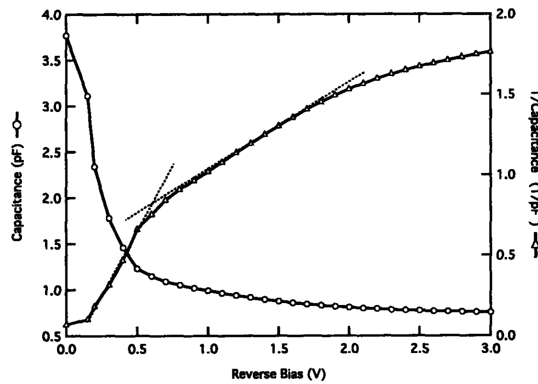
An interesting feature of these curves is that the intersection of the extrapolations of those linear regions occurs at approximately fixed points in either process, but is independent of the device layout. Therefore, it is determined primarily by the two-dimensional behaviour of the depletion region between the n and p implant regions (i.e. as in the schematic side view of the structure). For the H-GaAs-IV process, the two points of interest are approximately 0.8 V and 2.2 V. Simulations indicate that the former is the reverse bias at which the depletion region between the n finger and p- well and the depletion region between the p- well and n- substrate come into contact. The latter bias is the level at which the region under the n finger becomes fully depleted. For the H-GaAs-V process, the point of interest is approximately 0.6 V. Simulations indicated that at this level the region under the n-type implant becomes fully depleted, therefore the changes in slope between the two processes are qualitatively consistent.



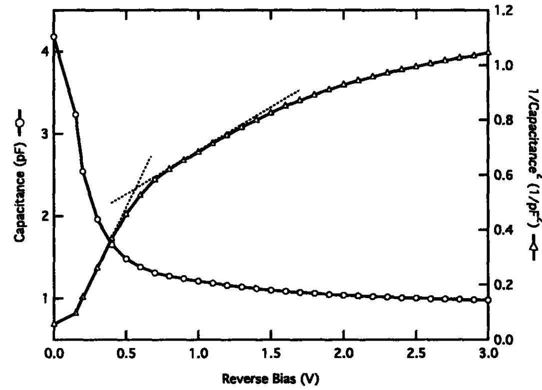
(a) H-GaAs-IV process, 2.5 μm i region



(b) H-GaAs-IV process, 2 μm i region



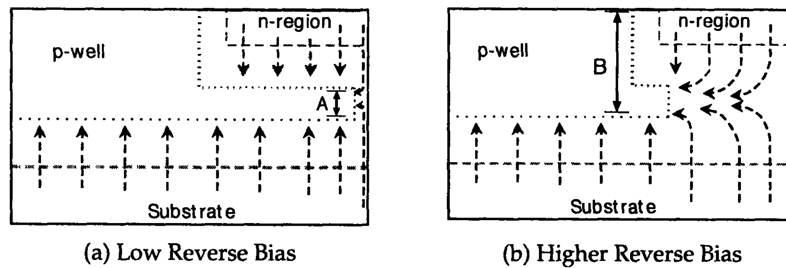
(c) H-GaAs-V process, 2 μm i region



(d) H-GaAs-V process, 0.9 μm i region

Figure 5-2: Capacitance (circles) and Capacitance^{-2} (triangles) as a function of reverse bias for devices of similar layout, but fabricated in two different GaAs VLSI processes and with different i-region widths.

A schematic explanation of this behaviour is shown in Figure 5-3. The electric field at the junction of the n and p-well regions has the opposite direction to the electric field at the junction of the p-well and substrate. Therefore when the reverse bias is such that the electric fields begin to interact (this may occur at zero reverse bias for appropriate doping levels and depths), the fields must bend in the absence of fixed charge. As reverse bias is increased, the region under the n-implant is depleted laterally outwards as more electric field lines interact and bend (this effective cross-sectional area is shown schematically as A). Once this region has been depleted, the depletion region continues to expand laterally with increased reverse bias, but now with a larger effective area (shown schematically as B). The C^{-2} vs. V behaviour has an area^{-2} dependence, hence the change in slope as seen in Figure 5-2.



(a) Low Reverse Bias (b) Higher Reverse Bias

Figure 5-3: Schematic diagram of simulated depletion region behaviour at low (a) and higher (b) reverse bias. The dashed arrows indicate electric field direction (not magnitude). The dotted line represents the depletion region edge.

5.1.3 Sensitivity

The photocurrent was measured as a function of incident light power at a constant reverse bias. A Thorlabs V3-780-TO-DA VCSEL with emission at nominally 780 nm was used as the light source, free-space coupled to detectors under measurement. The detectors exhibit a relatively linear intensity vs. photocurrent behaviour (from 0.5 μW to 1 mW of incident light power), with a small decrease in efficiency as incident power reaches the several mW range.

The photocurrent was also measured as a function of reverse bias at varying light intensities, as shown in Figure 5-4 for a 0.9 μm device. At low reverse bias, the photocurrent rises substantially with increasing bias. At higher bias, the photocurrent vs. reverse bias curve flattens significantly. At low light levels, the initial rise is steepest and the subsequent plateau is flattest. Extrapolating the linear regions of the curve at lower and high reverse bias results in an intersection at approximately 0.6 V, the same point observed in the C-V data. Sensitivity reaches a value of 0.17 A/W for a reverse bias of 5 V. Taking into account the metal fill-factor of each detector cell, the quantum efficiency of the devices is over 60% at this operating point.

The minority carrier diffusion length of electrons in GaAs is of the same order as the p-n spacing in the lateral p-i-n devices. Therefore, the same mechanism of rapid depletion of the area under the n-implant as identified regarding the capacitance data is believed responsible for the two regimes of sensitivity vs. reverse bias behaviour. As the area under the implant is depleted, more carriers are able to diffuse to the depletion region edge before recombining.

Devices with 2 μm i-regions were fabricated on the same chip. Their sensitivity also rises quickly at low reverse bias, then slowly at higher bias, with the change in behaviour also occurring

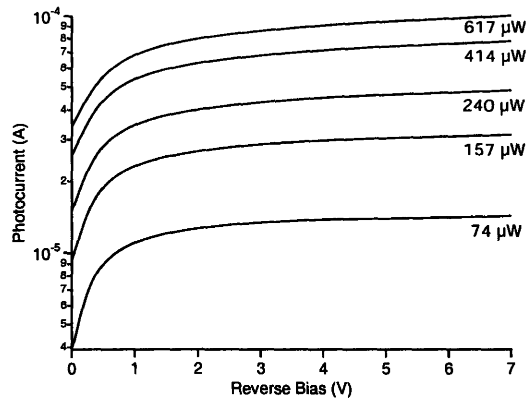


Figure 5-4: Photocurrent as a function of reverse bias and incident light power for a 0.9 μm device fabricated in the H-GaAs-V process.

at approximately 0.6 V, likewise as in their capacitance data. This confirms that the change in behaviour is due to the doping levels and depths of the n and p-well regions, not the physical layout of the device.

5.1.4 AC Characteristics

The GaAs devices were measured using a precursor of the apparatus described in section 3.3.1. A Thorlabs V3-780-TO-DA VCSEL and detector were biased using HP 6633A voltage/current sources through bias Ts with a specified bandwidth of 10 GHz. The VCSEL was modulated by an HP 8341B signal source in the frequency range 100 MHz to 5 GHz, free-space coupled to the DUT, and the small signal photocurrent was measured using a Minicircuits ZJL-6G amplifier and a HP 8510B network analyser.

The results of these measurements are shown in Figure 5-5 for a 0.9 μm devices, and for reverse biases varying from 0 V to 7 V. The data was fitted as discussed at the beginning of this chapter, and the fit line is represented by the thick lines in the figure. The 3dB bandwidth exceeds 2 GHz at 5 V reverse bias, and is 4.5 GHz at 7 V reverse bias which is the RC limit of the device and 50 Ω load. Because the capacitance is proportional to device area, smaller devices will have correspondingly larger bandwidths.

The fitting factor c increases with reverse bias, from 1.2 at 0 V to 2.7 at 7 V, indicating that the device is moving from diffusion-limited to RC-limited performance. The 2 μm detectors fabricated

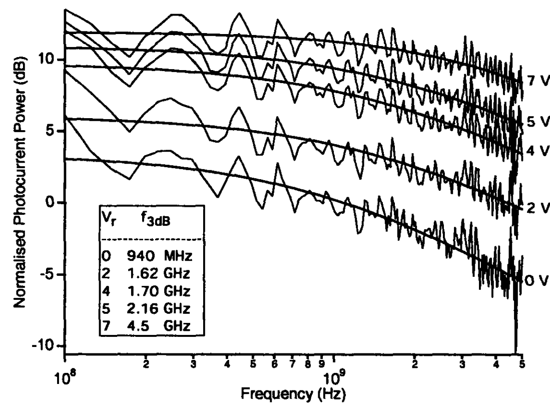


Figure 5-5: AC photocurrent response as a function of frequency at different reverse bias levels. The fitting parameter "c" increases with reverse bias (going from a value of 1.2 at 0 V to 2.7 at 7V)

on the same die showed similar response, but with lower f_{3dB} and lower c for each given reverse bias. This indicates that the response is more strongly dominated by carrier diffusion, as expected for a wider device.

5.1.5 GaAs Device Conclusions

High performance lateral p-i-n detectors were fabricated in an unmodified commercial GaAs VLSI process, with best performance demonstrated in the Vitesse H-GaAs-V process. Both the DC photocurrent data and the CV data exhibit a change in behaviour at a reverse bias of approximately 0.6 V. The AC response performance suggest that the devices are suitable for applications such as polymer fibre communication systems or gigabit Ethernet over fibre. Compared to other GaAs-based photodetectors, the current work demonstrates speed characteristics similar to those demonstrated for MSM diodes, but with a simpler fabrication process. For example, Vitesse Semiconductor produces a line of integrated receivers (VSC7807, VSC7809, VSC7810) featuring 70 or 100 μm diameter MSM detectors, based on process modifications licensed from IBM. The lateral device structure was considered as a replacement for the MSM detectors, but ultimately development on GaAs products was halted entirely due to unfavourable market conditions and the significant cost advantage of Si-based processes.

For a process/device design guideline, it is advantageous for both sensitivity and speed to operate the device in the regime where the region under the n implants is fully depleted. Optimally,

the device would be in this regime starting at zero reverse bias. This requires engineering the depth and doping of the p-well and n implant regions. In particular, the p-well should be doped at a low level to create a wide depletion region at all p-n junctions, but its depth should be at least on the order of an absorption length of the wavelength of interest to ensure good sensitivity.

5.2 CMOS-manufacturable Si lateral p-i-n photodetectors

A total of five Si device lots were processed, the details of which were presented in section 4.2. The first two lots were used to derive optimal processing conditions, lots 3 and 4 were used to produce "good" devices with different active region doping conditions, and lot 5 was a repeated version of lot 4 with further processing refinements primarily intended to reduce reverse leakage current. Short results of lots 1 and 2 will be presented to demonstrate processing progress, followed by detailed results of lots 3 and 4 demonstrating interesting and previously unreported sensitivity and bandwidth behaviours.

5.2.1 DC Electrical Characteristics

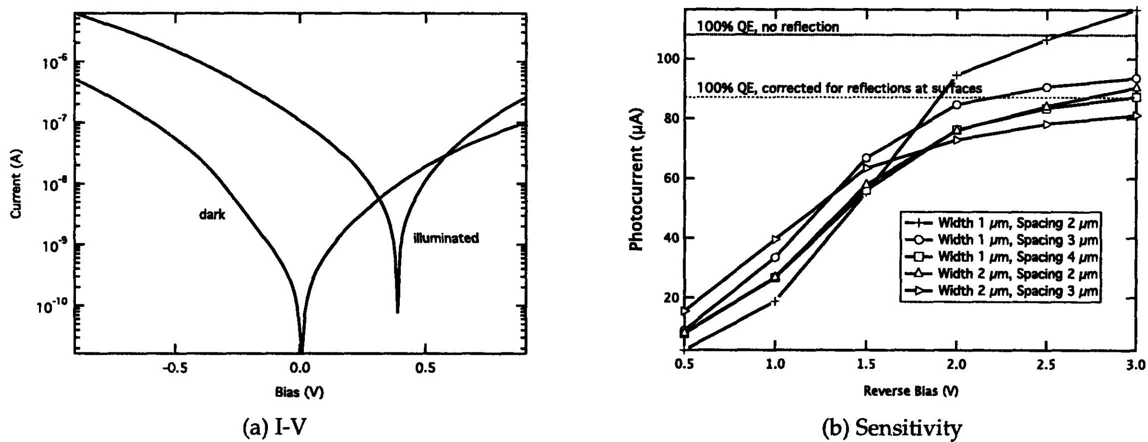
Si Lot 1

This lot was the first attempt at a full process, resulting in very poor devices but substantial processing insight. The I-V characteristics of a representative device and the sensitivity characteristics of several device geometries are shown in Figure 5-6.

The I-V relationship is very poor with a large reverse current and large series resistance. In fact, it is only the shift of the curve under illumination that definitively confirms which applied bias direction is positive. The sensitivity curve is promising since at some applied reverse bias, approximately an expected amount of photocurrent is collected; 100% internal efficiency is expected for the bulk p-type devices.

The following conclusions and/or process changes were reached.

- BOE etching should not be used to define the p and n finger patterns due to larger lateral etch and poor control of etch rate.



(a) I-V

(b) Sensitivity

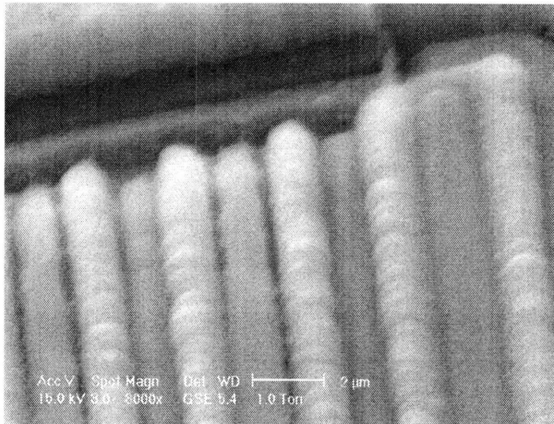
Figure 5-6: Si lot 1 DC device characteristics.

- The thermal passivation oxide thickness should be reduced from 750 nm to 250 nm. The large thickness results in significant topography and exacerbates the lateral etching problem.

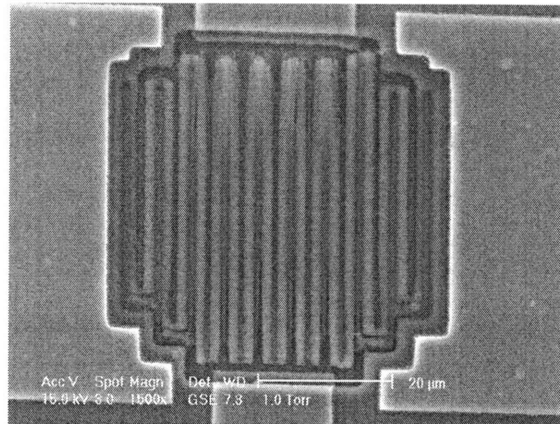
- The previous two points resulted in an unshielded silicon surface that was etched and roughened during the dry etching of the metal electrodes (clearly visible in Figure 5-7a). This created a large defective area, resulting in generation and recombination sites that affect the diode I-V characteristic, particularly increasing the reverse current. These areas are also expected to have a high carrier surface velocity, resulting in significant recombination under low bias, hence the behaviour of sensitivity at different bias.

- Dry etching was not available for the two-level metal process. The wet BOE etch was used as discussed in section 4.2.2 with poor results. Subsequently a thin Ti cap layer was deposited on the ohmic metal to prevent overetching, with slightly improved results. Dry etching was finally implement in Si lot 3.

- The contact pad metal acted as part of the ohmic metallisation after the desired metal was removed by the BOE via etch, but a large contact resistance remained likely due to residues remaining after the via etch and possibly insufficient contact sintering.



(a) Oxide openings defined (alternating) by wet and dry etch, followed by metal deposition and etch



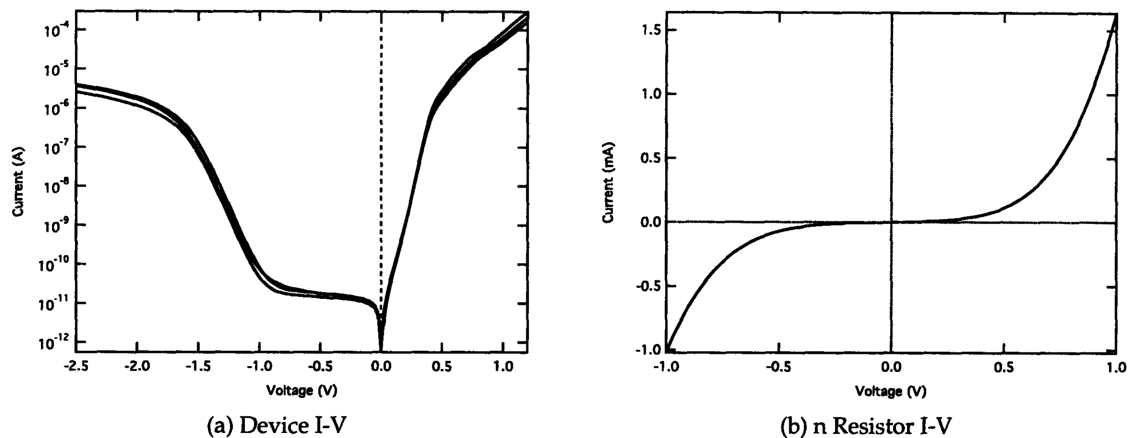
(b) Large lateral overetch of the via opening and ohmic contact metal

Figure 5-7: SEM pictures of lot 1 devices.

Si Lot 2

The devices of lot 2 were based on the changes to the process of lot 1, with the addition of a set of devices on n-type wafers with a p-type implanted "active" region. For the sections to follow, devices on p-bulk wafers will be called "control devices" while the devices with implanted active region will be called "isolated devices".

The I-V relationship of the lot 2 diodes was much improved as compared with lot 1, as shown in Figure 5-8a for an isolated device. The ideality factor is 1.4, in line with expectations for a lightly doped p side junction. There is a roll-off at higher currents that has a primarily resistive shape, with a calculated series resistance of approximately several kΩ. The reverse current between 0 and 1 V is excellent, then increases sharply. The increase appears similar to a diode behaviour with large series resistance, rather than a breakdown mechanism. This suggests a small Schottky junction is parallel with the desired p-n junction, likely an artefact of processing conditions, implant masking, or photolithography misalignment. The large and changing series resistance is due to a Schottky contact to the n-type region as shown in Figure 5-8b (implanted n-type resistor), a symptom of insufficient phosphorus implantation.



(a) Device I-V

(b) n Resistor I-V

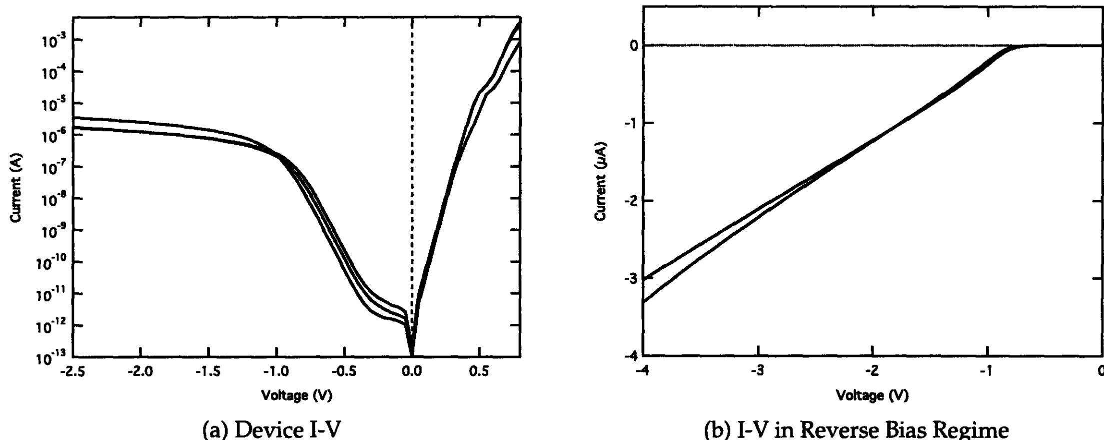
Figure 5-8: I-V properties of isolated devices. The three lines are the measurement of three devices of different dimensions (200 μm diameter, electrode spacings of 2-4 μm). The n resistor is a long narrow implantation region used to diagnose doping levels in the process.

Si Lot 3

The I-V relationship of the lot 3 diodes is shown in Figure 5-9. The three process changes as compared with lot 2 were a higher n region doping (dose raised by a factor of 10), slightly lower dopant anneal temperature (decreased by 50 degrees to 950°C), and the use of plasma etching to define the via between the ohmic contact and external pad metal. The first and third change contribute to a much lower series resistance as seen in the forward current region. The ideality factor of the forward current is 1.2. As seen in Figure 5-9b, the reverse current exhibits a behaviour resembling a diode in series with a large resistance on the other of 300 kΩ. The ideality factor of the reverse current rise between -0.5 and -1 V bias is 2, and saturation current is 10^{-3} times the forward saturation current. Analysis of the lot 2 devices showed a similar ideality factor in reverse bias, but with a lower saturation current. An equivalent circuit for this behaviour and discussion of process details creating this situation will be presented following the Si lot 4 results.

Si Lot 4

The I-V relationship of the lot 4 diodes is shown in Figure 5-10. The lot 4 process featured a new mask set designed for Ge-on-Si devices, but that could also be used with silicon devices and offered better alignment margin in the ohmic-metal etch pattern, as well as smaller contact pads for more reliable RF probing and smaller parasitic resistance between the signal pad and



(a) Device I-V

(b) I-V in Reverse Bias Regime

Figure 5-9: I-V properties of lot 3 isolated devices. The three lines are the measurement of three devices of different dimensions (200 μm and 50 μm diameter, electrode spacings of 2-4 μm).

substrate. Also, the lot consisted only of isolated devices with a lower active region doping and slightly larger active region thickness than the previous lots. Qualitatively the characteristics are similar to those seen in Figure 5-9. Series resistance is further reduced due to improvements in the via etch process. Forward ideality factor is 1.25, slightly higher than in lot 3 as expected for a lighter active region doping. Behaviour in reverse bias is the same as lot 3, with an ideality factor of 2 and a higher series resistance of approximately 1.5 MΩ.

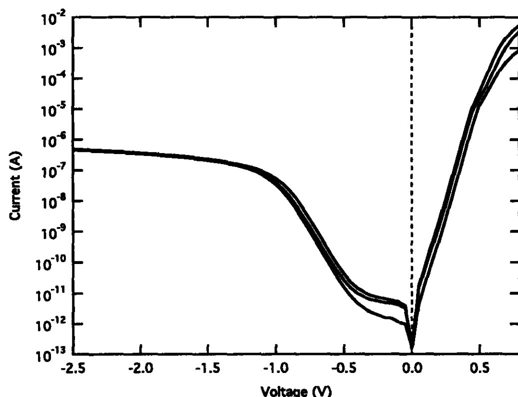
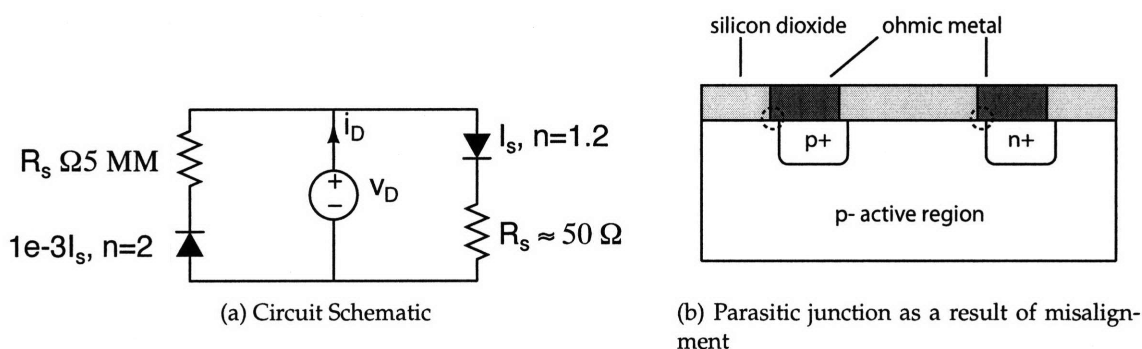


Figure 5-10: I-V properties of lot 4 isolated devices. The three lines are the measurement of three devices of different dimensions (200 μm and 100 μm diameter, electrode spacings of 2-4 μm).

Discussion of Lots 2-4 Reverse Current Characteristics

The characteristics of Figures 5-8, 5-9, and 5-10 may be represented by an equivalent circuit as shown in Figure 5-11. On the right, there is the desired junction with its series resistance, and on the left there is an undesired junction with a much smaller saturation current and much larger series resistance. It is believed that the undesired junction arises as a result of unexpected processing conditions. Specifically, the undesired junction is most likely a Schottky junction between the ohmic metal finger due to misalignment of p and n implants relative to the passivation oxide openings and ohmic metals pattern, as shown in Figure 5-11b.



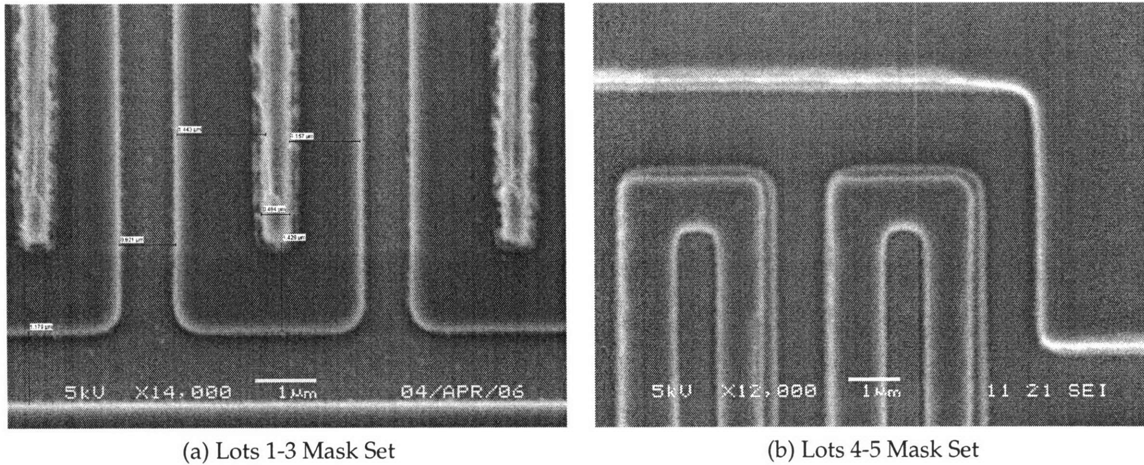
(a) Circuit Schematic

(b) Parasitic junction as a result of misalignment

Figure 5-11: Schematic diagrams representing the I-V characteristics as measured in lots 2-4, and the fabrication conditions that could lead to such a condition.

In lot 2, the saturation current of the parasitic junction is much lower than in the subsequent lots. This may be explained by two factors. First, the lower implant annealing temperature of the later lots results in less lateral dopant diffusion. Second, the phosphorus implant of lot 2 was carried out at 7° tilt, but with automatic rotation of the wafers. It was discovered after the fact that the implantations of lots 3 and 4 were carried out in a different apparatus with no rotation. Since a 1 μm photoresist mask was used, the shadowing effect left an approximately 100 nm region at one edge of the pattern with no implantation. These two conditions together contribute to exacerbate the misalignment problem. However, according to process simulation using TSUPREM4, these effects alone are insufficient to alone cause the misalignment. The additional factor is the systematic misalignment/offset problem described in section 4.2.2. Examples of test patterns exhibiting the systematic misalignment are shown in Figure 5-12. The error may be as large as 0.25-0.35 μm in the lateral direction depending on the reticle quadrants in question.

It should be noted that the magnitude of the reverse current is similar to that of otherwise



(a) Lots 1-3 Mask Set

(b) Lots 4-5 Mask Set

Figure 5-12: SEM pictures of misalignment of test patterns in the first and second mask sets used to fabricate the Si lateral p-i-n devices.

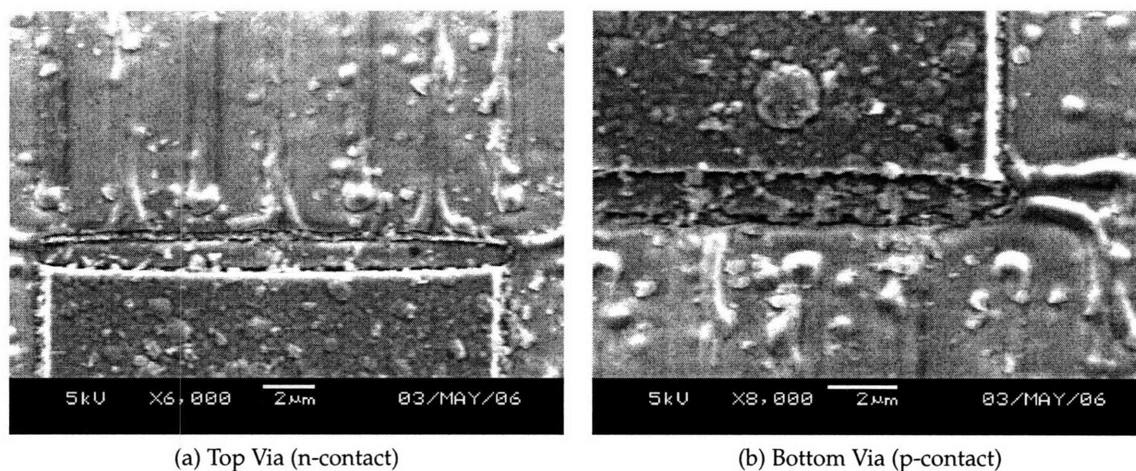
well-performing devices reported in the literature (for example [18]) where processing errors or problems with substrate quality were identified.

Si Lot 5

The Si lot 5 devices were affected by a processing problem during the metallisation process. After fabrication, it was found that a bias of 3-5 V (positive or negative) was required for any current to flow. This reason for this was the presence of some sort of dielectric that broke down at a particular field strength, most likely between the ohmic and pad metal layers, and perhaps also at the ohmic contact interface. This behaviour was confirmed with two other fabrication facility users using multiple metal layers [93]. One user, fabricating multi-level inductors, found that 20 V was needed for any current to flow in this passive device. The fabrication processes were all different in terms of deposited metal and etching, but shared the common characteristic of PECVD silicon dioxide and metal deposition by sputtering. The contamination was therefore identified as originating in one of these two instruments, though the exact cause is not fully understood. The material contributed to an abnormally high series resistance of the devices.

The reverse current was slightly lower than that of the Si lot 4 devices. However, it was still higher than expected (0.6μA at 3 V reverse bias), had the same diode-like characteristic, and was interestingly independent of device geometry. The only layout feature that is common among the device geometries is the inter-metal via. Upon inspection under SEM, it was found that the via

was etched much wider than defined by the mask pattern, as seen in Figure 5-13. The nominal via width is 1 μm, and the overlaying metal pad pattern is drawn to extend over and past the via. However, in the images the via appears to be somewhat oval with a width approaching 2 μm at the widest point. The pad metal is fully within the via hole. Depending on the rate of the via etch, this may result in a parasitic Schottky diode as described in the previous section. A solution to this problem is to increase the width of the bus metallisation and perhaps implant region in order to provide a larger pattern tolerance, at the price of a slightly higher parasitic capacitance.



(a) Top Via (n-contact)

(b) Bottom Via (p-contact)

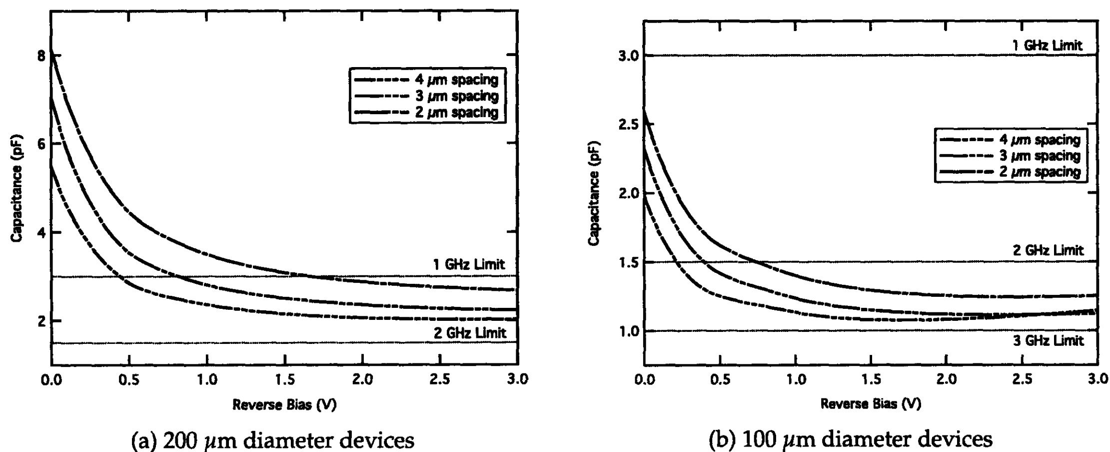
Figure 5-13: SEM pictures of n and p contact vias with pad metal. The roughness and dust-like particles are related to the metal deposition and anneal which left some residue on the field oxide surface.

5.2.2 Capacitance

Due to the high reverse currents present in the Si lots 1-4, measurement of the capacitance was difficult in certain cases. As the reverse current rises, the measurement bridge is not able to distinguish between the current through the effective parallel resistance and the capacitor under measurement. The result is a change in direction of the C-V characteristic (i.e. a kink followed by rapid decline, or a minimum followed by a rise in capacitance with increasing bias). The best and most consistent measurement results came from lots 3 and 4.

Si Lot 3

The C-V relationship of the lot 3 isolated devices is shown in Figure 5-14. The low-bias capacitance value is somewhat higher than expected. The capacitance falls rapidly at low reverse bias, reaching levels where RC bandwidth, assuming a 50 Ω load, is above 1 GHz for 200 μm diameter devices and approaching 3 GHz for 100 μm diameter devices. These capacitance levels are indicated by dotted lines in the figures. The p- active region is a blanket region over the entire die, therefore there is a parasitic capacitance between the signal contact pad and the active region that is in parallel with the junction capacitance. The value of this parasitic capacitance is 0.45 pF.

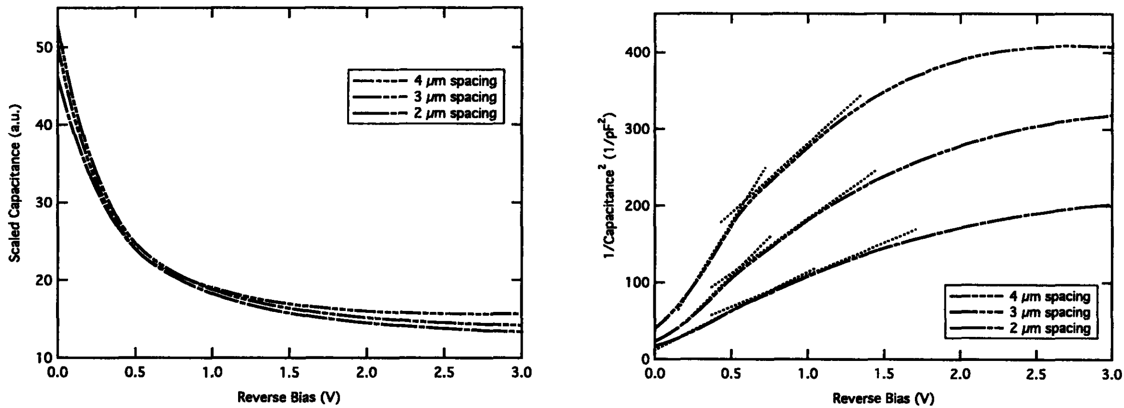


(a) 200 μm diameter devices

(b) 100 μm diameter devices

Figure 5-14: C-V properties of isolated lot 3 devices. The three lines indicate different electrodes spacings as shown in the legends.

The C-V behaviour is investigated in two additional ways. First, the contact pad parasitic capacitance is subtracted from the measured values, and the resulting data is scaled by the proportion of n+ finger area to the entire device area. This is shown for the 200 μm diameter devices in Figure 5-15a. The characteristics for the three devices are nearly coincident, indicating that the capacitance behaviour is, as expected, dominated by the area adjacent to the n+ regions. The capacitance-voltage relationship is also not affected by the electrode spacing. The error between the lines may be attributed to several factors: the diameters of the devices are only nominally equal (as the devices are designed based on small rectangular unit cells), the perimeter effects of the devices, and measurement error due to reverse current (particularly in the case of the 4 μm spacing devices that exhibit the lowest capacitance).



(a) Capacitance scaled to account for different n region fill factor, 200 μm diameter devices

(b) $1/C^2$ vs. reverse bias data, 200 μm diameter devices

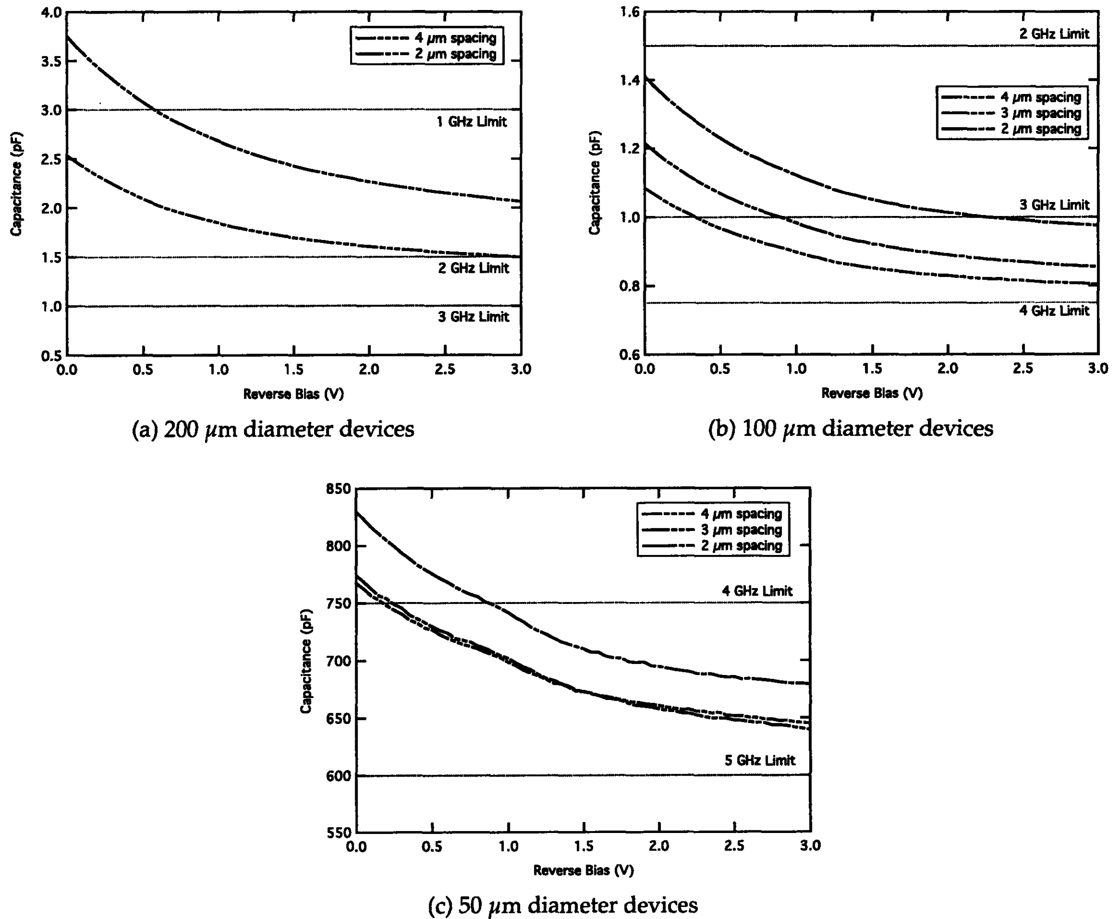
Figure 5-15: C-V properties of isolated lot 3 devices. The three lines indicate different electrodes spacings as shown in the legends.

The inverse capacitance squared as a function of reverse bias is shown in Figure 5-15b. Several behaviours are visible. First, the measurement error of the 4 μm spacing devices is apparent beyond 2.5 V as an abrupt flattening of the curve. Second, in all curves it is possible to identify approximately linear regions at lower biases. For all devices, the intersection of these regions occurs at 0.5 V. This is qualitatively the point at which the capacitance drop with bias slows significantly. A similar explanation maybe employed as in the case of section 5.1.2, where the change of slope is associated with a change of shape of the depletion region as it expands first under the n region with a small effective area, then laterally towards the p+ region with a larger effective area. The relative slopes of the three curves correspond to the different junction areas of the devices (smaller junction area results in a larger slope).

Si Lot 4

The C-V relationship of the lot 4 isolated devices is shown in Figure 5-16. The effect of the lower active region doping is clearly visible as the capacitances are significantly lower than in the case of the lot 3 devices. A positive side effect of this observation is that the width of QNR is also smaller than in the case of the lot 3 devices. The lower reverse current leakage also allows for a cleaner measurement. The measurement of the 50 μm devices were more strongly affected by the reverse current. The capacitance reaches levels where RC bandwidth, assuming a 50 Ω load, is up

to 2 GHz for 200 μm diameter devices, approaching 4 GHz for 100 μm diameter devices, and 5 GHz for 50 μm devices. The lot 4 devices were designed with smaller contact pads, therefore the value of the parasitic capacitance in parallel with the junction capacitance is 0.26 pF. This suggests that over 5 GHz bandwidth is possible for the 100 μm devices if the contact pads were eliminated, for example in an integrated application.



(a) 200 μm diameter devices

(b) 100 μm diameter devices

(c) 50 μm diameter devices

Figure 5-16: C-V properties of isolated lot 4 devices. The different lines indicate different electrodes spacings as shown in the legends.

The inverse capacitance squared as a function of reverse bias is shown in Figure 5-17. It is possible to identify two approximately linear regions. The intersection of these regions occurs at 1.5 V. The explanation is a change in effective junction area. Since the capacitance decreases similarly for two different electrode spacings, it is apparent that the entire QNR has not been depleted. The relative slopes of the curves correspond to the different junction areas of the devices (smaller junction area results in a larger slope).

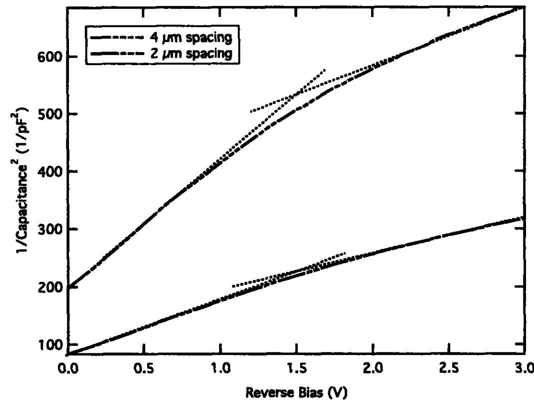


Figure 5-17: Lot 4, $1/C^2$ vs. reverse bias data, 200 μm diameter devices.

5.2.3 Sensitivity

Si Lot 2

Both control and isolated devices were measured under illumination to determine device sensitivity using a 780 nm VCSEL. The sensitivity vs. bias characteristic of the control devices is shown in Figure 5-18. The sensitivity is uniform over reverse bias, as expected (and contrasted to the characteristic of the lot 1 devices), and the internal sensitivity is close to the 100%. Differences between the electrode spacings may be attributed to non-idealities in electrode width definition, and additional reflection due to surface topology that is less uniform as the inter-electrode spacing decreases.

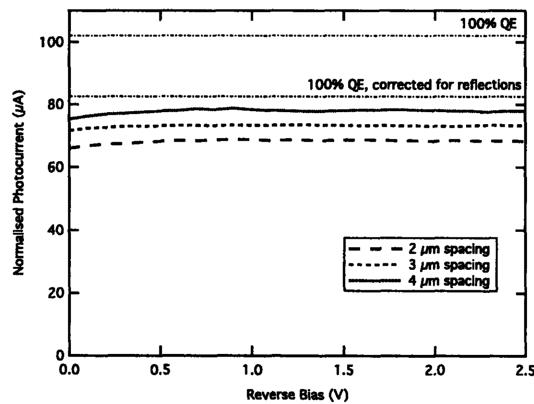


Figure 5-18: Photocurrent vs. reverse bias, normalised to account for electrode shadowing and surface reflections. The devices has 1 μm wide electrodes, with electrode spacing as shown in the figure.

In device simulations, it was consistently found that the photocurrent was much higher than expected and that consequently the active region to substrate isolation junction was not behaving as expected, i.e. it was not blocking diffusing carriers generated below the active region. This is contrary to the results of other reported photodetectors with similar structures. However, the behaviour was confirmed upon measuring the lot 2 isolated devices. Device sensitivity at 0 V reverse bias (expressed as quantum efficiency), is shown in Fig. 5-19. The sensitivity at low intensity is approximately as expected from the device structure and the absorption length of 780 nm light in Si (8 μm). At high intensity, the efficiency rises rapidly and substantially. The figure inlay contains the raw measurement data, showing linear sensitivity over nearly 3 orders of magnitude of light intensity.

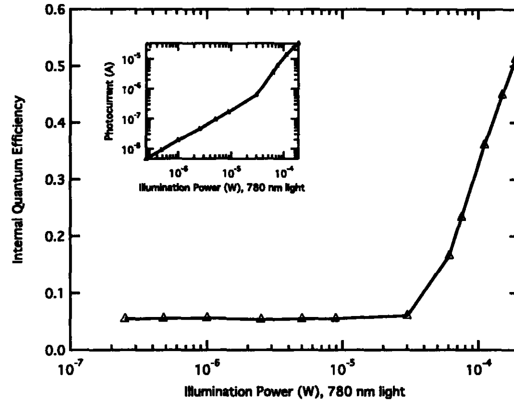


Figure 5-19: Sensitivity vs. illumination for the isolated Si lot 2 devices. The inset shows the raw photocurrent data, demonstrating linear sensitivity over three orders of magnitude followed by an abrupt change in behaviour. The spot diameter was approximately 70 μm.

Si Lot 3

The sensitivity of the lot 3 devices was measured using an 850 nm light source, with absorption length in Si of approximately 15 μm. The rise in sensitivity is not as abrupt as in the case of the 780 nm light source, as shown in Figure 5-20.

The phenomenon may be explained by considering the isolated detector as an npn phototransistor with a hanging collector or emitter. In practice, a fabricated phototransistor for commercial use may be quite similar to the lateral p-i-n detector structure with large electrode spacing, with the exception that the p active region (transistor base) is left as a hanging node. Therefore an ef-

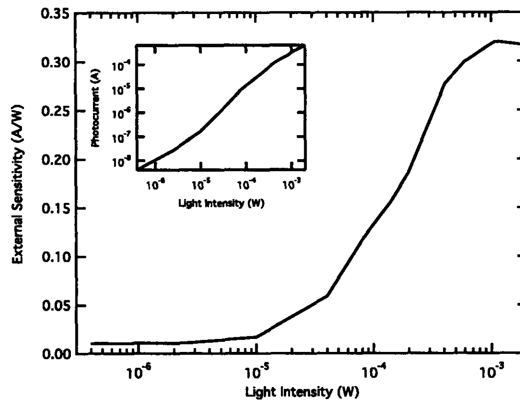


Figure 5-20: Sensitivity vs. illumination for the isolated Si lot 3 devices. The inset shows the raw photocurrent data, demonstrating linear sensitivity over two orders of magnitude followed by a strong change in behaviour. The spot diameter was approximately 70 μm.

fective current source is formed by the photodiode junction between the collector (substrate) and base, and the current is amplified by the current gain β of the transistor. In the case of the lateral p-i-n of this work, the base is connected electrically and the collector node is hanging. This may be represented by the Ebers-Moll model as shown in Figure 5-21

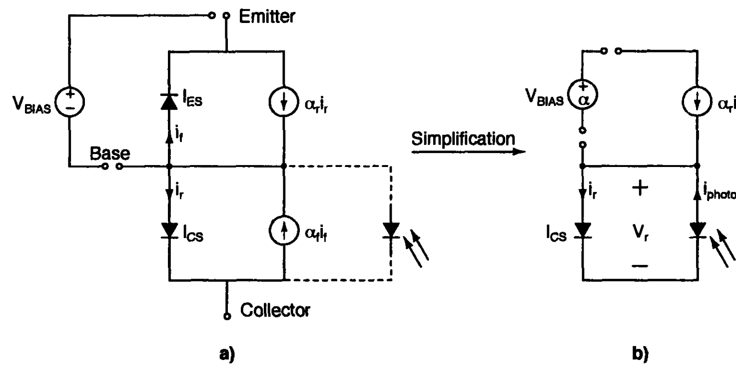


Figure 5-21: Ebers-Moll model representation of the lateral p-i-n device as a phototransistor. The model may be simplified as shown assuming that the detector is in reverse bias.

As the illumination intensity increases, the operating point of the pair of diodes changes. Physically, the diffusion of carriers into the hanging substrate results in a change of the substrate-active region bias that reinjects the carriers back in the active region to maintain charge neutrality. However, because the area under the electrodes that is not illuminated is different than the inter-finger area, the I_S of the two diodes in the figure is not equal. The operating point could be determined by a load line analysis as shown schematically in Figure 5-22. In the model, the operating point

current generates an emitter-base current with a factor α. Physically, the npn junction under the shadowed areas becomes forward biased and results in electron injection (with a corresponding hole injection across the junction to maintain charge neutrality). This is only a first-order explanation, as it ignores several conditions for example the fact that the npn junction under the electrodes is biased in punch-through (i.e. base fully depleted).

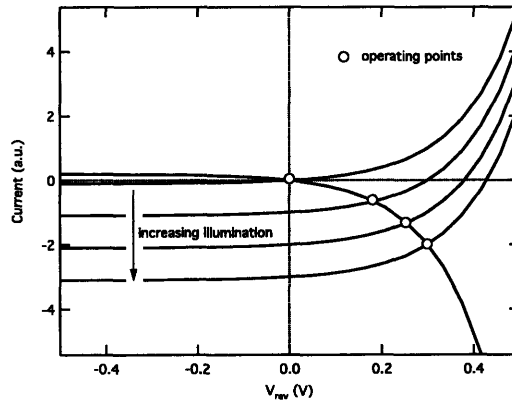


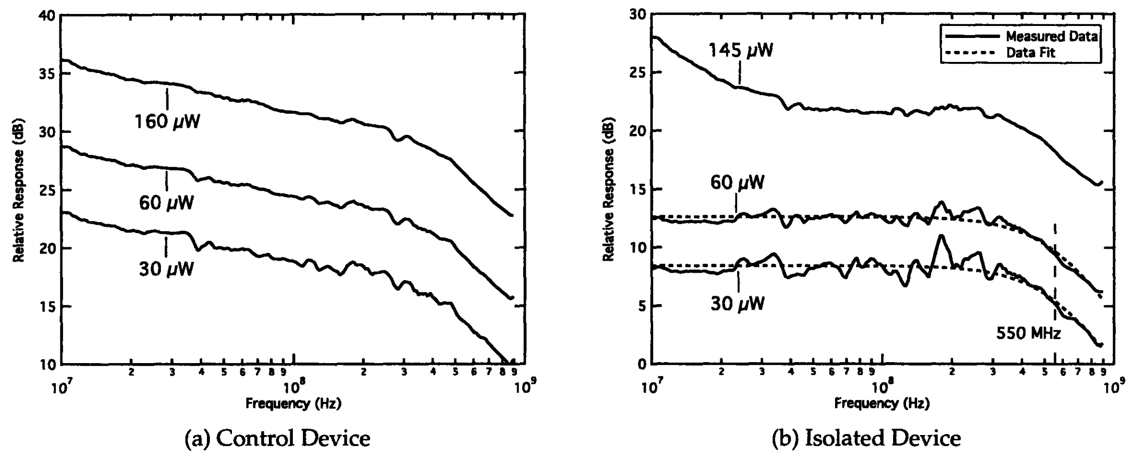
Figure 5-22: Schematic diagram of the load line operating point of the illuminated isolated devices.

5.2.4 AC Characteristics

Si Lot 2

Device bandwidth was measured at 1V reverse bias, and at three different intensities corresponding to three different sensitivities as seen in Figure 5-19. The results of the control devices are shown in Fig. 5-23a. The measurement is similar to previously reported results of lateral device structures [14]. There is a slow decay from the starting frequency and a faster decay beginning near 500 MHz. All measured intensities exhibit the same behaviour. The results of the devices with the buried junction are shown in Fig. 5-23b. The low-intensity measurements show a flat characteristic with a bandwidth of 550 MHz. The high-intensity measurement is similar to the other curves above 50 MHz, but exhibits a low-frequency decay reminiscent of the control devices.

This is an important and previously unreported effect that indicates the presence of two independent photocurrent sources (fast lateral, slow vertical), the latter of which is intensity-dependent. Both sources are limited by the device RC time constant, and the slow substrate current strongly degrades device bandwidth.



(a) Control Device (b) Isolated Device

Figure 5-23: Frequency response of the control and isolated devices at different average illumination powers as indicated in the figures (70 μm spot size).

A further interesting observation from the control devices is the frequency decay of the photocurrent, dominated in that case by vertical carrier diffusion from the substrate. The response falls at approximately 5 dB/decade, different than the RC case of 20 dB/decade or the lateral diffusion case of 10 dB/decade. The behaviour is consistent, however, with the solution of the diffusion behaviour presented in section 2.3.3 for the case of non-uniform generation, as suggested in Figure 2-9c.

Si Lot 3

Bandwidth measurement of the control devices (200 μm diameter, 2 μm electrode spacing) is shown in Figure 5-24a. Similarly as in the Si lot 2 devices, the photocurrent decays at approximately 5 dB/dec at lower frequencies, and more quickly following a certain corner frequency. The corner frequency is independent of average light intensity, occurring at approximately 650 MHz for 1 V reverse bias. More generally, the entire frequency response is independent of illumination. The corner frequency is somewhat lower than expected for a purely RC effect based on the measurements of Figure 5-14, therefore it is reasonable to conclude that the approximately 20 dB/decade decay is a combination of vertical (5 dB/decade) and lateral (10 dB/decade) diffusion time constant (combining additively), and the RC pole at slightly higher frequency. This decay point is at a higher frequency than in the corresponding Si lot 2 device, due to improvement in the ohmic contact to the n+ fingers.

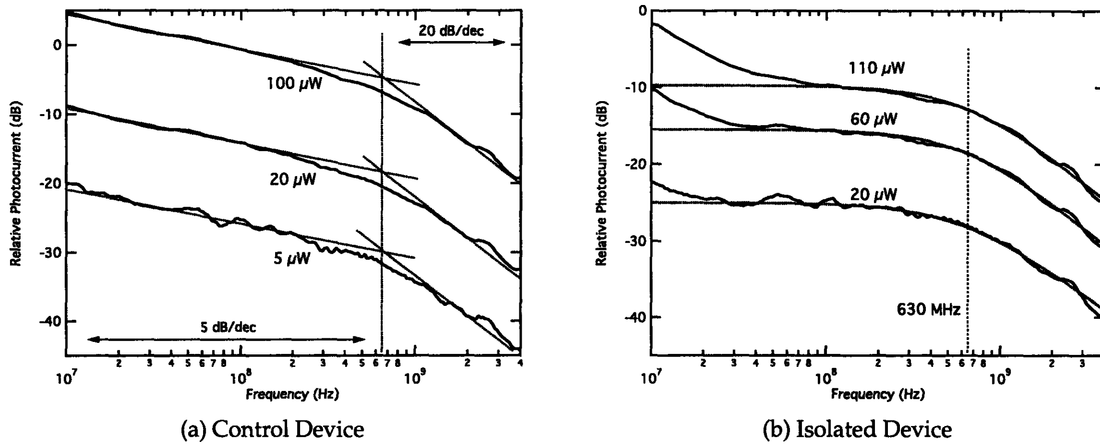
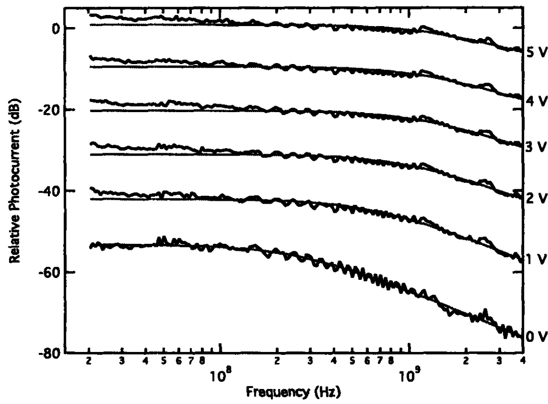


Figure 5-24: Frequency response of the control and isolated devices at different average illumination powers as indicated in the figures (70 μm spot size). The devices geometry was 200 μm diameter, 2 μm electrode spacing.

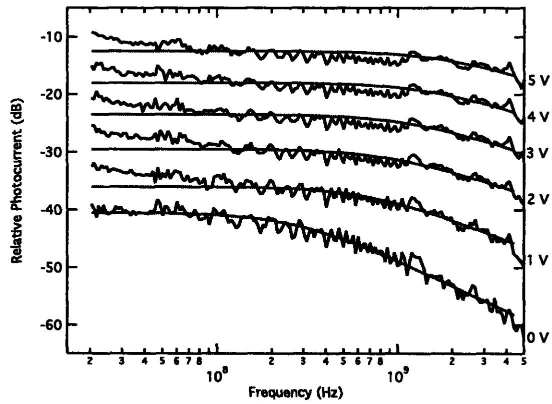
Isolated devices with the same geometry and reverse bias show a similar response as observed in Si lot 2. The fitted corner frequency is 630 MHz, approximately the same as observed in the simple two-line fit of the control devices. The corner frequency is consistent over different illuminations, and the substrate current effect increases with increasing illumination. Additionally, the effect is stronger in the lot 3 measurement than in the lot 2 measurement due to the VCSEL source (850 nm vs. 780 nm respectively) that results in greater carrier generation below the isolation junction.

Bandwidth measurement of the isolated devices was performed for a large variety of device diameters and electrode spacing as fabricated. Yield of the smallest devices (50 μm diameter, 1 and 2 μm electrode spacing) was poor and the frequency responses were inconsistent, therefore the smallest devices are not considered in the analysis of the lot 3 devices. Typical frequency response characteristics of the isolated devices are shown in Figure 5-25.

The measured data was fit using the function presented in equation 5.3. The value of the fitting coefficient c in this formulation may vary between 1 and 3, indicating the dominant time constant (diffusion, RC, or coincident time constants). It should be noted that the corner frequency of the fitting function is not necessarily the 3 dB bandwidth frequency - this condition is only met for $c = 2$. The 3 dB bandwidth frequencies and values of the fitting coefficient for 100 and 200 μm diameter devices are shown in Figure 5-26.

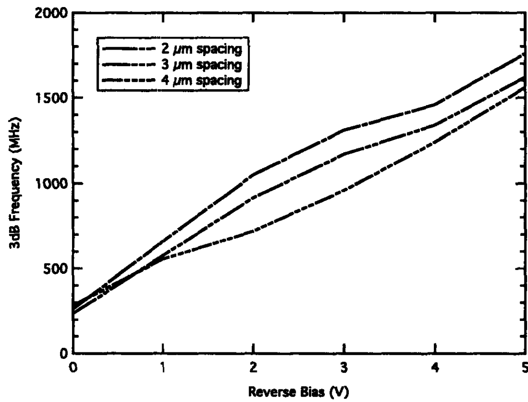


(a) 200 μm diameter, 2 μm electrode spacing

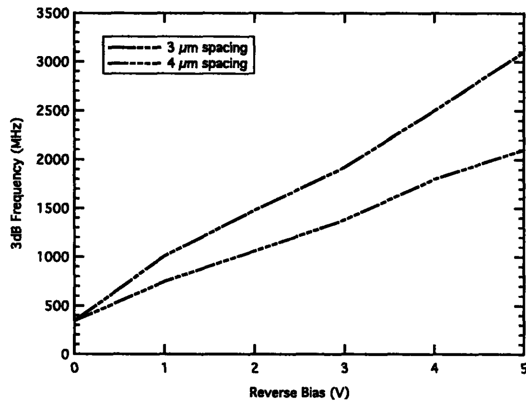


(b) 100 μm diameter, 3 μm electrode spacing

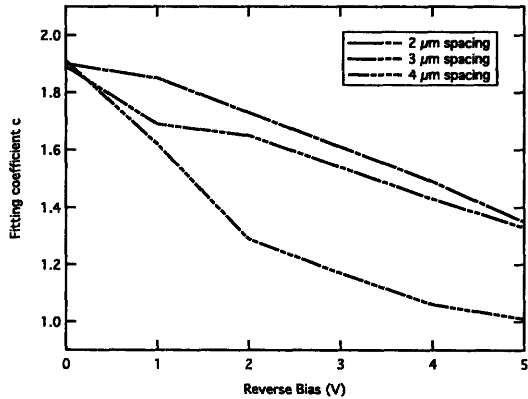
Figure 5-25: Frequency response of isolated lot 3 devices at different reverse biases and low illumination (i.e. as in lowest curves of Figure 5-24). The measured data is shown in the dark lines, and the curve fit in the grey overlaid lines.



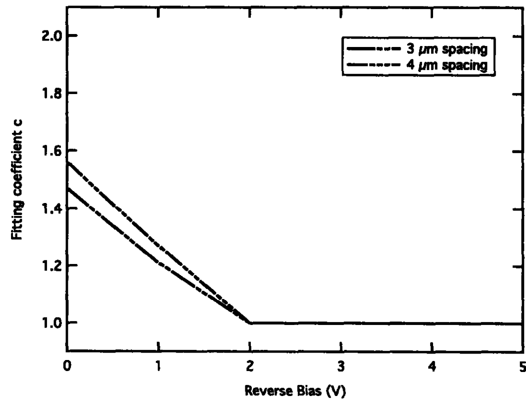
(a) 200 μm diameter, 3 dB bandwidth



(b) 100 μm diameter, 3 dB bandwidth



(c) 200 μm diameter, fitting coefficient c



(d) 100 μm diameter, fitting coefficient c

Figure 5-26: Extracted 3 dB frequencies and fitting coefficients of lot 3 devices with geometry as noted below and within each diagram.

The results indicate several important points. For 200 μm diameter devices, the bandwidth is dominated by the RC time constant at low reverse bias, increasingly moving towards a diffusion-dominated behaviour at higher reverse bias. The devices with larger electrode spacing are more strongly dominated by the diffusion effect, which is particularly visible in the fitting coefficient of the 4 μm spacing devices. The bandwidth at 5 V reverse bias indicates that all three presented device geometries are suitable for systems above 2 Gbps. Unexpectedly, the devices have similar bandwidths despite the very different geometries. Therefore, as long as the system bandwidth specification is met, it is desirable to increase electrode spacing in order to reduce shadowing (thereby increasing sensitivity and sensitivity-bandwidth product).

The 100 μm diameter devices have a much smaller capacitance than the 200 μm diameter devices, as seen in Figure 5-14. Therefore, it is expected that these devices would be diffusion-dominated, and this is the case as may be seen in Figure 5-26b. In this case, as the 3 and 4 μm spacing devices are both operating in the diffusion-limited regime, the difference in bandwidth is strongly apparent. The bandwidth of the 200 and 100 μm diameter devices with 4 μm spacing is similar, consistent with the fact that both devices are diffusion-limited and this process is independent of device area.

Bandwidth, fitting coefficients, and expected RC time constants based on direct capacitance measurements are presented in Table 5.2. The possible errors in the 100 μm device capacitance measurements were discussed earlier. Comparing the measured and calculated RC bandwidths, it is apparent that the devices with larger electrode spacing are certainly diffusion-limited. The discrepancy in the 200-2 device at 3 V reverse bias is not understood. A full discussion of the device behaviour will follow in chapter 6.

Si Lot 4

The lot 4 devices feature a lighter active region doping, and thus for an identical geometry as in the lot 3 devices, a higher bandwidth was expected. The capacitance and lateral extent of the depletion region should both improve with lighter doping (since the capacitance is in fact a function of the extent of the depletion region). Capacitance was indeed lower, as shown in Figure 5-16. Control devices, similarly as in lot 3, exhibited a frequency response dominated

Device	Reverse bias (V)	f_{3dB} (MHz)	c	f_{RC} (MHz)
200-2 μm	0	263	1.9	392
	1	662	1.85	910
	2	1050	1.73	1110
	3	1310	1.61	1200
200-3 μm	0	236	1.89	452
	1	578	1.69	1130
	2	917	1.65	1350
	3	1170	1.54	1430
200-4 μm	0	284	1.91	582
	1	556	1.62	1350
	2	720	1.29	1550
	3	960	1.17	1600
100-3 μm	0	342	1.47	1370
	1	1010	1.21	2570
	2	1480	1	2840
	3	1920	1	2840
100-4 μm	0	340	1.56	1610
	1	750	1.27	2820
	2	1060	1	2940
	3	1380	1	3000

Table 5.2: Comparison of measured bandwidths and fitting coefficients with calculated RC bandwidth based on measured capacitance and 50 Ω load.

by substrate diffusion photocurrent. The effect of the breakdown of the isolation junction with increasing illumination intensity was also observed as in the lot 3 devices.

Bandwidth measurement was performed for devices with diameters of 50, 100, and 200 μm, and electrode spacings of 1, 2, 3, and 4 μm. Due to the use of a revised mask set, reverse current was reduced compared to previous devices and the yield was improved allowing for a greater number of reliable measurements. The measurements were fit as described previously. As shown in Figure 5-27, the 200 μm diameter devices exhibit a rich set of varying behaviours. The 1 μm spacing devices were generally dominated at all biases by device capacitance. Despite the shortest inter-electrode distance, they were only faster than the 3 and 4 μm devices at low bias where those devices exhibit a large diffusion time constant. The fitting coefficient maintained a value close to 2 for all biases. The 2 μm devices exhibited the highest bandwidth, and operated within the range $2 < c < 3$ over the reverse bias range. The value $c = 2$ is ambiguous as it may signify a purely RC behaviour, or a particular spacing of diffusion and RC time constants on either side of the derived

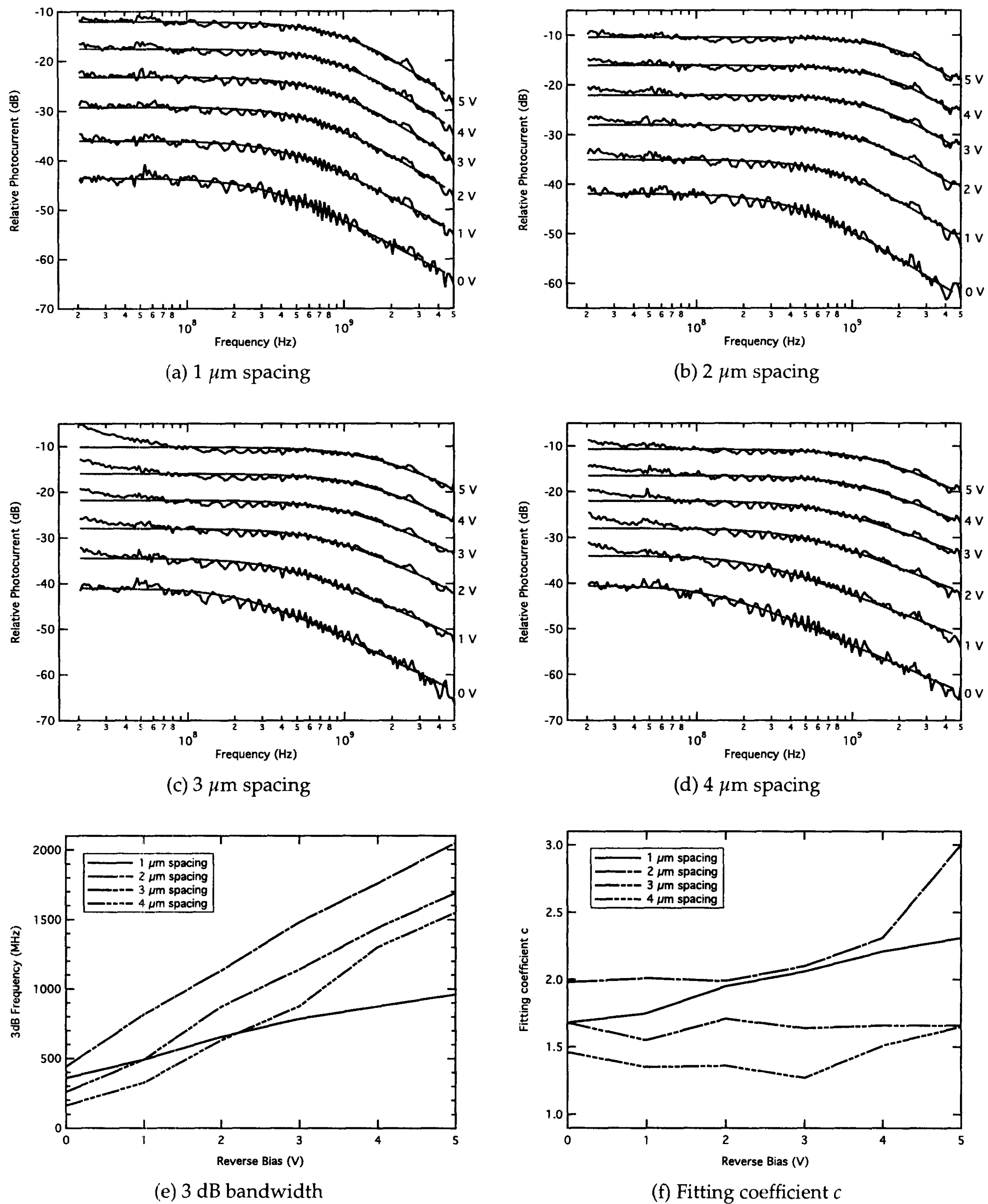
bandwidth that approximates a pure RC behaviour in the fit curve. The latter explanation is likely in the case of the 2 μm devices at low bias, whereas at higher bias the time constants coincide ($c = 3$). The wider devices exhibit a bandwidth influenced more strongly by carrier diffusion. As in the case of the lot 3 devices, the diffusion-dominated bandwidths are higher than expected. A comparison of fit bandwidths and measured RC time constants is shown in table 5.3. The measured bandwidths suggest that these devices are suitable for 2-3 Gbps applications depending on the precise geometry, which may be very useful considering the large device diameter.

Device	Reverse bias (V)	f_{3dB} (MHz)	c	f_{RC} (MHz)
200-2 μm	0	440	1.98	850
	1	815	2.01	1190
	2	1130	1.99	1410
	3	1480	2.1	1540
200-4 μm	0	161	1.46	1250
	1	325	1.35	1720
	2	630	1.36	1980
	3	875	1.27	2120

Table 5.3: Comparison of measured bandwidths and fitting coefficients with calculated RC bandwidth based on measured capacitance and 50 Ω load.

The frequency response of selected smaller devices is shown in Figure 5-28. It should be expected that the majority of these devices are limited by diffusion rather than capacitance due to their much smaller size, and this is indeed the case. Because of the low signal level required to remain in the isolated-substrate regime, and the low sensitivity of the detectors (due to the use of 850 nm light), it was not possible to measure beyond 5 GHz due to noise and crosstalk issues. The measurement limit of the system for a higher signal level was previously shown (section 3.3.4) to be approximately 8 GHz. Therefore, the data fits demonstrating bandwidths above 4 GHz may be questioned, since for such a bandwidth the trace appears nearly flat for the entire measurement window. In particular, some devices where an RC limit of around 3.5-4 GHz was expected actually exhibited a fitting coefficient $c = 1$, likely influenced by noise or calibration errors in the system (though the behaviour was repeatable between devices and measurement sessions).

The 100 μm diameter, 1 μm spacing device showed improved bandwidth compared to the 200 μm diameter device. The bandwidth is still clearly limited by capacitance, as the increase of bandwidth with reverse bias is slower than for the 3 and 4 μm spacing devices. The very strong



(a) 1 μm spacing

(b) 2 μm spacing

(c) 3 μm spacing

(d) 4 μm spacing

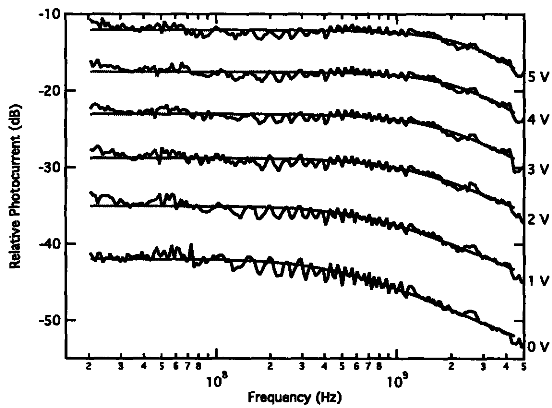
(e) 3 dB bandwidth

(f) Fitting coefficient c

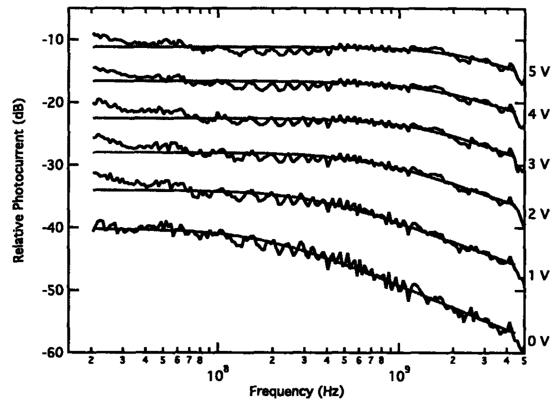
Figure 5-27: Extracted 3 dB frequencies and fitting coefficients of 200 μm diameter lot 4 devices with electrode spacing as noted below and within each diagram. In subfigures a-d, the data is shown by the black lines and the curve fit is shown by the grey overlaid lines.

match between the 50 and 100 μm diameter devices with the 3 and 4 μm spacings, respectively, conclusively demonstrates the diffusion-limited performance regime. It should be noted that the 4 μm device curves shown in Figure 5-28f are also very similar to the curve in Figure 5-27e indicating that the 200 μm devices are mostly dominated by diffusion, with only a small RC contribution (as could be concluded from the capacitance data alone). The diffusion bandwidth limit for the 4 μm devices is approximately 2 GHz at 5 V reverse bias, and approximately 4 GHz at 5 V reverse bias for the 3 μm devices. These values are far higher than originally expected during device design, and will be discussed in the context of device simulation in section 6.3 . They suggest a wider range of operation than previously envisioned, and also an additional contribution of the isolation junction to device performance. The devices with larger electrode spacing have a very low metal fill factor (0.25 and 0.2 for 3 and 4 μm spacing respectively), and may thus be considered good candidates for communication over POF using red (660 nm) VCSELs, currently commercially available at speeds approaching 2 Gbps (achievable in these devices for a convenient reverse bias of 3-4 V). The devices with 1 μm spacing clearly have the highest potential bandwidth, but achievable only for device sizes that are too small for POF system applications.

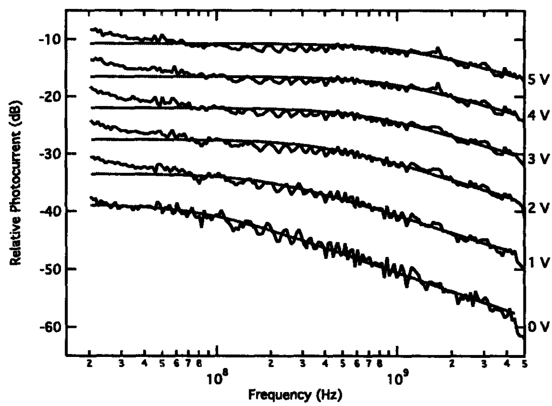
A comparison of fit bandwidth and calculated RC bandwidth for selected 50 and 100 μm lot 4 devices is shown in Table 5.4. The fitting factors suggest a strong diffusion time constant domination, and the measured capacitances (and time constants calculated therefrom) confirm that this should be the case. However, in some of the higher bandwidth cases (for example the 50-3 devices at 3 V bias) one would have expected a higher c that was derived from the fit. Conversely, the c for some zero bias measurements is above 1 whereas the 3 dB bandwidths and RC bandwidths differ by over a factor of ten (for example the 100-4 and 50-3 devices). These behaviours are examined further by simulation in section 6.3 .



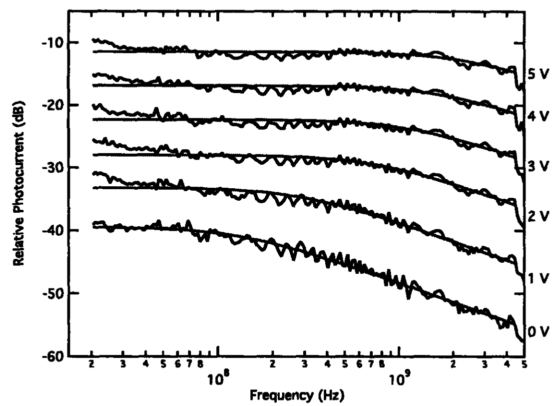
(a) 100 μm diameter, 1 μm spacing



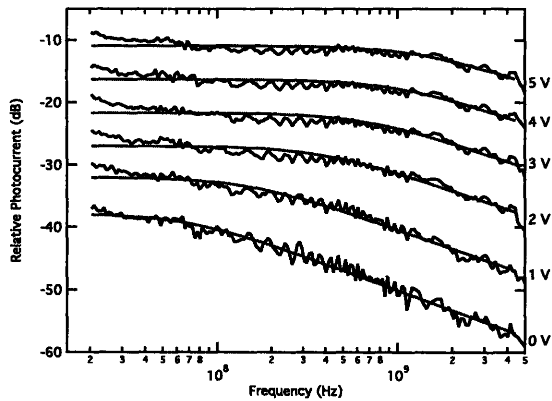
(b) 100 μm diameter, 3 μm spacing



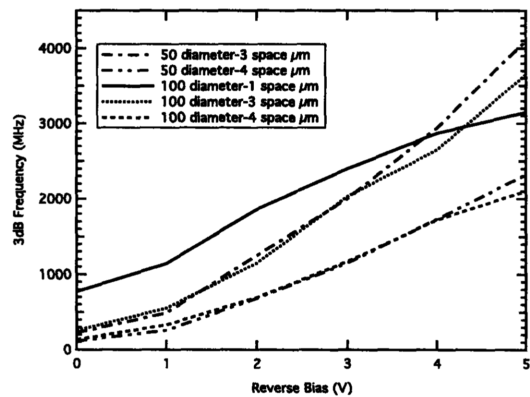
(c) 100 μm diameter, 4 μm spacing



(d) 50 μm diameter, 3 μm spacing



(e) 50 μm diameter, 4 μm spacing



(f) 3 dB Bandwidth

Figure 5-28: Extracted 3 dB frequencies and fitting coefficients of 200 μm diameter lot 4 devices with electrode spacing as noted below and within each diagram. The data is shown by black lines and the curve fit is shown by the grey overlaid lines.

Device	Reverse bias (V)	f_{3dB} (MHz)	c	f_{RC} (MHz)
100-3 μm	0	259	1.18	2620
	1	552	1.05	3240
	2	1150	1.01	3580
	3	2030	1.13	3720
100-4 μm	0	135	1.09	2950
	1	332	1	3540
	2	695	1	3840
	3	1170	1	3950
50-3 μm	0	226	1.47	4080
	1	490	1	4480
	2	1250	1	4820
	3	2000	1	4970

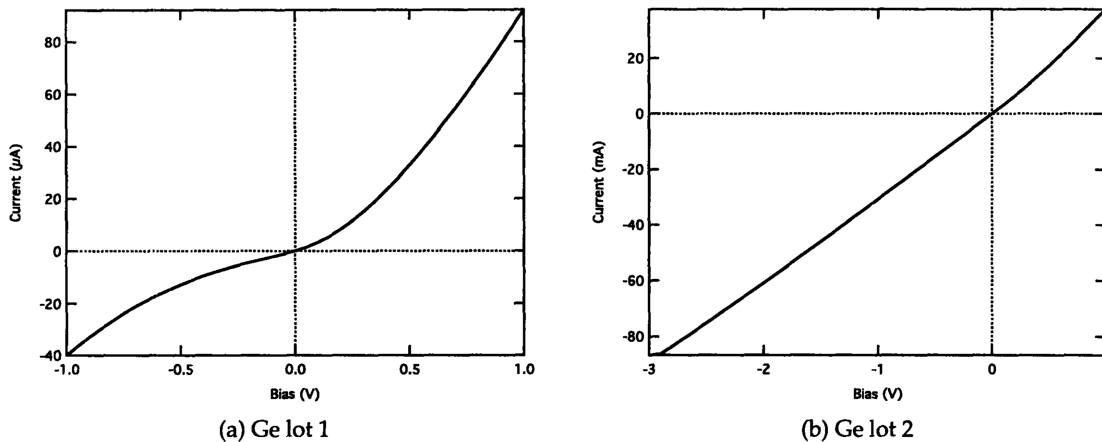
Table 5.4: Comparison of measured bandwidths and fitting coefficients with calculated RC bandwidth based on measured capacitance and 50 Ω load.

5.3 Germanium-on-Silicon lateral p-i-n photodetectors

5.3.1 DC Electrical Characteristics

Both of the germanium lots produced dominantly or entirely linear I-V characteristics as shown in Figure 5-29. The lot 1 devices exhibited a resistance on the order of tens of kΩ. This was believed to be caused by the misalignment of the implant definition masks that effectively introduced a short-circuit through the nominally undoped polysilicon electrodes. However, in lot 2 where alignment was controlled more precisely the observed resistance fell to the order of 50-300 Ω. The behaviour was consistent among devices, dies, and wafers. The resistance was qualitatively lower for larger devices. It was not possible to identify the precise leakage mechanism, but the possible explanations are conduction through dislocations/defects in the germanium, a surface potential pinning causing the formation of an inversion layer at the germanium surface, or a breakdown effect due to the apparently high epilayer doping [48] measured by four point probe method. Parallel currents of a similar order of magnitude were observed in devices from another project within the research group, originating from the same epitaxial growths as Ge lot 2, however the processing of those devices was much different and it is therefore not possible to make a direct comparison between the conduction mechanisms.

The strong dominance of an apparent leakage path over a diode-like behaviour made the direct



(a) Ge lot 1

(b) Ge lot 2

Figure 5-29: IV characteristics of typical devices fabricated in Ge lot 1 and lot 2.

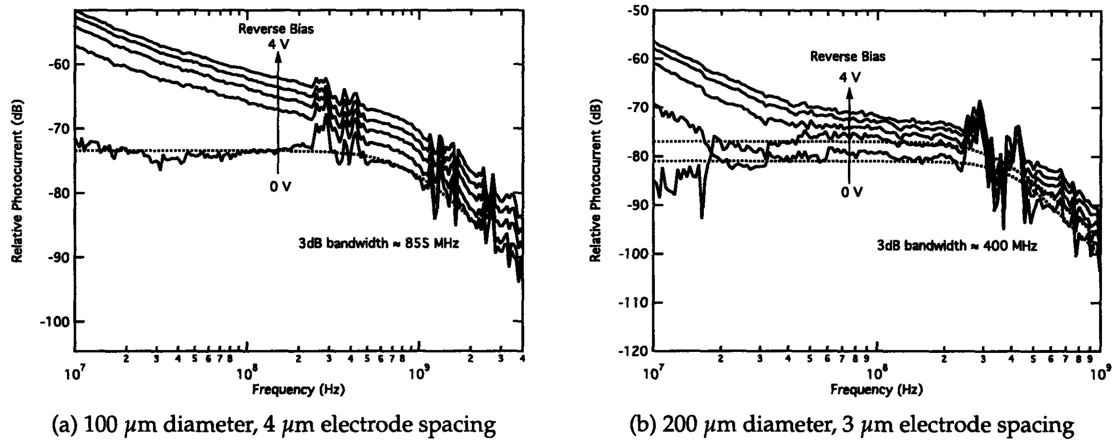
measurement of capacitance impossible in these devices. Some lot 1 devices had characteristics that were visually closer to a diode behaviour with a very large parasitic resistance at low bias levels. However, the application of a larger bias caused the devices to change to an IV curve as shown in Figure 5-29, indicating that some sort of annealing or degradation of the device material took place at the larger voltage and/or current. The lot 1 devices exhibited no response to illumination that could be measured in DC. On the other hand, the lot 2 devices exhibited such a strong parallel current that an extremely strong illumination would have been required to generate an observable level of photocurrent. The responsivity and capacitance of the Ge lot 1 and 2 devices will be discussed further in the next section with reference to the frequency response measurement.

5.3.2 AC Characteristics

Ge Lot 1

The frequency response of the Ge lot 1 devices was measured despite the fact that a DC photoresponse was not observed. Unexpectedly, devices with an IV characteristic, as shown in Figure 5-29, produced the most measurable photocurrent. For all devices, the signal level was very low corresponding to below 1% quantum efficiency. Typical responses are shown in Figure 5-30.

The data contain a low-frequency decay similar to that seen in silicon devices, which is therefore believed to originate from a diffusion mechanism. For both graphs, the curves have not



(a) 100 μm diameter, 4 μm electrode spacing

(b) 200 μm diameter, 3 μm electrode spacing

Figure 5-30: Typical frequency response measurements of Ge lot 1 devices. The measurement was performed using a Minicircuits ZKL-2R7 amplifier with a gain of 24 dB.

been offset as in previously presented data in order to emphasise the effect of interference and/or crosstalk due to the low signal level. A fit to the data reveals a roll-off varying with device size, and not changing significantly with reverse bias. This suggests a high doping level in the germanium layer such that capacitance and/or diffusion distance do not change substantially with increasing bias. Additionally, the very low signal level suggests that there is a very high level of recombination (i.e. short minority carrier lifetime) in the germanium layer. A very short recombination lifetime would not be able to support the measured frequency response decaying even at low frequencies. Therefore it is believed that the collected photocurrent in fact originates from the silicon substrate, but that a large proportion of the incoming light is absorbed by the germanium epilayer. In order to confirm this hypothesis, the photocurrent at $f = 10$MHz was plotted as a function of epilayer thickness for various reverse bias. The fabrication lot included wafers with three different epilayer thicknesses. Assuming a decay of the form $e^{-L/\alpha}$, the semi-logarithmic plot of photocurrent as a function of germanium thickness has a linear fit with slope corresponding to the absorption distance $1/\alpha$. This data is shown in Figure 5-31.

At 1 V reverse bias, the effective absorption length from the linear fit is 487 nm. This is a very good approximation to the actual value of 350 nm, particularly considering that the stated germanium thickness is only nominal (i.e. etching losses during processing are neglected, particularly the Ge RCA cleaning process) and that some photocurrent may originate in the Ge layer near the contacts. The effective absorption distance rises with reverse bias as expected, indicat-

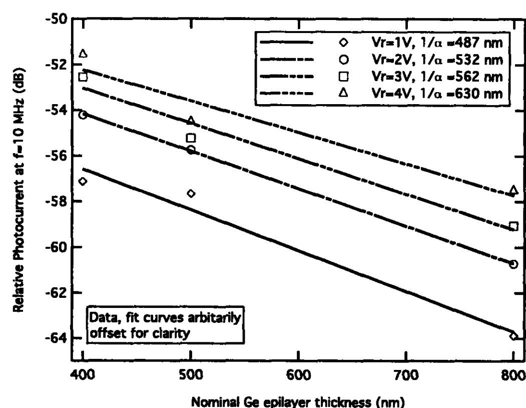


Figure 5-31: Photocurrent as a function of epilayer thickness for different reverse biases. The data is taken from measurement of 100 μm diameter, 4 μm finger spacing devices from three different wafers in the fabrication lot.

ing a small but increasing photocurrent collection in the Ge layer. This analysis together with the shape of the frequency response confirms that the majority of collected photocurrent is generated in the silicon growth substrate rather than in the Ge epilayer. Therefore, it must be concluded that the germanium material quality was extremely poor and unsuitable for high-performance device fabrication.

Ge Lot 2

Due to the large parallel current of the lot 2 devices, it was not possible to measure a DC photocurrent directly. A lock-in technique was found to be problematic due to the large difference in dynamic range between the "dark" current and the photocurrent. However, in the frequency measurement apparatus the bias voltage and DC current are separated from the AC signal by the Bias T at the receiver side, making it possible to directly measure the optoelectronic transfer function despite the difference in signal magnitudes, assuming that the device is a junction in parallel with a resistance as shown in Figure 5-32. The measurement is made difficult by the magnitude of the parallel current which results in substantial local heating of the sample, inducing drift in the measured trace and sometimes burning of the pad-ohmic contact connections.

It is possible to estimate the DC device responsivity from the frequency response [14]. Knowing the specified VCSEL source slope efficiency (i.e. the electrical-optical transfer function), losses in the optical path (3 dB for the optical U bench with attenuator as described in section 3.3.1), and

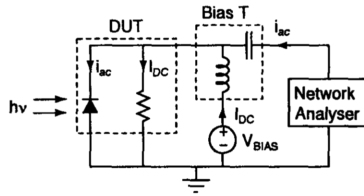


Figure 5-32: Schematic representation of the device measurement, with the DUT assumed to be a diode in parallel with a parasitic resistor.

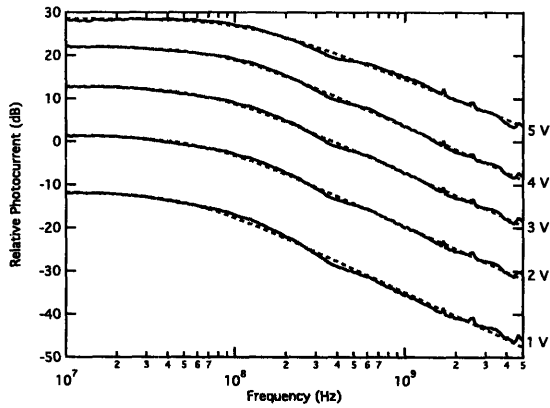
the measured S12 value, the optical-electrical transfer function may be calculated. It was found that the measured values varied significantly between devices of different sizes and spacings, due to the problems of germanium device fabrication described in section 4.3.3. The results of the best devices are summarised in table 5.5.

Electrode Spacing	S12 at f=10 MHz	Sensitivity	External QE (850 nm)	Internal QE (850 nm)
2 μm	-33.5 dB	0.4 A/W	0.59	0.85
3 μm	-37 dB	0.25 A/W	0.37	0.49
4 μm	-36.5 dB	0.28 A/W	0.41	0.51

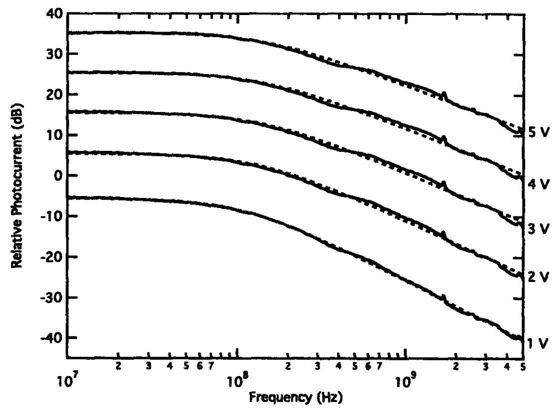
Table 5.5: Responsivity figures extracted from the optoelectronic frequency response measurement of Ge lot 2 devices.

The data indicates very good quantum efficiency. The result of the 2 μm device is close to the estimate of approximately 0.65 external QE for this device layout. The wider electrode spacing devices show an unexpectedly lower efficiency, though the ratio of the 3 and 4 μm devices is as expected. The photoresponse data suggest that the Ge lot 2 epitaxial material is of much higher quality than the lot 1 germanium.

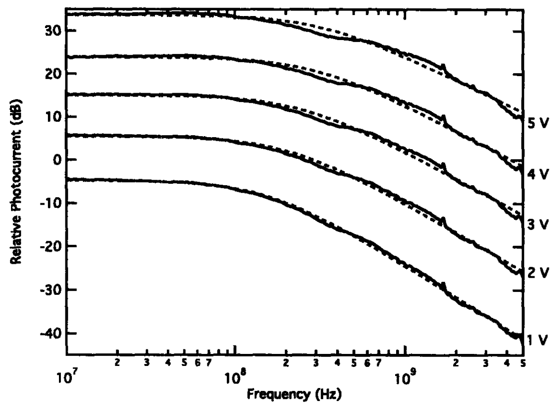
Typical frequency response curves of the Ge lot 2 devices are shown in Figure 5-33. Data is presented for the 100 μm diameter devices with three different finger spacings in order to demonstrate the clean and consistent data for the three devices geometries. The fitting function is shown in the dashed lines and corresponds very well to the data. All of the curves has a fitting coefficient c of approximately 2. This suggests an RC-dominated behaviour, confirmed by Figure 5-33d where the bandwidth increases with increasing finger spacing (i.e. lower capacitance per device area). The carrier diffusion delay of equation 2.32, assuming no lateral depletion, predicts a higher bandwidth than measured for the 2 and 3 μm spacings: 950 MHz and 425 MHz respectively. The diffusion bandwidth for a 4 μm diffusion distance is estimated to be 240 MHz, or lower than the



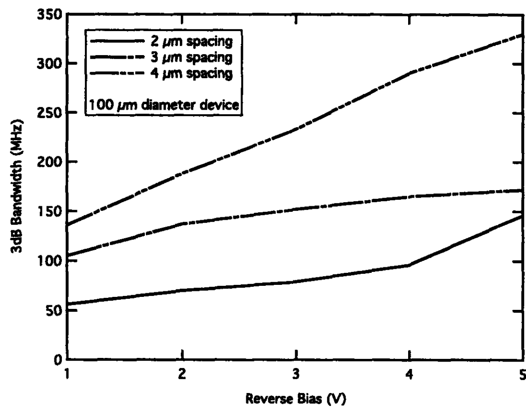
(a) 100 μm diameter, 2 μm spacing



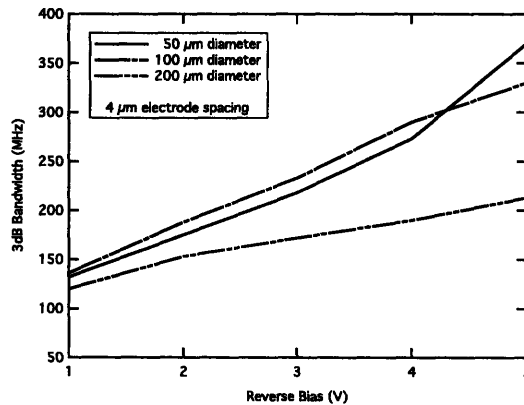
(b) 100 μm diameter, 3 μm spacing



(c) 100 μm diameter, 4 μm spacing



(d) 3dB bandwidth: 100 μm diameter devices, different electrode spacings



(e) 3dB bandwidth: 4 μm electrode spacing, different device diameters

Figure 5-33: Frequency response measurement of selected devices, and extracted bandwidth as a function of electrode spacing for constant device diameter and as a function of device diameter for constant electrode spacing.

measured bandwidth of 4 μm spacing devices. However, some lateral extent of the depletion region would explain this discrepancy.

The RC-dominated behaviour suggests a high germanium layer doping, consistent with four point probe measurement, where the measured resistivity corresponds to a mid-10^{18} cm^{-3} doping level due either to unintentional dopants or crystal defects. Such a high doping level would also result in a decrease in mobility and therefore minority carrier diffusion coefficient [94], such that the diffusion bandwidths estimated above might require reestimation. Interestingly, as seen in Figure 5-33e, even though the shape of the frequency response and corresponding fitting factor c indicate an RC-dominated response in the case of all devices, the measured bandwidth value does not scale as expected with device diameter. The 50 μm and 100 μm diameter devices are nearly identical bandwidths despite a factor of 4 difference in area, likewise the 200 μm device bandwidth does not scale as expected compared to the 100 μm device. There are several possible explanations for this effect.

1. The devices parallel current tends to scale with device size, therefore the junction sees a smaller parallel resistance. The very large devices may therefore have a smaller RC time constant due to the small effective load resistance.

2. Even the best devices contained some pinhole defects in the germanium epilayer following metal sintering, as discussed in section 4.3.3. During the measurement, the fibre probe was positioned over the devices in order to maximise the received signal. It was found that the optimum position, was often off-centre. It is possible that the holes in the germanium result in effective open circuits or recombination centres, cutting off parts of the device. Therefore, the effective areas, in the cases of all devices, may be similar despite the actual device sizes. This may also explain the wide range of signal levels and discrepancy noted in the responsivity discussion above where the 2 μm spacing devices showed higher photoresponse than the 3 and 4 μm devices.

3. A trapping mechanism associated with defect and/or dislocations in this material has approximately a first-order response. This is not expected, since trap time constants are generally several orders of magnitude below the frequencies of interest in the measurement.

Chapter 6

Simulation and Discussion

Measurement of the fabricated devices in chapter 5 revealed two unexpected device behaviours: the intensity-dependence of the responsivity with a consequent effect on the frequency response of the isolated detectors, and the smaller than expected dependence of electrode spacing on the 3 dB bandwidth. The latter is particularly interesting since the resultant device bandwidths for smaller diameter 3 and 4 μm spacing devices were an order of magnitude greater than expected. Simulations of the device structure, were carried out to confirm the intensity dependence and bandwidth versus finger spacing dependence, to further understand the device performance limits, and to confirm expected performance of related structures including ones fabricated on germanium epilayers. Finally, based on the measurement and simulation, a new device structure is proposed combining SOI with a buried junction. Additionally, the optimal Si substrate doping polarity is identified for subsequent Ge growth and lateral device fabrication.

Simulation structure

Simulation was carried out using MEDICI software from Synopsys Inc., a two-dimensional Poisson equation and electron and hole continuity equation solver, incorporating a number of physical models for recombination, photogeneration, lifetime, mobility, and other important phenomena. It includes the ability to study devices under DC, transient, and sinusoidal steady-state conditions. Additional useful features include the ability to attach lumped passive elements to contacts, to specify boundary conditions, and to define incident illumination by ray-tracing in-

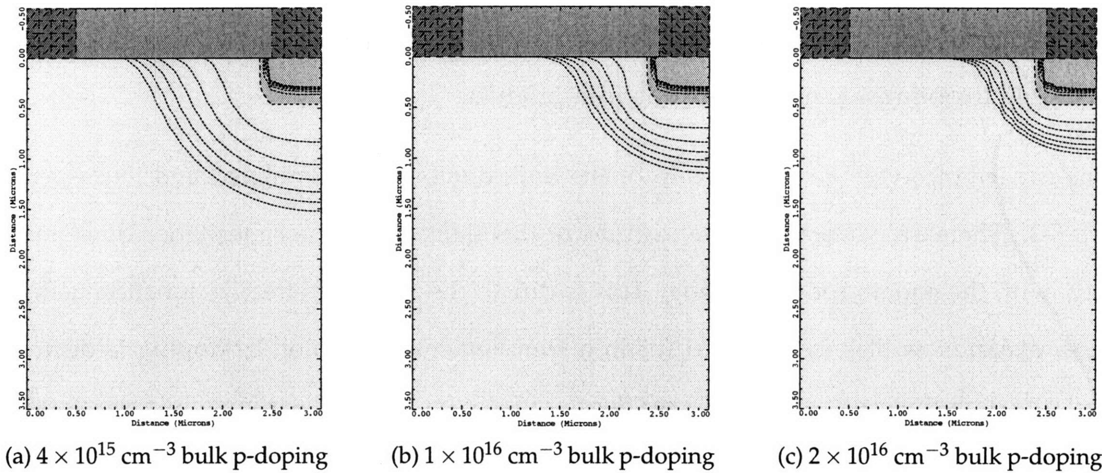
cluding the calculation of transmission through a stack of layers by a transmission matrix method.

The basic device structure used in the simulations is a simplified version of the fabricated device. A bulk substrate doping is specified, and the active region is defined as a flat doping profile with Gaussian tails in the top 0.2 μm (i.e. at the semiconductor surface) and in the 0.25 μm before the specified metallurgical junction depth. This approximates the condition of a well created by implantation. Alternatively uniform active region doping may be used to represent the condition of epitaxial growth. The implanted finger regions are Gaussian profiles with 0.45 μm deep junctions. It is also possible to import an impurity profile generated by the SUPREM device process simulator corresponding exactly to the fabricated devices, however the additional complexity is not particularly instructive. The electrodes are defined as non-transparent ohmic contacts to avoid the simulation complexity of the metal-semiconductor junction, and a silicon dioxide passivation layer is defined. The structure is a two-dimensional representation of the device, assumed to be infinite in the third dimension. A full device simulation including the third dimension is achieved by specifying a lumped passive element at the contact with the units $\Omega \cdot \mu$m. The substrate node is left unconnected unless otherwise specified.

6.1 Depletion Region Shape, Capacitance

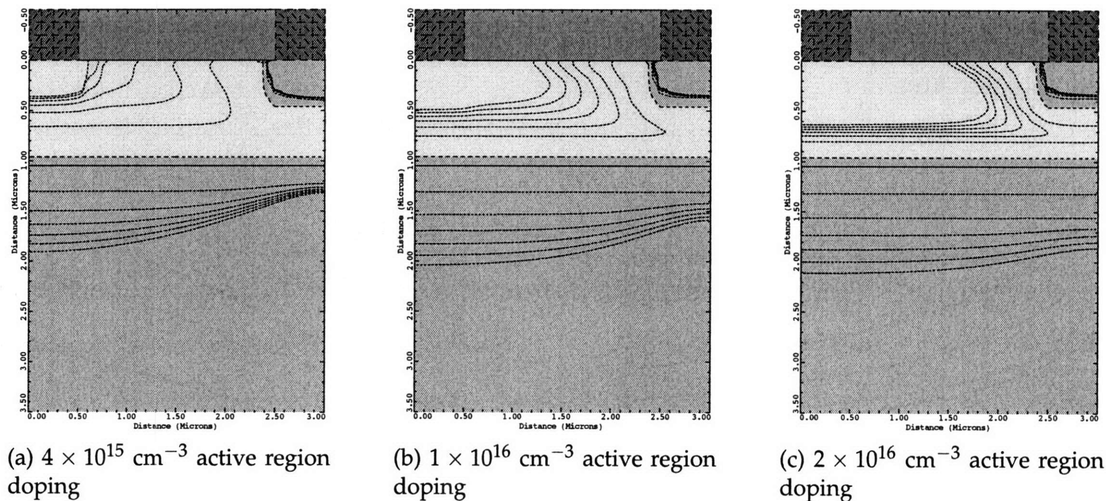
Control devices (bulk p-type substrates) were simulated to observe the expansion of the depletion region under reverse bias for different doping conditions as shown in Figure 6-1. As expected, the depletion region expands approximately uniformly around the n+ finger region. From these images, it is apparent why published reports of simple interdigitated structures include measurements at large reverse biases of 10 to 40 V [14], particularly when the application is an 850 nm system. The isolated structure discussed in previous sections serves two purposes: it blocks carriers generated in the substrate, and it forces the device depletion area to expand laterally rather than vertically or uniformly. The mechanism for the lateral expansion was discussed with reference to fabricated GaAs devices in section 5.1.2, and may be confirmed by observing the simulated devices in Figure 6-2.

Comparing bulk and isolated devices with the same active region doping (where the bulk is the active region in the former devices), the isolated devices have much greater lateral depletion



(a) 4×10^{15} cm^{-3} bulk p-doping (b) 1×10^{16} cm^{-3} bulk p-doping (c) 2×10^{16} cm^{-3} bulk p-doping

Figure 6-1: Simulation of a lateral p-i-n structure assuming a bulk p-type substrate of different doping concentration. The dotted lines are the edges of the depletion region for reverse bias from 0 to 5 V.

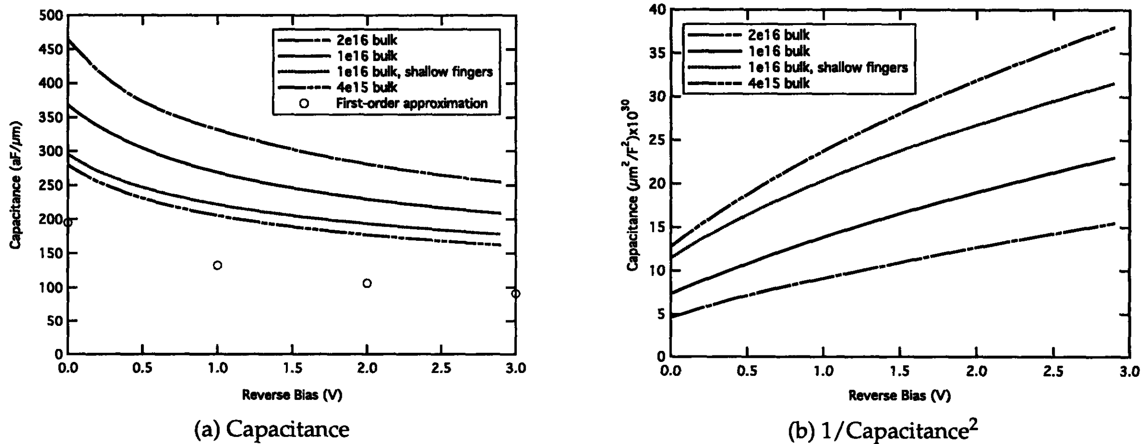


(a) 4×10^{15} cm^{-3} active region doping (b) 1×10^{16} cm^{-3} active region doping (c) 2×10^{16} cm^{-3} active region doping

Figure 6-2: Simulation of a lateral p-i-n structure assuming an isolated structure with different active region doping concentration and constant bulk doping concentration of 4×10^{15} cm^{-3}. The dotted lines are the edges of the depletion region for reverse bias from 0 to 5 V.

than the bulk devices for the same reverse bias. This points to the second advantage of the isolated structure over the bulk structure stated above, i.e. that the entire region between the electrodes may be depleted with carriers travelling solely by drift, or alternately that a larger finger spacing may be used to optimise sensitivity-bandwidth product.

The capacitance-voltage relationship of the bulk devices was simulated and is presented in Figure 6-3. There are several notable features of this data. First, the capacitance does not scale directly with the square root of doping. This is due to the factor of effective junction area, since the area expands as bias is increased (or in a complementary fashion as doping is decreased). Second, the dark lines all assume a finger depth of 0.45 μm as noted earlier, while the dotted line represents a shallower finger implant of 0.15 μm (with 10^{16} cm^{-3} bulk doping). The capacitance is significantly lower due to the smaller sidewall length. This result directly suggests against deep trench strategies as employed for example in [16], because the additional sidewall area increases capacitance significantly compromising the available design space. Finally, the circles indicate the first-order estimates of section 4.1.2 assuming 10^{16} cm^{-3} doping. They are based on a vertical junction under the finger area, ignoring implant finger sidewall and uniform expansion of the depletion region, and therefore significantly under-estimating device capacitance. The $1/C^2$ data is also shown. The data is not linear despite the uniform doping due to the increasing effective depletion region area.



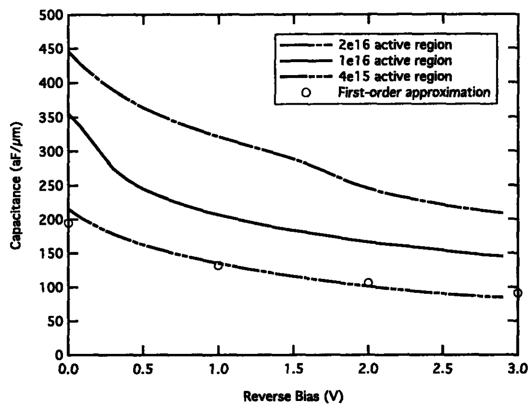
(a) Capacitance

(b) $1/\text{Capacitance}^2$

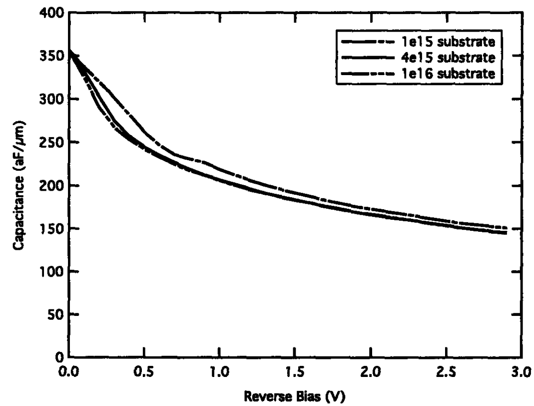
Figure 6-3: Simulation of bulk device capacitance as a function of voltage for different bulk doping levels.

The capacitance-voltage relationship of the isolated devices is presented in Figure 6-4. The isolated device capacitances are lower than those of bulk devices with the same active region doping level. Additionally, the capacitance decay is not necessarily a smooth curve, depending on the doping profile of the active region. For example, the 1×10^{16} cm^{-3} curve exhibits a sharp drop at very low bias reminiscent of the measured Si lot 3 devices. Referring to the image of depletion region progression with bias (Figure 6-2), there is a significant change in region shape between 0 and 1 V reverse bias as the region under the n finger is depleted laterally for higher doping levels. The capacitance is lower than in the case of the bulk devices due to the smaller effective depletion region area and longer depletion distance. The first-order approximation points for 1×10^{16} cm^{-3} doping are again shown as circles, and the match between the approximation and simulated values is much closer than in the case of the bulk devices. Additionally, it is important to note that the substrate doping does not have a significant impact on device capacitance as shown in Figure 6-4b for a range of an order of magnitude in substrate doping. For cases where the entire region between the implant fingers is not depleted (such as for the case of active region doping of 1×10^{16} cm^{-3} and substrate doping of 4×10^{15} cm^{-3} as shown in Figure 6-2), the normalised device capacitance is essentially independent of finger spacing. The depletion region expands outwards from the n finger and interacts with the substrate-active region junction. The p finger is an ohmic contact to the active region setting the potential of the QNR portion of the active region. Therefore the precise location of that contact does not affect the junction until the depletion region extends to the boundary of the p finger. This is a useful device design tool since the finger spacing may be optimised without recalculating the device capacitance at each step.

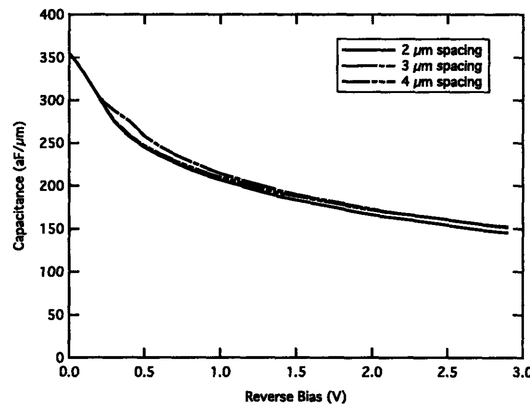
The capacitance simulation data shown in the previous figures is presented in units of aF/μm. For device design consideration it may be divided by the width of the simulation unit cell to yield a capacitance per device area or multiplied by a length to yield a capacitance for a particular device size. That analysis would yield an expected separation of the three finger spacing curves of Figure 6-4c. A convenient multiplication factor to convert the 2 μm spacing simulation data to a 200 μm device diameter is 10^4. From an RC perspective, over 2 GHz of bandwidth may be achievable for that device diameter which is promising for POF applications. Wider spacings decrease capacitance further, and lower active region doping decreases both capacitance and carrier diffusion distance.



(a) Varied active region doping



(b) Varied substrate doping



(c) Varied finger spacing

Figure 6-4: Simulation of isolated device capacitance as a function of voltage for 2 μm finger spacing and a fixed substrate doping of 4×10^{15} cm^{-3} (a), a fixed active region doping of 1×10^{16} cm^{-3} (b), and fixed active region and substrate doping (1×10^{16} cm^{-3} and 4×10^{15}cm^{-3} respectively) (c).

As shown in Figure 6-5, the simulated capacitances compare favourably to the measured results of Si lots 3 and 4. The simulated structures are chosen to have similar but not necessarily indentical active region doping as the measured devices. Additionally, a parasitic pad capacitance as discussed in section 5.2.2 is added to the simulated values. The precise details of the metallisation and doping concentrations explain the slightly different shape of the lot 3 device curve at low bias. Otherwise the match is excellent, particularly for the lot 4 devices, confirming that simulation may be used to examine and predict device behaviour.

The effect of substrate doping on the depletion region may also be seen qualitatively in Figure 6-6. Somewhat counter-intuitively, a higher substrate doping leads to greater lateral depletion

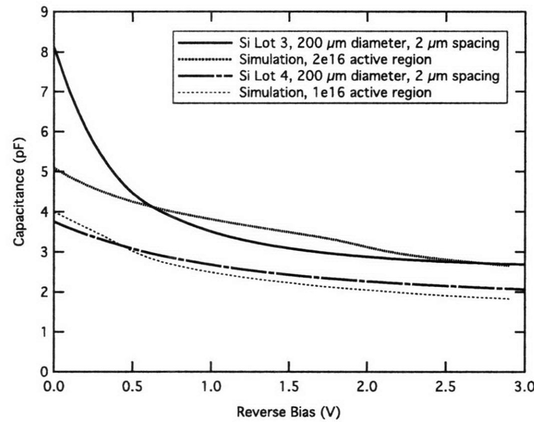
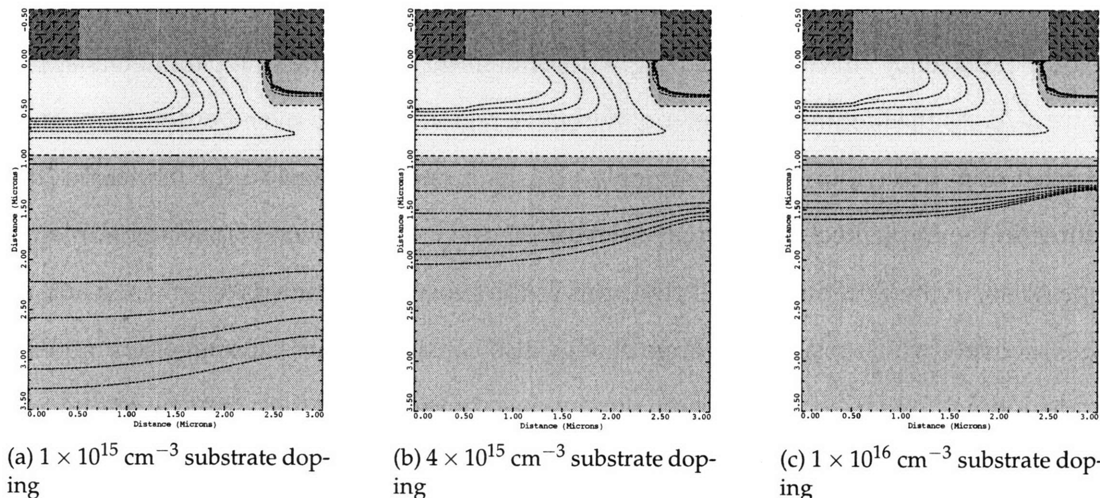


Figure 6-5: Measured capacitance curves of 200 μm diameter, 2 μm finger spacing devices from Si lots 3 and 4 (dark lines) and simulated capacitance for similar structures (dotted lines).

of the active region. This arises from the properties of the p-n junction, where the majority of potential drop occurs on the lightly-doped side of the junction. Therefore, the depletion region edge on the lightly doped side moves farther than on the higher-doped side when bias is applied as is clearly seen in the figure. This suggests that an epitaxial growth of lightly-doped active region on a highly doped substrate may be an interesting device structure in terms of lateral depletion at low reverse bias.



(a) 1×10^{15} cm^{-3} substrate doping

(b) 4×10^{15} cm^{-3} substrate doping

(c) 1×10^{16} cm^{-3} substrate doping

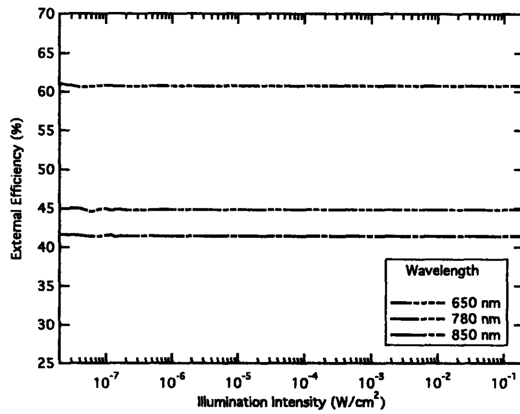
Figure 6-6: Simulation of a lateral p-i-n structure assuming an isolated structure with different active region doping concentration and constant bulk doping concentration of 4×10^{15} cm^{-3}. The dotted lines are the edges of the depletion region at reverse bias from 0 to 5 V.

6.2 Isolation Junction Behaviour

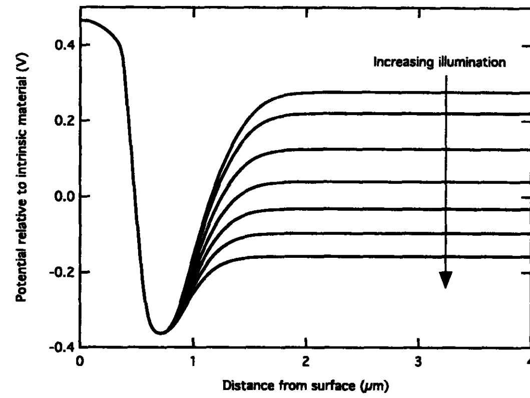
The behaviour of the bottom isolation junction was modelled to confirm the measurement results and explanation as presented in section 5.2.3. DC results are shown in Figure 6-7. First, control devices (i.e. lateral geometry on a bulk p-type substrate) were simulated as a baseline to confirm that sensitivity is independent of illumination intensity. This was the case, and the external efficiency under 650 nm illumination is approximately 100% as expected. However, the external efficiency of longer wavelengths was lower than expected considering a minority carrier diffusion length on the order of 100 μm in lightly doping silicon. It was found that the simulation program assumed a base carrier lifetime of 100 ns, and changing this to a larger value recreated the high efficiencies at 780 and 850 nm as measured in section 5.2.3. The potential as a function of position (relative to intrinsic silicon) was calculated for different illumination intensities (Figure 6-7b), and the substrate potential was found to change with illumination, consistent with the explanation presented earlier. Because the substrate is a hanging node, its potential must change in order to counteract the photocurrent crossing the active region-substrate junction.

The external efficiency as a function of illumination was simulated for devices with active region doping of 2×10^{16} cm^{-3} and substrate doping of 1×10^{15} cm^{-3}. The transition from low to high sensitivity moves to higher illumination as finger spacing is increased, as expected. The shape of the efficiency curve is also strongly dependent on carrier lifetime parameters as shown in Figure 6-7c. As the lifetime parameter of the active region is varied over 2 orders of magnitude, the photocurrent behaviour changes strongly. This is a relevant effect to the fabricated devices as lifetime in the implanted regions may be reduced due to incompletely annealed implantation damage. Also, in the extreme left and right curves the substrate lifetime is lengthened and shortened respectively which results in a large change also consistent with the suggested model. The substrate doping (for a fixed active region doping) has some effect on the curves by setting a larger thermal equilibrium substrate potential. The effect is coupled to the concentration-dependence of lifetime.

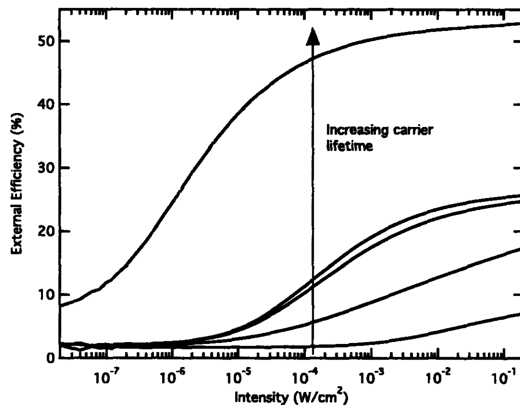
The effect of the isolation junction behaviour on the device frequency response is shown in Figure 6-8. The noted rise in sensitivity means that carriers generated in the substrate are collected at the finger electrodes. These carriers travel primarily by diffusion, and when they add to the



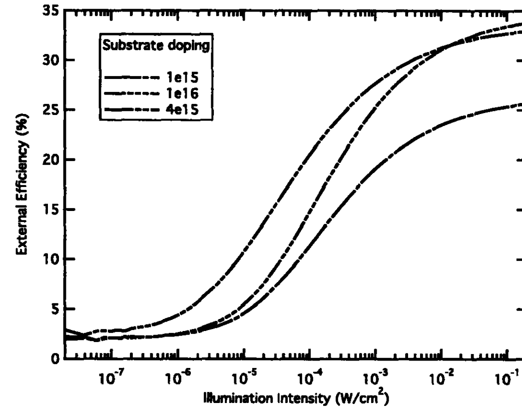
(a) Control Devices



(b) Substrate self-biasing



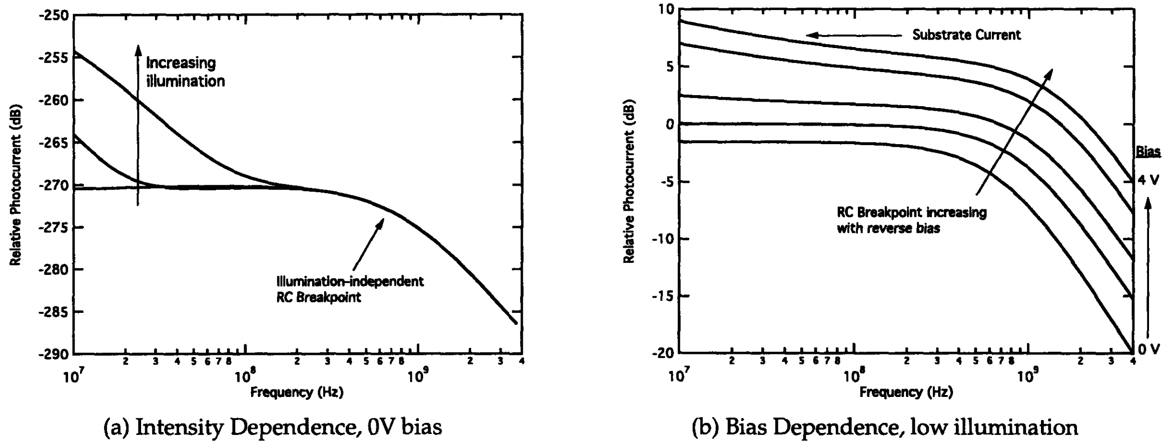
(c) Simulation lifetime parameter dependence



(d) Substrate doping dependence

Figure 6-7: Simulation of device intensity dependence. The substrate equilibrium potential changes with illumination and the behaviour is strongly dependent on carrier lifetime.

lateral current response they create an undesirable low frequency peak as seen in the figure. The RC-related breakpoint remains at the same frequency, independent of the substrate current. This is the identical behaviour as observed in Figure 5-24. A further effect may be noticed in the frequency response: as reverse bias is increased, carriers from the substrate appear at low frequencies. This is also observed in the measured data, and may be eliminated by lowering the illumination intensity (the higher the bias, the lower the maximum intensity for the effect not to be noticeable in the measurement). The other solutions to the substrate current problem are a physical isolation barrier (e.g. SOI) or electrical pinning of the substrate potential (i.e. a backgate). Also, the severity of the problem decreases as a wavelength of light is chosen such that the absorption length is close to the active region depth.



(a) Intensity Dependence, 0V bias

(b) Bias Dependence, low illumination

Figure 6-8: Simulated effect of substrate current on bandwidth as a function of illumination and bias.

6.3 Frequency Response

The simulations were performed using parameters similar, but not identical, to the fabricated devices as discussed in section 6. This was done in order to concentrate on trends in the simulation rather than a precise recreation of the intended device structure. The resulting match between simulated and measured structures was excellent, as shown in Figure 6-9 for both Si lot 3 and lot 4 devices. There is a small deviation at zero bias, particularly in the lot 3 devices. This comes from the large capacitance at very low bias as seen in Figure 5-15. Otherwise, the simulations recreate all important dynamic features of the response including the consistent RC and lateral-diffusion influenced breakpoint and low-frequency contribution of long-distance current at higher bias. The simulation as shown was performed assuming a 50 Ω load and 200 μm diameter device. The excellent match between the fabricated and simulated devices (for two different fabrication lots with different implant parameters) validates the fabrication, measurement, and simulation procedures by demonstrating self-consistency.

Simulations of device bandwidth as a function of various device parameters are shown in Figure 6-10. In each subfigure, one parameter is altered and may be compared with the measured data or expected result. The device bandwidth as a function of electrode spacing is similar to the measured results of Si lots 3 and 4 presented in section 5.2.4. The bandwidth decreases with finger spacing, but not as significantly as expected. The bandwidth as a function of electrode spacing

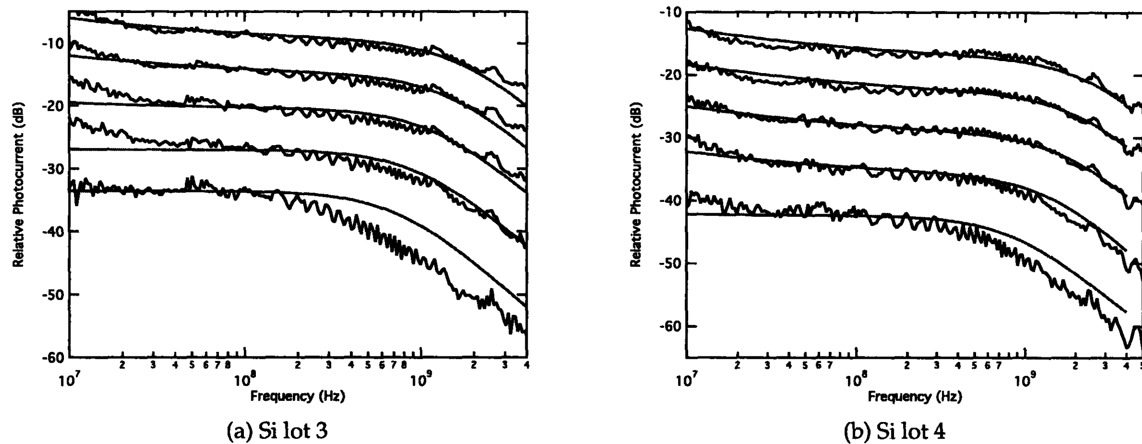
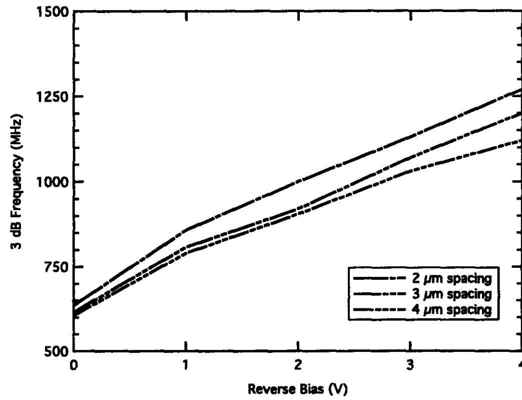


Figure 6-9: Measured (dark) and simulated (smooth overlay) characteristics of Si lot 3 and 4 devices. The device geometry in both cases is 200 μm diameter, 2 μm electrode spacing.

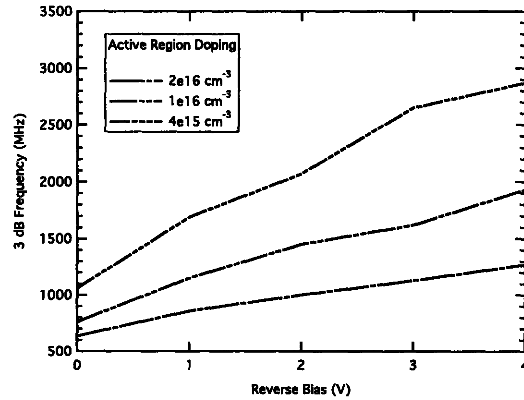
will be discussed further below. Bandwidth increases with decreasing active region doping, as expected. Lower doping decreases capacitance and increases the size of the QNR for a given reverse bias. Likewise, the slope of the bandwidth versus reverse bias curves increases with decreased doping. This is consistent with the expressions for depletion capacitance and QNR width which are dependent on the square root of doping concentration. On the other hand, bandwidth is only weakly dependent on substrate doping. This is also logical given the pictures of the depletion region shown in section 6.1 . The doping of the substrate has a weak influence on the vertical extent of the isolation junction QNR, however this influence is only on the order of 100 nm for a wide range of doping levels.

The influence of device diameter on bandwidth is seen in Figure 6-10d as well as in the measurements of the silicon device in section 5.2.4. For larger devices, RC effects play a strong role in device behaviour. As the device shrinks, bandwidth rises due to the smaller capacitance and may reach very high values (4 GHz in the measurement of 3 μm spacing Si lot 4 devices, and above 5 GHz in the simulated Si devices with 2 μm spacing and with doping structure similar to Si lot 3). Once again, this is a much higher bandwidth than was expected considering that the device is operating in a mode where the depletion region does not extend through the entire width of the inter-finger space.

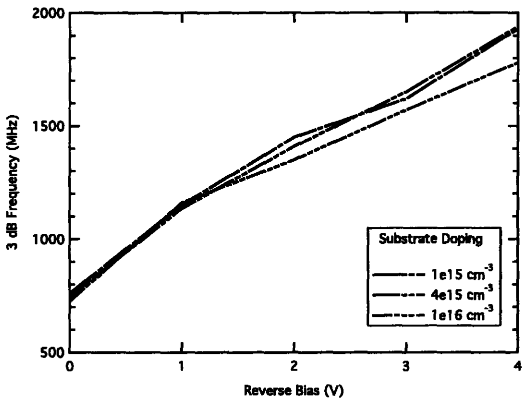
In order to eliminate the intensity-dependence of the devices, a potential may be applied to the



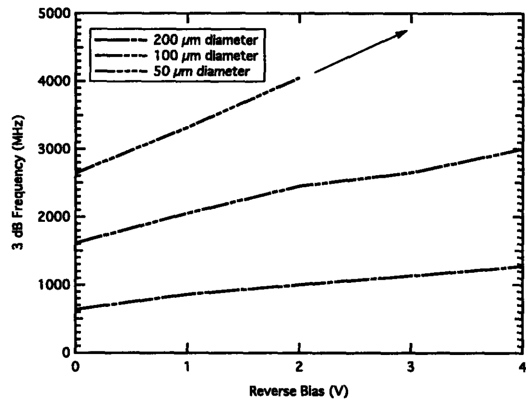
(a) Active region doping 2×10^{16} cm^{-3}, substrate doping 10^{15} cm^{-3}, 200 μm diameter



(b) substrate doping 10^{15} cm^{-3}, 200 μm diameter, 2 μm finger spacing



(c) Active region doping 10^{16} cm^{-3}, 200 um diameter, 2 μm finger spacing



(d) Active region doping 2×10^{16} cm^{-3}, substrate doping 10^{15} cm^{-3}, 2 μm finger spacing

Figure 6-10: Simulation of device bandwidth sensitivity to various fabrication parameters.

substrate node to prevent forward-biasing of the substrate-active region junction. If this is done, the sensitivity is uniform over many orders of magnitude of illumination intensity, and the small signal response at low frequencies is flat for all bias voltages up to approximately the substrate voltage, at which point a similar behaviour begins as already described for the floating substrate devices. The bandwidth is slightly improved as more of the active region is depleted vertically due to the applied bias. The disadvantage of such an arrangement is a lower responsivity as the substrate competes for generated carriers with the top electrodes. Increasing finger spacing does not improve responsivity in the expected way, since the gradient of minority carrier laterally decreases with increasing spacing.

These simulated bandwidths may be compared with simulations and reported experiments of

devices fabricated on SOI. In that case, there is no bottom junction and the SOI layer is simply a thin bulk material with the physical barrier to carrier transport. Such devices were reported by Schaub *et al.* multiple times [21, 18]. Several structures are simulated, starting with the one reported to result in a bandwidth of 3.4 GHz at 5 V reverse bias (2.7 μm SOI thickness, 3 μm BOX, 50-100 $\Omega \cdot$ cm resistivity). Then, the same SOI thickness was maintained, but doping increased to 10^{15} and 10^{16} cm^{-3}. The results are shown in Figure 6-11.

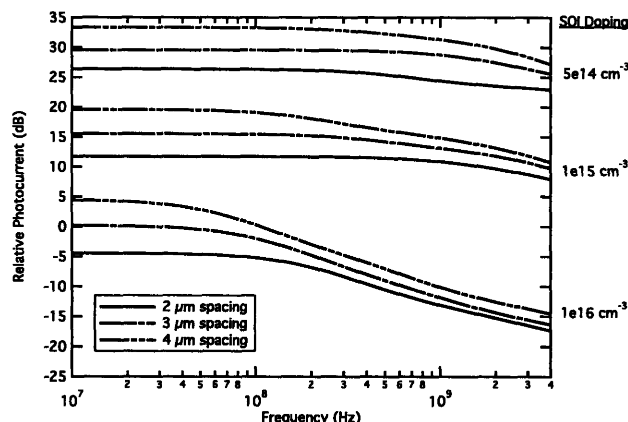


Figure 6-11: Simulated SOI detectors based on the structure and results reported in [18], and including different finger spacings and doping levels. The three groups of devices based on doping level are offset for clarity.

The first report of the SOI detectors [18] used a 2 μm finger spacing. The simulation confirms a bandwidth of approximately 2.5 GHz at 5 V reverse bias, consistent with the reported measurement. Interestingly, as a result of the very high resistivity material, the entire SOI region is depleted at that bias for both 2 and 3 μm finger spacings. However, the 3 μm spacing has a slightly higher bandwidth due to an unfavourable electric field profile under the p finger in the narrow devices. This was likely realised after publication, and the subsequent report [21] used a wider spacing. The high bandwidth, relative insensitivity to finger spacing, and high sensitivity due to the thick SOI layer are a consequence of the high resistivity SOI substrate. Figure 6-11 also shows the effect of increased doping on bandwidth. The junction area is kept constant in the calculation, i.e. the device area grows as the electrode spacing is increased, therefore the RC time constant is the same among all devices of a given SOI doping level. When doping is increased to 1×10^{15} cm^{-3}, the 2 μm device bandwidth is not substantially changed, but the 3 μm and especially the 4 μm bandwidths are degraded. When doping is increased to 1×10^{16} cm^{-3} all devices bandwidths

are significantly reduced, with bandwidth varying inversely with electrode spacing. Also, the frequency decay slope is approximately 10 dB/decade, consistent with the derivations presented in earlier sections and with measurement and curve fitting, indicating a carrier diffusion limited response that is slower with increasing QNR extent (i.e. increasing finger spacing).

This simulation is presented for two reasons. First, to again validate the frequency response derivation of previous chapters with respect to diffusion and RC bandwidth limits. Second, to confirm that the frequency response of the SOI devices at moderate doping and large electrode spacing is quite different from the measured and simulated response of the fabricated lateral p-i-n devices. The reason for this different behaviour is the buried junction in the devices in this work. The junction sets a boundary condition on minority carriers, resulting in vertical carrier diffusion (and as a consequence an autobiasing of the substrate as discussed in section 6.2). Due to the electric field line bending that must take place when the two depletion regions of the device meet, a largely lateral electric field pointing outwards from the n finger as expected, but there is also some lateral component of the electric field in the depletion region around and mostly importantly below the remaining QNR. This is quite apparent for example in Figure 6-6 where the edges of the depletion region at different biases may also be considered as electric potential contours, and therefore perpendicular to the electric field direction. Therefore, a carrier generated in the QNR instead of diffusing laterally for a certain distance may diffuse diagonally or vertically a shorter distance (therefore a shorter diffusion time constant), and then be carried by the lateral component of the electric field. The total time constant for this movement must be calculated numerically due to the two-dimensional problem statement and movement by drift in relatively low electric fields. This leads to the proposal of a novel device structure combining the advantages of the electrical isolation of SOI with the increased bandwidth at higher doping levels of the lateral pin with buried junction structure.

6.4 Novel device structure: buried junction on SOI

The SOI structure has the significant advantage of complete carrier isolation by the oxide layer as well as some enhancement of efficiency due to the reflection at the silicon-oxide interface. However, to achieve high bandwidths the SOI layer must be very lightly doped which may be undesir-

able in a CMOS process. Also, the bandwidth degrades significantly with increasing finger spacing for moderate doping levels. The lateral structure with buried isolation junction as fabricated and simulated in this work is able to overcome the lateral diffusion limit by allowing carriers to diffuse a short distance vertically, then drift laterally. However, it suffers from low sensitivity due to the relatively short active region depth that may be achieved without epitaxy (i.e. by implantation or diffusion alone) and from the hanging substrate effect when illuminated by a source with absorption length much longer than the active region depth. A combination of the two structures is shown in Figure 6-12 in a simulation of the depletion region edge as a function of reverse bias. The simulated structure is a 1.75 μm thick SOI layer with 2 μm BOX. The SOI layer has a uniform n-type doping of 5×10^{15} cm^{-3} and an implanted p-type region with net doping concentration of 1×10^{16} cm^{-3} and a junction depth of 0.95 μm . The fingers are n+ and p+ implants approximately 0.4 μm deep, as in the previous devices.

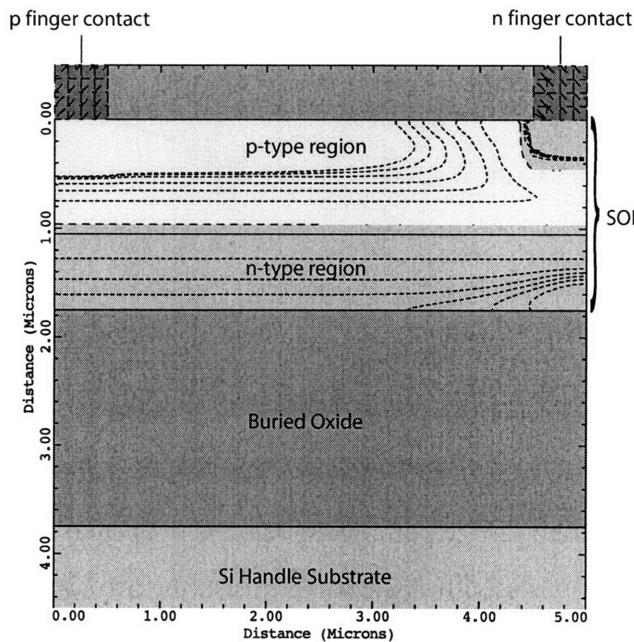
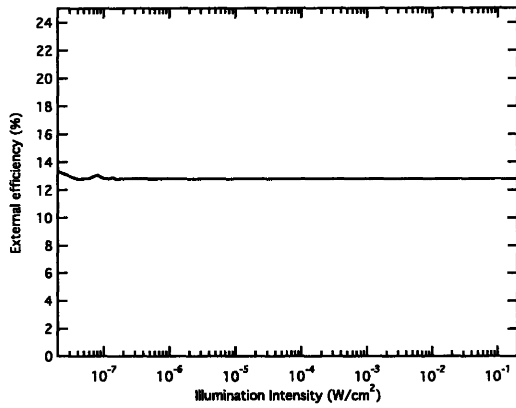


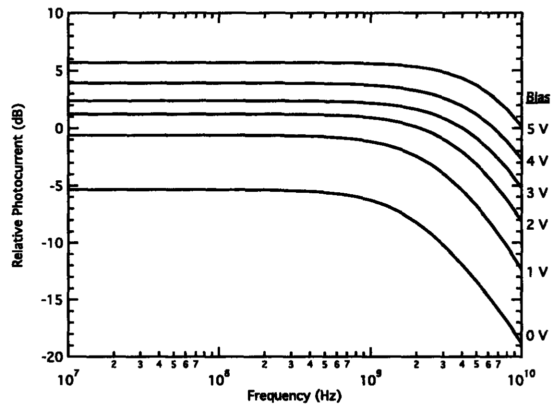
Figure 6-12: New device structure with a simulation of the depletion region edge (dotted lines) as a function of reverse bias.

For a low reverse bias, the unimplanted part of the SOI layer is depleted and the collection efficiency is very high, similar to the high resistivity SOI devices. Likewise, the efficiency does not change with illumination intensity. Typical frequency response curves are shown in Figure 6-13b (for a 100 μm diameter device with 2 μm finger spacing) demonstrating isolation of the carriers

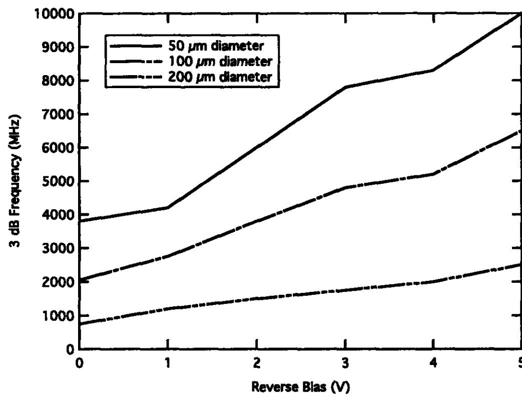
generated in the substrate.



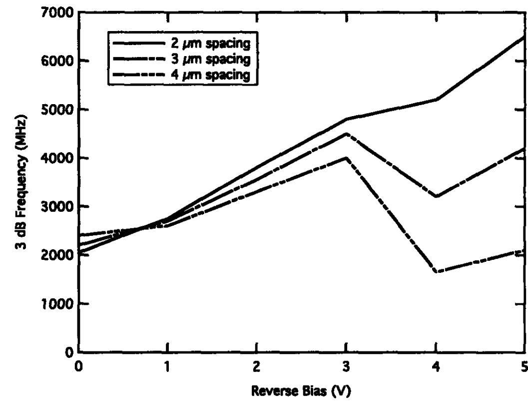
(a) DC Responsivity



(b) Typical Frequency Response, 100 μm diameter, 2 μm finger spacing



(c) 2 μm spacing, various device diameters



(d) 100 μm device diameter, various finger spacings

Figure 6-13: Simulation results of novel proposed buried junction SOI detectors.

The bandwidth for different device parameters may be extracted. For 2 μm spacing devices, the bandwidth as expected varies with device area due to the increasing RC time constant. For the 50 μm diameter devices, the RC bandwidth is greater than the final simulated bandwidth meaning that the devices are limited by drift and diffusion processes. In that case, the bandwidth reaches very high values approaching 10 GHz at 5 V reverse bias. This is over a factor of 2 higher bandwidth than that of SOI devices on high resistivity substrates. The change in bandwidth as a function of reverse bias includes an unexpected behaviour at higher biases for spacings of 3 and 4 μm. The frequency response develops a low-frequency plateau which decays somewhat and then forms another plateau before the ultimate decay at high frequencies. The net effect is a lower 3

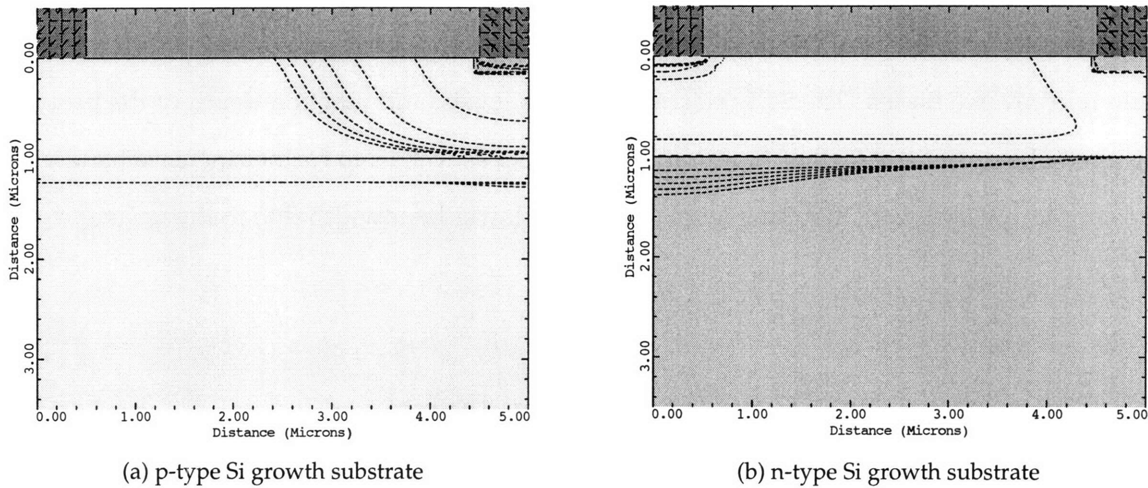
dB bandwidth. The curve shape suggests a sum of fast and slow current components, with the slow component slightly larger (approximately 1dB) than the fast component. Likely due to the exact geometry and shape of the depletion region, these components are found in different ratios at different reverse biases. The SOI thickness, doping levels, and junction depth of the proposed device may be optimised further to minimise this effect as well as to further increase bandwidth. For example, a very low SOI doping level results in a lower bandwidth due to the weaker electric field.

6.5 Germanium-on-silicon lateral p-i-n devices

Lateral p-i-n devices were simulated using the structure of epitaxial germanium on silicon as described in section 4.3. The simulation results are not directly comparable to measurements due to the apparent poor quality and/or high doping concentration of the grown germanium films, however they do allow for the evaluation of the general design as in the case of the previously simulated devices. In particular, there is a question arising from the best-case estimate in section 4.1.1, of whether the germanium devices are capable of reaching 10 Gbps operation at the desired 200 μm device diameter. Also, due to the band discontinuities at the junction of the germanium epilayer (generally unintentionally p-type) and the growth substrate (either p or n), it is instructive to examine the difference in device performance depending on the growth substrate.

In the case of a p-type substrate and assuming a p-type epilayer (in the simulations with a doping concentration of 1×10^{15} cm^{-3}), the band discontinuity results in a hole accumulation region on the Ge side of the interface and a depletion region on the silicon side of the interface (similarly as in a p-p+ homojunction, but increased by the band offset). For an n-type substrate, the heterojunction is Type 2 with a depletion region in the Ge epilayer. The consequence of these two alignments is visible in the simulation of the depletion region edge with increasing reverse bias, Figure 6-14. The sample with the p-type substrate shows an interaction with the hole accumulation region at the epilayer-substrate junction that results in some undepleted region near the interface. At 5 V reverse bias in the 4 μm spacing example, there is a substantial distance of QNR between the electrodes. The sample with the n-type substrate on the other hand exhibits an interaction between the junction depletion region and n-finger depletion region that results in very fast depletion of

the entire inter-finger space.



(a) p-type Si growth substrate

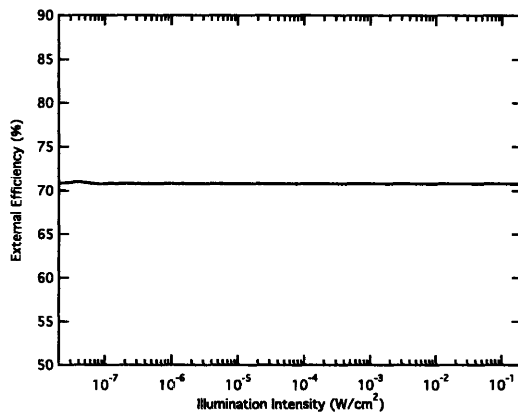
(b) n-type Si growth substrate

Figure 6-14: Simulation of depletion region edge (dotted line) in the lightly p-type Ge epilayer for increasing reverse bias. The depletion of the entire inter-finger space occurs at approximately 1 V reverse bias for the n-type substrate case.

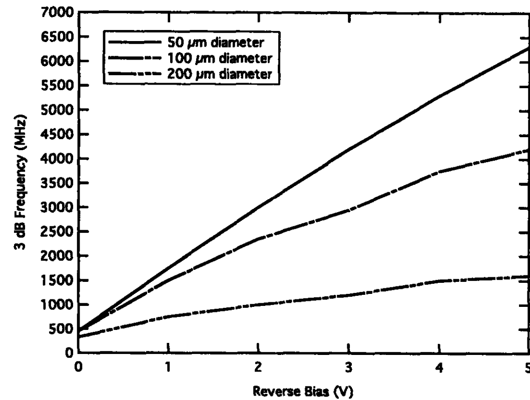
During the Ge lot 2 fabrication, both p and n-type silicon growth substrates were used. Due to the problems associated with blanket metal annealing, only devices on p-type substrates were successfully fabricated. This structure is therefore simulated first, using the nominal 800 nm germanium thickness of the fabricated lot 2 devices. The efficiency and bandwidth characteristics are shown in Figure 6-15.

The normalised external efficiency is 71%. To translate to actual device efficiency this value is multiplied by the metal fill factor of the electrodes, equalling over 45% for a 2 μm spacing device with 1 μm wide electrodes. This figure may be improved further by optimisation of the dielectric stack thickness to reduce surface reflections.

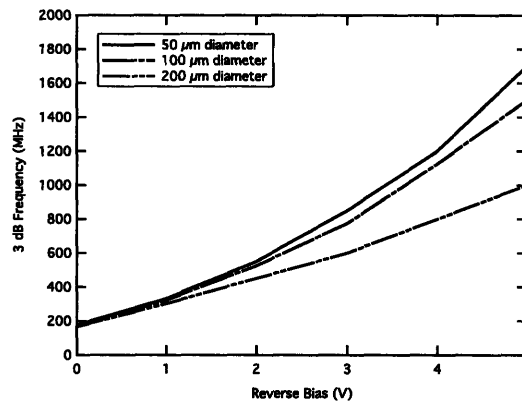
The bandwidth was examined for approximately circular devices with three diameters. The 2 μm spacing devices exhibit low bandwidth at zero bias, and a similar value for all three device sizes, indicating a carrier diffusion time limit due to the large extent of the QNR at zero bias. Bandwidth increases with reverse bias at different rates, indicating that bandwidth is RC-dominated for the larger 2 μm devices beyond the low bias regime. The smallest devices at 5 V bias have an RC time constant (calculated based on capacitance simulation and assuming a 50 Ω load) that corresponds to a larger bandwidth than the simulated value, therefore it is assumed that they are



(a) Normalised external efficiency



(b) 2 μm finger spacing, varying device diameter



(c) 3 μm finger spacing, varying device diameter

Figure 6-15: Simulation of lateral Ge-on-Si devices with p- epilayers (10^{15} cm^{-3}) on p-type Si substrates (10^{16} cm^{-3}) .

limited by diffusion and drift time. This intrinsic limit is quite high, approaching 6.5 GHz. The 200 μm diameter devices are dominated by RC resulting in a relatively modest 1.5 GHz, and the 100 μm devices are dominated by a combination of the two. In fact by using the 50μm device bandwidth, scaling the 200 μm device RC bandwidth to the 100 μm size, and using the "risetimes add quadratically" rule one estimates a bandwidth of approximately 4 GHz which is 5% from the simulated value.

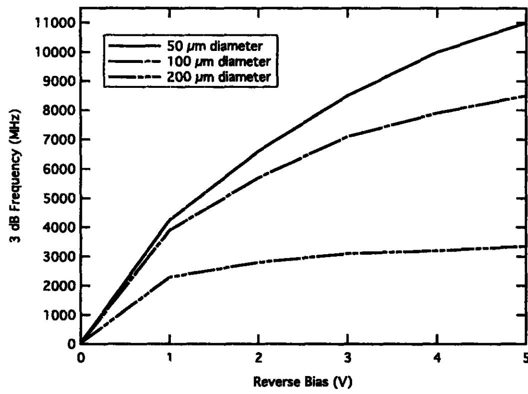
The 3 μm devices are more strongly dominated by diffusion time constant. The 50 and 100 μm diameter devices have similar bandwidths indicating the weak role of the RC time constant. The 200 μm devices, despite having a lower capacitance than the 2 μm spacing device for the same area, also has a lower bandwidth due to the additional effect of diffusion time constant

arising from the large QNR diffusion distance. It is apparent that the devices on p-type silicon substrates are not suitable to meet the goal of 10 Gbps speed in a 200 μm diameter device. Further simulations were performed to determine the sensitivity of the bandwidth to lower germanium region doping. For example a 1×10^{14} cm^{-3} doping was assumed (likely unrealistic from an epitaxy point of view) and it was found that bandwidth improved less than expected. Therefore, the interaction with the hole accumulation layer at the heterojunction is a performance limiter for the germanium devices.

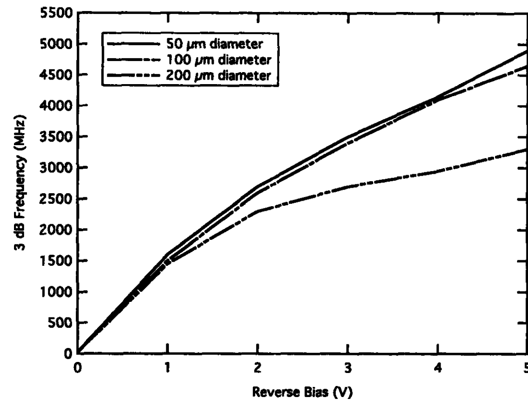
Devices assuming an n-type substrate were simulated, with a higher bandwidth expected due to the full depletion of the inter-finger region. The results for different device diameters and finger spacings are shown in Figure 6-16. The external efficiency of the devices is the same as for the p-type substrates. Since the region between the electrodes is apparently fully depleted, the limiting bandwidth factors in the case of these devices are RC time constant and drift time. The latter is dependent on the magnitude of the electric field if the carrier velocity is below the saturation velocity.

The 2 μm spacing devices show the familiar relationship of the previous devices where the 50 μm diameter bandwidth is dominated by delay time, in this case drift delay. The delay decreases with increasing bias as the field in the depletion region increases, with the rate of increase slowing down with increasing bias as the field approaches the velocity saturation condition. The 200 μm diameter devices are dominated by capacitance, but with a bandwidth twice as high as in the case of the devices on p-type substrates. The effect of capacitance is clearly seen as the bandwidth does not substantially increase beyond 1 V reverse bias when the entire region is depleted, i.e. following the discussion of p-i-n devices in section 2.3.1. The 100 μm diameter device bandwidth at 5 V may once again be derived similarly as in the previous devices using the quadratic sum of time constants method indicating a contribution of both effects to the final bandwidth.

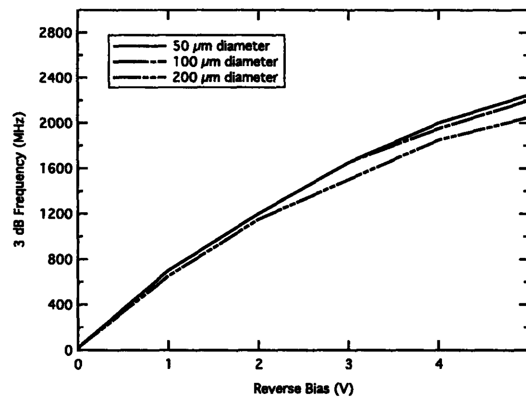
The 3 μm spacing devices are more strongly dominated by drift time. The 50 and 100 μm devices have essentially the same bandwidth behaviour with reverse bias, while the 200 μm devices feel a contribution from capacitance, resulting in approximately the same bandwidth as in the case of the 2 μm spacing devices. Finally, the 4 μm devices are entirely dominated by drift delay as all three device sizes show the same behaviour. The resultant bandwidth is relatively low and not influenced significantly by capacitance, whose time constant in the case of the 200 μm diameter



(a) 2 μm finger spacing, varying device diameter



(b) 3 μm finger spacing, varying device diameter



(c) 4 μm finger spacing, varying device diameter

Figure 6-16: Simulation of lateral Ge-on-Si devices with p- epilayers (10^{15} cm^{-3}) on n-type Si substrates (10^{16} cm^{-3}). Three different finger spacings exhibit different degrees of RC and drift time influence on bandwidth.

devices is approximately a factor of 2 above the drift delay time constant. As a final check, one may calculate the carrier velocity using equation 2.18, and the result for 4 V reverse bias is approximately one third of the saturation velocity (limited by holes), consistent with the field strengths present in the device.

Clearly the germanium epitaxial device on n-type silicon substrates offer a superior bandwidth to that of devices on p-type substrates. Based on the simulations, a 200 μm diameter device is believed to be suitable only up to 5 Gbps applications based on the achievable 3-3.5 GHz bandwidths for 2 or 3 μm electrode spacing devices. To meet a 10 Gbps specification, the 3 μm and larger spacing devices are not suitable. The 2 μm device area should be reduced approximately by a factor of 0.5, suggesting an acceptable device diameter of approximately 150 μm . Alternately, there is likely

a more optimal design point in the vicinity of 2 μm spacing that my permit a slightly larger diameter or a slightly larger sensitivity-bandwidth product for a given area. These two design points are not coincident, therefore one would have to take into account the sensitivity-bandwidth-mode overlap product to determine the optimal device geometry.

The simulation results also suggest that it is more desirable to fabricate devices on n-type substrates (or n-type implanted regions) than on ultra-thin SOI as in [40]. In the latter case, there is no interaction between the two depletion regions that expands the extent of the depletion region at low bias. In cases where the possible interdiffusion of the substrate and epilayer are not a concern, for example the 850 nm regime explored in this work, the SOI substrate adds cost and complexity to the process while in fact reducing device performance. This is shown in Figure 6-17 for 2 μm spacing devices. The effect is more pronounced for larger electrode spacing.

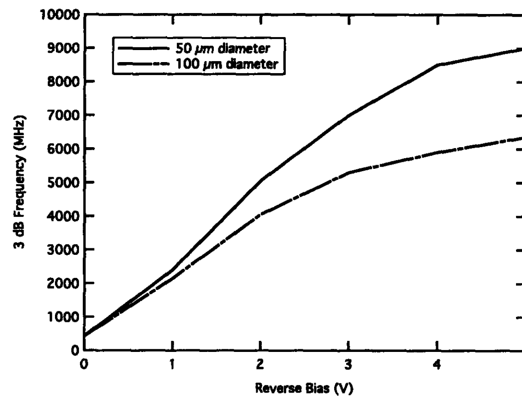


Figure 6-17: Simulation of lateral Ge-on-I devices with p- epilayers (10^{15} cm^{-3}). The finger spacing is 2 μm. These curves may be directly compared with the 2 μm finger spacing curves of Figure 6-16.

Chapter 7

Photodetector and System Design Summary

The measurement of chapter 5 and simulations of chapter 6 indicate a variety of bandwidths and sensitivities depending on the precise device structure and device material. In particular it was shown by simulation that similar bandwidths may be achieved for silicon and germanium devices of similar geometries. In fact, it was shown in section 4.1.1 that the maximum possible bandwidth of silicon devices is higher than that of germanium devices due to a lower dielectric coefficient (resulting in lower capacitance) and higher saturation velocity (resulting in shorter drift delay).

Silicon has however traditionally been relatively unattractive for high performance applications due to the low absorption coefficient of wavelengths of interest. Silicon of course exhibits no absorption in the longer-distance telecom wavelength regimes of 1310 and 1550 nm, and weak absorption in the common VCSEL wavelengths of 780 and 850 nm. This results in carriers travelling a long distance by diffusion, and therefore a slow frequency response as discussed in section 2.3.3. In order to counteract this problem, variety of structures have been proposed as discussed in section 1.2 and presented in the body of this work. These structures may result in devices with very high bandwidth, but with low responsivity since the bulk of carriers generated far from the junction are not collected such that they do not contribute a slow component to the frequency response. A weak signal may be amplified, but such an operation introduces noise. As discussed

in section 3.2.1, the important system parameter in a digital communication system is the bit error rate which may be directly related to the signal to noise ratio (SNR) by equation 3.4. The BER of interest in modern systems is BER $< 10^{-12}$, corresponding to SNR $= 7$. The SNR is related to the noise and the bandwidth of the system.

System considerations

The communication systems of interest in this work operate over multimode connections, with the possibility for use of silicon as a photodetector material. Therefore, the wavelengths of interest are in the area of 650 nm (RCLED, recently VCSEL), 780 nm (VCSEL), and 850 nm (VCSEL). For commercial applications, light intensity in the system is limited by eye safety considerations to operation at so-called Class 1 or Class 1M levels [95]. Class 1M signifies a system that includes a monitoring component that reduces power to Class 1 levels if a break in the transmission system is detected. The relevant maximum admissible powers for the different wavelengths are shown in table 7.1. Class 1 is preferred for two reasons: the additional circuitry required to implement the safety circuit in Class 1M, and the necessity for more than one source in order to reach Class 1M power levels (with a typical VCSEL providing less than 1 mW of fibre-coupled power).

Wavelength (nm)	Class 1	Class 1M
633	-4.1 dBm	5.9 dBm
780	-2.5 dBm	7.5 dBm
850	-1.1 dBm	8.9 dBm

Table 7.1: Eye safety levels for Class 1 and Class 1M systems in mulitmode optical fibre (diameter $< 150\ \mu$m) [95].

The loss budget of the transmission system is determined by fibre coupling, connector/adapter losses, optical fibre loss, and margin for parameter drift due, for example, to ageing and temperature variation. This work was performed in cooperation with a commercial sponsor, and as such the author was involved in discussions of a proprietary nature regarding the definition of the loss budget for the sponsor's project. In order to avoid any release of proprietary information, the determination of the loss budget will not be discussed further in this document. A rule of thumb of 13 dB loss for a short-distance link will be assumed for the system loss (i.e. 95% loss from optical source to TIA input). An equivalent distance glass multimode fibre link would have a several dB

lower loss budget owing to lower fibre loss and more stringent connector and coupling tolerances. For discussion in this section, a loss budget of 10 dB is assumed.

7.1 Ideal system - Quantum Limit of Detection

A light flux incident onto a photodiode generates carrier pairs in independent, random events. This results in a shot noise where the probability of error may be described by the Poisson probability distribution. This relationship is derived precisely in [96]. The minimum mean received power in the quantum limit for BER $< 10^{-12}$ is given by

$$\overline{\Phi}_R > \frac{13.5\, \mathcal{E}_{\mathrm{ph}} B}{\eta} \tag{7.1}$$

where B is the bitrate, η is the collection/conversion efficiency, and $\mathcal{E}_{\mathrm{ph}}$ is the photon energy. For $\eta = 1$ and $\lambda = 850\mathrm{nm}$, $\overline{\Phi}_R > 3.1\mathrm{pW/(Mb/s)}$, with an appropriate scaling for other wavelengths. This relationship is shown along with expected incident power (based on eye safety limits and loss budget) in Figure 7-1. The system operating range is the area of the graph below the eye safety limits (horizontal lines) and above the quantum detection limit (diagonal lines). It is apparent that the expected received power is far above the quantum detection limit up to 10 Gb/s. A BER of 10^{-12} is maintained for all wavelengths for efficiencies as low as 0.2%.

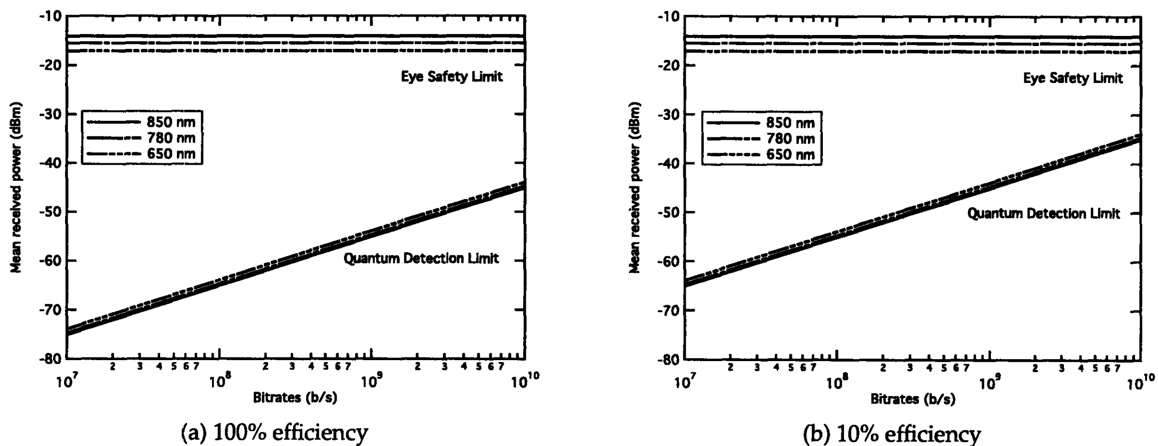


Figure 7-1: Quantum detection limits compared to maximum incident power for different wavelengths and efficiencies.

7.2 Realistic System - Detector and Receiver Noise

As discussed in section 2.5, the shot noise of the detector (caused by dark current) is generally much smaller than the noise of the transimpedance amplifier. In cases of devices with very high dark currents, a more serious problem is the dynamic range of the amplifier. Reports of such devices generally include a capacitor coupling of the device to the amplifier/receiver (for example [43]) in order to reject the large DC current. However, this is not a practical solution for the integrated transceivers in future applications, as coupling capacitors consume a large amount of IC die area.

It is not possible to make a general statement regarding the noise behaviour of transimpedance amplifiers. The circuits may be optimised for example to reach a particularly high bitrate or for low noise behaviour at a low bitrate, and the performance is a strong function of the fabrication and device process. Additionally, performance of the system varies with the photodetector capacitance. This refers both to the frequency-domain behaviour (more precisely the bandwidth or stability as discussed in sections 2.4 and 2.4.1 as well as the noise behaviour. The calculation will be not repeated here, but a simplified treatment of TIA noise in [96] includes the device capacitance in one of the terms of the denominator in the expression for the signal to noise ratio, and TIA parts are specified with a nominal and/or maximum detector capacitance for stability and noise concerns.

As a design example, one may consider the state of the art at time of writing Analog Devices ADN2821 transimpedance amplifier [63]. The device is designed and fabrication in a SiGe BiCMOS process, specified to operate at in a transceiver system at 11.1 Gb/s (typical TIA bandwidth 8.5 GHz). It is possible to compare the expected received intensities to the specified noise perfomance of the part as was shown above for the quantum limit. In this case, $Q = 7$ is applied for an appropriate BER. To convert from bandwidth to bitrate, a conversion factor of 4/3 or 1.5 is used (depending on the details of the system) as discussed in section 3.2.2. The curves of Figure 7-2 also includes a 1 dB noise penalty for a typical 10 dB modulation depth.

From this analysis, the difficulty of designing a system operating at 10 Gb/s is apparent. For an 850 nm light source system, an external efficiency of 25% is required for a POF system. A glass MMF system would require slightly lower efficiency due to its several dB lower loss budget.

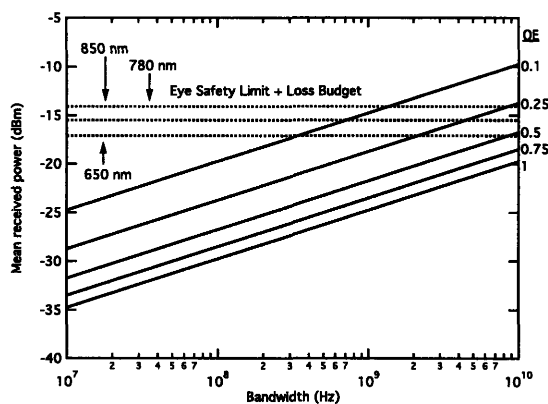


Figure 7-2: Detection limits of an ADN2821 TIA assuming $Q = 7$ and a 1 dB penalty for a typical extinction ratio. The dark lines are varying system external quantum efficiency curves as noted on the right of the figure.

The required efficiency rises as the system wavelength is reduced due to eye safety standards. Additionally a system bandwidth of above 6.5 GHz is required. In practice, this requires one or both of the photodetector and TIA to have a higher bandwidth, such that the combination of the two system functions will yield the required composite bandwidth.

Though the eye safety power limit is lower by 3 dB for 650 nm light as opposed to 850 nm light, the absorption coefficient is higher by a factor of 5 in Si. For the devices presented in this work, the external efficiency is most strongly influenced by the proportion of the active region thickness to the absorption distance of the wavelength of interest in the photodetector material (Si, SOI, Ge, etc.). Therefore, moving to 650 nm may open the path to a wider range sensitivities and bitrates for silicon OEICs.

Finally, this work concentrated strongly on the issue of large area detectors in the context of POF applications. The overlap of the photodetector and the fibre core of course directly relates to the external quantum efficiency since light may only be collected in the overlapping regions. The overlap may be approximated by simply comparing fibre core and device areas. This is not precisely correct since more of the light generally travels in the mode near the centre of the core, particularly in the case of graded index fibre. In that case, the mode profile must be calculated numerically, particularly for such highly multimode fibre as POF, or measured directly using a near-field technique. On the other hand, higher-order modes may be excited by non-optimal source-fibre coupling or strong fibre bending, therefore the ratio of areas is a simple and useful

approximation. For example, the modal overlap of 200 μm POF and the 50 $\times$ 50 μm SOI detectors discussed earlier is approximately 8%. Clearly, even if they exhibitied 100% light collection efficiency, these detectors are not suitable for multiple Gb/s systems over such a large diameter fibre.

7.3 Detector Design for Particular Applications

This section will present the design decisions required to implement a photodetector for several particular applications. The initial assumptions are Class 1 operation, loss budget for POF and glass MMF as described above, and TIA bandwidth and noise characteristics as described above for the ADN2821. A further assumption is that device capacitance does not change the bandwidth or noise characteristics. In practice this is a false assumption, but necessary in order to avoid introducing too many variables into the discussion. A full review of optimal TIA designs for different applications is beyond the scope of this work.

The devices presented in this work are intended for integrated applications. The devices with buried isolation junctions or fabricated on SOI substrates (with or without buried junction) should require at most minor fabrication process changes consisting of one or two additional implantation steps and, for the SOI devices, a more expensive substrate. The Ge detectors is intended for integration, though at time or writing there are many challenges remaining in order to integrate the epitaxial material into the front end of an IC process, many of which were discussed in section 4.3 and appendix A. For Si or SOI devices, the designs, fabricated detectors, and simulations were focused around material regions with doping levels on the order of $10^{15} - 10^{16}$ cm^{-3} in order to maintain compatibility with typical doping levels of substrates and n and p wells in CMOS processes. This was also one of the motivations behind the novel buried junction on SOI structure, as contrasted with the literature results of devices on very high resistivity SOI. Likewise Ge lateral devices were chosen for study partly because of their planar structure that should simplify process integration as compared with vertical devices. The fabrication procedure of the Ge devices specifically avoided unrealistic process steps, such as metal deposition by lift-off. Such considerations should play a role during device design, because too many required process modifications will result in a device or system that does not benefit from the economies of the standard CMOS

process.

Two design criteria must be met: external quantum efficiency and bandwidth. The first is a function of the material absorption coefficient and the active region thickness. It is also a function of the metal fill factor of the contact electrodes. The second is a function of doping, finger spacing, active region depth, active region doping, and electric field strength. The extent of the depletion region between the fingers and the magnitude of the electric field may not be simply calculated, and must be extracted from a numerical two-dimensional solution. This is particularly true in the case of a device with a buried junction which involves an interaction of two depletion regions, as was seen dramatically in the examination of the optimum doping type for Ge-on-Si growth in section 6.5. Finally, the bandwidths discussed thus far in this work have assumed a 50 Ω load. As discussed in section 2.4, the actual load seen by the detector may be less than 50 Ω, depending on the details of the TIA. The standard load is used in order to provide a standard reference for published results and in order to take advantage of standard RF measurement tools. However, a device that is RC-limited in a measurement apparatus may behave slightly differently connected to a particular TIA.

7.3.1 10 Gb/s over glass MMF

The specification defines the source device: VCSELs operating at 850 nm. VCSELs at 780 nm or 650 nm are not readily commerically available with an appropriate lifetime and/or bandwidth for a commercial application at this bitrate. The glass MMF core diameter is typically 62.5 μm with 50 μm also available. This implies a detector area of approximately 2500-3000 μm^2. Using the relationships introduced earlier, a system bandwidth above 6.5 GHz is required. The design space is shown in Figure 7-3.

An examination of the design space indicates a minimum external quantum efficiency of 0.1. This efficiency may be achieved by a number of devices studied in this work: high resistivity SOI [18], the novel buried junction SOI structure of section 6.4 for finger spacings above approximately 1.5 μm, and the Ge-on-Si structures fabricated in this work and simulated in section 6.5. The silicon devices with buried junction do not meet the quantum efficiency specification due to the thickness of the active region. An active region with a depth of 2 μm is not easily fabricated by implantation

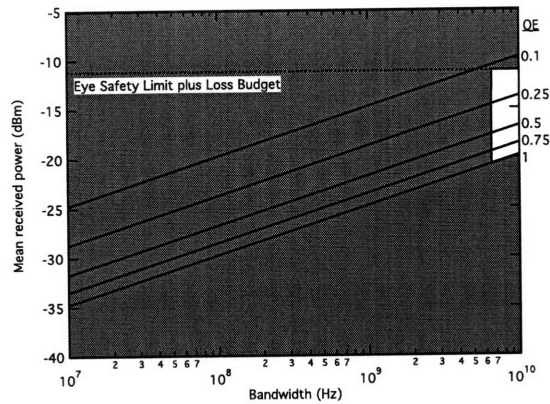


Figure 7-3: Design space (clear area) for a 10 Gb/s optical link over glass MMF.

or diffusion, requiring epitaxial growth.

The reported and simulated bandwidth of the high resistivity SOI structures is approximately 3 GHz for 2 μm spacing, limited by a weak electric field far from the surface that results in a significant drift delay. A narrowing of the finger spacing would increase capacitance without substantially increasing the magnitude of the electric field. In fact, as shown in section 6.3, increasing the finger spacing to 3 μm results in a small increase in device bandwidth due to decreased capacitance. In any case, published reports suggest that a 6.5 GHz bandwidth is difficult to achieve while maintaining relatively high responsivity (i.e. an SOI layer with thickness above 2.5 μm).

Simulated buried junction SOI as simulated in section 6.4 exhibits a bandwidth approaching 10 GHz at 5 V reverse bias for a 50 μm diameter devices with 2 μm finger spacing. The responsivity is at the limit of acceptability due to the metal fill factor. Since the simulated bandwidth is above the minimum requirement for a modest reverse bias, some numerical optimisation may be used to increase the responsivity while maintaining an acceptable bandwidth. This is accomplished by increasing the SOI thickness and/or increasing the finger spacing.

Germanium devices with 50 μm diameter and 2 μm finger spacing on n-type silicon exceed both the responsivity and bandwidth requirements. As shown in section 6.5, devices on p-type silicon have inferior frequency response due to interaction of the depletion region with the substrate heterojunction. The excess bandwidth may be traded for larger device area, enabling a more robust and less expensive alignment and assembly of the transceivers, or higher responsivity by increasing finger spacing (and decreasing metal fill factor). The higher responsivity may then be

applied to a more liberal loss budget or a redesign of the TIA in order to reduce IC area or power consumption.

7.3.2 10 Gb/s over POF with 100 μm core diameter

The bitrate specification again defines the source device: VCSELs operating at 850 nm. The design space is similar to the the previous diagram above, but with a high loss budget, as shown in Figure 7-4. In this case, a larger detector is required for an optical mode match. The device areas of the previous example would result in external quantum efficiencies on the order of 3-5%, far below the required value. The choice of device for appropriate efficiency is the same as in the previous example, assuming a larger device area.

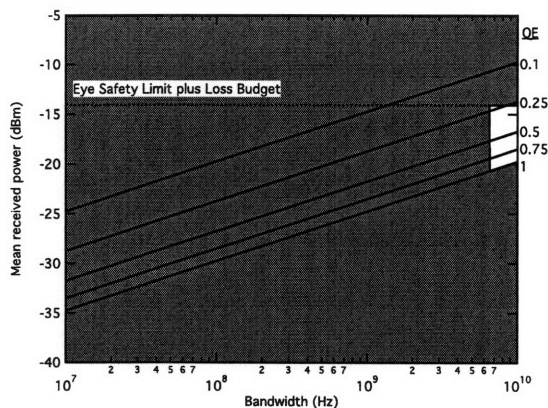


Figure 7-4: Design space (clear area) for a 10 Gb/s optical link over 100 μm diameter POF.

The high resistivity SOI structures with larger area are limited by a combination of capacitance and transit time delay, and are not capable of meeting the bandwidth specification. As shown in Figure 6-13, and buried junction SOI device with 100 μm diameter and 2 μm spacing is capable of 6.5 GHz bandwidth. However, the high loss budget requires a 3 dB higher efficiency than in the glass fibre case. Numerical optimisation of the structure would yield some improvement but, would is not expected to yield a structure with the appropriate efficiency while maintaining the required bandwidth.

Germanium devices with 100 μm diameter and 2 μm finger spacing on n-type silicon exceed both the responsivity and bandwidth requirements (0.5 external efficiency and 8.5 GHz bandwidth) as shown in Figure 6-16. Germanium on a p-type substrate is not acceptable, with band-

width approximately 50% below the required value. Again, with the device structure far exceeding specifications, optimisation may be performed on the device and/or the system as a whole.

7.3.3 1.5 Gb/s over POF with 200 μm core diameter

Such a system enables a wider variety of sources: VCSELs at 850, 780, and recently 650 nm. This section will concentrate on the latter, as it is approximately this wavelength and POF diameter that may be prevalent in automotive applications in later generations of the MOST standard as discussed in section 1.1. Red sources are desired in these applications due to the ease of identifying connection problems by automobile mechanics (as opposed to the IR wavelengths, requiring more extensive training, IR viewing cards, etc.). The incident light limit is lower than in the previous design space. The device diameter is nominally 200 μm for good modal overlap.

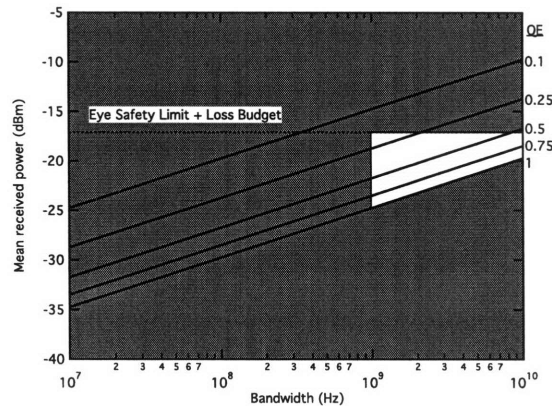


Figure 7-5: Design space (clear area) for a 1.5 Gb/s optical link over 200 μm diameter POF using 650 nm VCSEL sources.

The bandwidth and efficiency requirements of this system are less severe than in the previous examples. In particular, all of the silicon devices (both and bulk with isolation junction) are able to meet the external efficiency and bandwidth specifications as shown in chapter 5 In the case o the isolation junction devices, efficiency is at the edge of the design space and may be improved further by using a deeper active region. This may however not be practical for a processing point of view.

Both varieties of SOI devices offer sufficient efficiency and bandwidth for this application. In fact, there is a range of acceptable thickness and doping levels that satisfy the specification,

allowing for easier integration with an existing fabrication process.

In this application, germanium on silicon devices can easily satisfy all requirements. However, there may be a question of whether the additional cost and complication to integrate germanium into the fabrication process is warranted for this application. The use of SOI substrates with either a bulk or buried junction design offer more than sufficient performance with much smaller process disruptions.

7.4 Detector and System Design Charts

The discussion of the previous section identified three useful application areas, derived the available design space for each application (constrained by Class I operation limits, system bandwidth, and signal to noise ratio), and based on the above suggested possible detector designs for the different applications. In this section, the design space is rewritten as a minimum sensitivity for a given bandwidth, and the different device designs of this work are plotted in the design space. This exercise serves two purposes: to graphically demonstrate the advantage of lateral over vertical detector designs in the large-area regime, and to graphically demonstrate the usefulness (or lack thereof) of different detector materials for different applications

One parameter in this discussion is the minimum sensitivity at a given bandwidth. The design space shown in Figure 7-2 indicates an upper power limit defined by Class I limits and system loss budget, and a lower power limit defined by SNR and quantum efficiency. The ratio of the SNR line for an efficiency of 1 to the maximum power line defines the minimum sensitivity at a given bandwidth. The minimum sensitivity of course increases with increasing bandwidth as seen in Figure 7-2. The detector system efficiency is defined by the device external efficiency and the fibre-detector modal overlap. Dividing the detector system efficiency by the minimum efficiency yields the figure of merit for this discussion: the system efficiency factor (also called margin below). A value above 1 indicates that the SNR requirement (to achieve a particular BER) for a given bandwidth is met, and a value below 1 indicates a higher than desired BER. Obviously, for a constant device efficiency this factor decreases with increasing bandwidth due to increasing noise as discussed above. For a successful system design, the detector must satisfy the conditions of minimum bandwidth as discussed earlier, and efficiency margin factor of greater than one. In

the figures to follow, the design space with insufficient SNR (i.e. a forbidden region for the design) is indicated by a grey background.

All assumptions listed in the previous section apply to this discussion as they are implicit in the derivation of the minimum efficiency specification. Several devices are investigated: lateral and vertical p-i-n detectors of varying diameters, i region thicknesses, and material. The lateral detectors include the simulated germanium devices on n-type silicon substrates discussed in section 6.5, and the buried junction SOI devices discussed in section 6.4. The vertical detectors are specifically: silicon vertical p-i-n with a 1 μm thick i region, silicon vertical p-i-n with a 5 μm thick i region, and a germanium vertical p-i-n with a 1 μm thick i region. The thick silicon p-i-n is essentially the detector discussed in section 1.2.3 and the germanium p-i-n is essentially the detector discussed in section 1.3.1. Two important assumptions are made. First, only light absorption in the i region is considered. In other words slowly diffusing carriers are not considered in the discussion of the silicon vertical p-i-n devices. The impact of slowly diffusing carriers has been discussed at length throughout this work, and it should be obvious to the reader that this assumption is very liberal and serves to inflate the silicon device bandwidth significantly, particularly for the 850 nm regime. Second, a bias of 5 V is assumed. Therefore carriers in the silicon devices do not travel at their saturation velocity, and are limited by the hole drift velocity calculated from [48]. At this bias, the 5 μm thick Si vertical p-i-n devices are limited solely by drift transit time and not by capacitance. This is the reason behind the higher bias schemes developed, for example, in [29]. Also, the simulated germanium lateral devices and SOI with buried junction lateral devices are considered. Devices with diameters from 50 to 200 μm are considered.

The first design chart is shown in Figure 7-6. It illustrates communication systems operating at 850 nm over three types of fiber: 50 μm diameter glass MMF, 100 μm diameter POF, and 200 μm diameter POF. There is a visible offset between the POF and MMF lines, due to the lower loss budget as discussed in the previous section. There is one very important point that the reader must keep in mind when considering the design chart. The plotted points indicate the device bandwidth and system margin factor for a given fibre core diameter, fibre material, wavelength, and device diameter. However, the bandwidth is a *maximum* specification. Therefore a device geometry that is within the forbidden region for a given bandwidth may be within the allowable design space for a lower bandwidth due to the higher minimum sensitivity as discussed above.

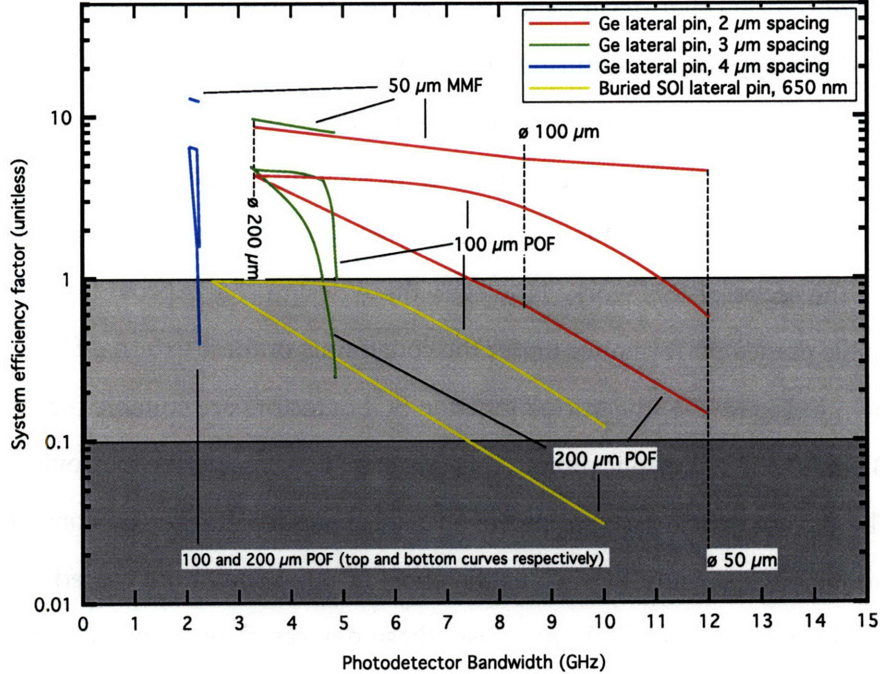


Figure 7-6: Design chart of lateral germanium and buried junction SOI devices. The germanium devices assume a 850 nm transmission system, and the SOI devices assume a 650 nm transmission system. A factor of 1 represents the class I limit, and a factor of 0.1 represents the class IM limit (also shown as a darker shade of grey).

The first important point in the figure is the rapidly decreasing efficiency factor as the device diameter decreases below the fibre core diameter. This is particularly apparent in the case of the 4 μm finger spacing device whose bandwidth is essentially unchanged with device area, but whose efficiency factor changes rapidly as the fibre-detector modal overlap decreases.

The germanium devices are suitable for over 10 Gb/s operation in various system configurations. The 200 μm diameter POF design space extends to approximately 7.5 GHz, sufficient for 10 Gb/s operation as predicted in section 6.5. For 100 μm diameter POF, the system efficiency factor is well above 1 in the regime 7-9 GHz that is relevant for 10 Gb/s operation. This confirms that there is a large design space of POF fibre core diameters and detector dimensions that may be used to optimise the 10 Gb/s system from the point of view of loss budget, chip area, coupling complexity, or any number of other parameters.

The larger germanium finger spacings result in lower bandwidths as shown earlier. The lower metal fill factor results in only somewhat higher external efficiencies. The decision to use such a

geometry is influenced by other design choices that are not capable of being represented in the figure, for example an application requiring only a 5 Gb/s bitrate, or a TIA architecture that benefits from a lower device capacitance due to noise or stability issues as discussed in section 2.4.1. As discussed many times in this work, the decision for a particular device geometry or design is governed not only by bandwidth and external efficiency, but also by external design considerations far outside the scope of this work. However, these design charts provide a mechanism for evaluating possible device performance under the constraints of these external factors.

As also shown in Figure 7-6, the buried junction SOI detectors are suitable for approximately 2 GHz (up to 3 Gb/s) 650 nm applications. This performance is superior to previously reported SOI results. The design space may be improved and expanded further by considering device optimisations to improve sensitivity, such as the effect of reflection at the buried oxide interface or the precise SOI thickness choice. In any case, these devices show significant promise in 650 nm, large diameter applications. Finally, as shown by the light and dark grey areas, even higher bitrates are possible by moving from a class I to a class IM system, due to a 10 dB increase in allowed source power.

Figure 7-6 presented the design space of devices discussed in this work, but it is instructive to also compare these devices to more common structures. Vertical pin structures as described at the beginning of this section are shown in Figure 7-7 using the same design space visualisation. The 1 μm silicon devices are not useful for POF applications, and may be applicable only in 1-2 GHz MMF applications. As noted earlier, these curves assume a complete blocking of all diffusing carriers generated in the substrate, and therefore the real silicon device performance would be much worse than shown unless the devices were to be fabricated on an electrically isolated (e.g. SOI) substrate. On the other hand, 5 μm silicon devices may be useful in a limited bandwidth regime. Since the device bandwidth is limited by drift transit time, the design space is essentially a thin vertical slice, i.e. the device area does not strongly influence the bandwidth. However, the external efficiency does scale with device area (due to modal overlap), hence the rapid drop-off of the curve for smaller device diameters and large POF cores.

Figure 7-7 also indicates that germanium vertical p-i-n devices are suitable for high-speed MMF and POF applications. Due to the high absorption coefficient of 850 nm light, glass MMF applications to approximately 10 GHz show a high design margin. For 100 μm diameter POF ap-

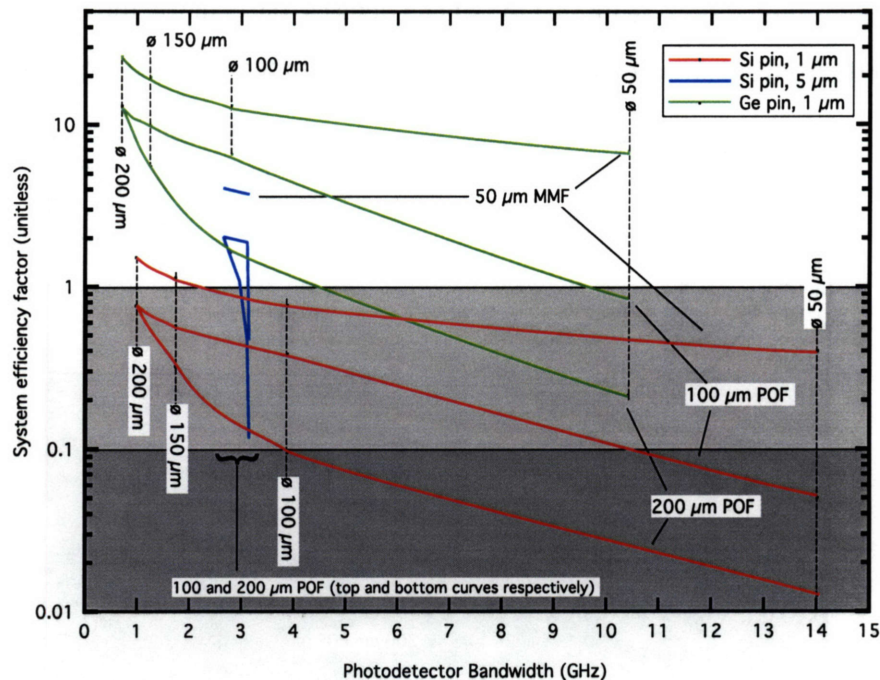


Figure 7-7: Design chart for an 850 nm data transmission system including different photodetector and fibre geometries. A factor of 1 represents the class 1 limit, and a factor of 0.1 represents the class 1M limit (also shown as a darker shade of grey).

plications, devices slightly larger than 50 μm diameter may provide up to 9-10 GHz performance at the limit of efficiency margin. Larger POF (200 μm diameter) applications are limited to below 4.5 GHz at the limit of efficiency margin.

The next system under consideration is a 650 nm communication system operating over POF with core diameter from 100 to 200 μm, as shown in Figure 7-8. A visible difference compared with the above system is the downwards shift of the germanium detector curves. There is no efficiency advantage realised by moving from 850 to 650 nm for a 1 μm thick germanium film, but there is a 3 dB eye safety limit penalty. Therefore, though the device bandwidths are unchanged, the applicable design space is much smaller. The silicon devices see a net advantage from the change of wavelength due to a much stronger absorption coefficient. This raises the silicon 1 μm devices into the allowable design space for bandwidth of 1-2 GHz. The silicon 5 μm devices also take advantage of the improved absorption coefficient, moving vertically to a higher efficiency factor. The 650 μm POF application is precisely the target of [28] as discussed in section 1.2.3. The detector bandwidth in this case is slightly higher than that of currently commercially available red

VCSELs. It should be noted once again that the silicon devices (particularly the 1 μm devices) may suffer significantly from substrate carrier diffusion without proper electrical isolation.

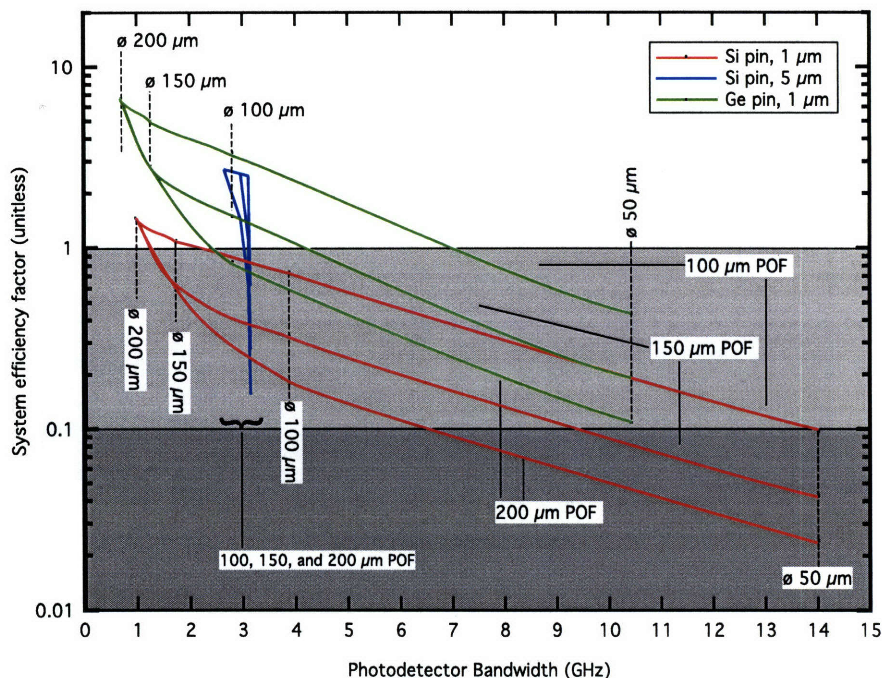


Figure 7-8: Design chart for an 650 nm data transmission system including different photodetector and fibre geometries. A factor of 1 represents the class I limit, and a factor of 0.1 represents the class IM limit (also shown as a darker shade of grey).

Finally, the curves of vertical and lateral germanium detectors may be presented together on one design chart in order to highlight the advantage of the lateral structure, as shown in Figure 7-9. There are two notable features in this design space. First, the sensitivity of the lateral devices is lower than that of the vertical devices due to the metal electrode shadowing of the lateral case. This is a traditional criticism of interdigitated detector structures. However, as seen from the design space, in both cases the device efficiency is high enough to meet the minimum system efficiency specification for a wide range of frequenies. The second notable feature is that for a given device area, the lateral photodetectors have a significantly higher bandwidth than their vertical counterparts due to a lower capacitance per unit area. Therefore, it is possible to use larger detectors in order to achieve a higher modal overlap and therefore a higher effective efficiency in large core diameter applications. This is illustrated by the MMF design (lateral detector curve below and to the right due to metal shadowing and lower capacitance), as compared with the

POF design (lateral pin curves above and to the right due to better modal overlap and lower capacitance).

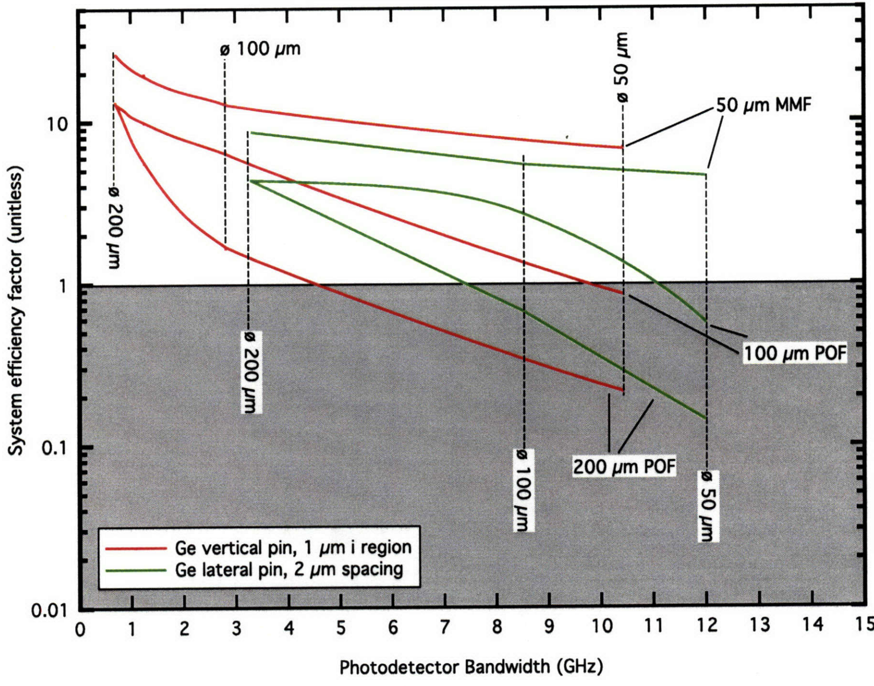


Figure 7-9: Design chart for an 850 nm data transmission system including germanium lateral and vertical detectors, and different fibre geometries. A factor of 1 represents the class I limit.

Chapter 8

Conclusions and Future Work

This work investigated the design, measurement, and fabrication of photodetectors with large active areas for applications in polymer optical fibre systems. Although the principle of operation of photodiodes is well known, most prior work in the field was directed towards achieving speed records or achieving a certain specification in a glass fibre system. The conventional wisdom states that to achieve high bandwidths one must operate in a regime where the entire i region of the p-i-n structure is depleted, and that there is a direct capacitance-transit time trade-off related to the thickness of the i region. This work has demonstrated a much richer set of device behaviours and a much wider design space that may be used to construct the optimum device for a large-area application.

8.1 Summary

Optoelectronic frequency response measurement apparatus

An apparatus was designed an implemented by the author for the direct measurement of the frequency-domain response of optoelectronic devices. As described in this work, the apparatus was used to measure all of the fabricated photodetectors. Through a simple change consisting of the reconnection of certain elements and the incorporation of some additional micro-positioning instruments, the apparatus may be used to measure the response of modulators or active devices.

The effective measurement range of the apparatus is 10 MHz-8 GHz. This range is defined by the low-frequency limit of the network analyser, and on the high-frequency side by the sources

(VCSEL or external lithium niobate modulator) and by the reference photodetector and TIA used to calibrate the sources. The bandwidth limit of the network analyser is 20 GHz, which could be achieved for the entire apparatus with an appropriate choice of fast reference detector and sources. Depending on the precise device in question, the maximum repeatable measurable bandwidth may be slightly lower depending on the signal level, or more precisely the signal to noise ratio at bandwidths of interest.

Also, an impulse response and eye diagram measurement system was assembled or future use with integrated devices in this and other research projects. A discussion of the various methods of frequency response measurement, and the relative advantages and disadvantages of each method, was given in chapter 3.

Lateral GaAs p-i-n photodetectors

Lateral photodetectors were fabricated in a standard commercial VLSI process, characterised, and simulated. The devices achieved very high quantum efficiency and bandwidths of 4.5 GHz at 7 V reverse bias, for a device area of approximately 75 μm diameter. These devices are good candidates to replace the traditional GaAs MSM photodetectors in integrated applications, as they may be integrated into a standard MESFET process with no additional mask steps, whereas an integrated MSM typically requires at least one additional mask to define an n- implant region in the semi-insulating substrate.

Lateral silicon p-i-n photodetectors

Silicon lateral photodetectors were fabricated and characterised. These devices represent the most precise alignment tolerance devices fabricated in the Microsystems Technology Laboratories in recent years, and as such exposed many weaknesses of the fabrication equipment and necessitated the development of a variety of new processing techniques in order to achieve reasonable device yield.

The devices were characterised using the apparatus described above. Device diameters of 50-200 μm and finger spacings of 1-4 μm were investigated. The best 200 μm diameter devices had a bandwidth of over 2 GHz at only 5 V reverse bias. The best 50 and 100 μm devices had

a bandwidth of over 4 GHz at only 5 V reverse bias. The latter represents the highest ever reported bandwidth of silicon devices fabricated in CMOS-manufacturable process at a reverse bias accessible for high-volume applications. A common figure of merit in photodetector design is efficiency-bandwidth. However, efficiency only accounts for the light incident on the photodetector surface. Efficiency is also defined at a particular wavelength, and as discussed previously some photodetectors may operate at a high speed at one wavelength and poorly at another (section 1.2.3). One may also define a figure of merit area-bandwidth, which represents the product of bandwidth and modal overlap to a given source. Using such a definition, the 200 μm silicon device fabricated in this work have the highest reported figure of merit. Admittedly, the sensitivity at 850 nm is poor, but for applications in the red and, in the future, yellow light regimes over POF these devices have excellent performance.

A dependence of the device behaviour (both DC and AC) on the devices was observed for the first time as a consequence of the floating substrate node, explained with a qualitative model, and confirmed by simulation. Device bandwidths were also confirmed by simulation with an excellent match between fabricated and simulated performance.

Lateral germanium p-i-n photodetectors

Epitaxial germanium lateral photodetectors were fabricated and characterised. The second set of devices exhibiting excellent quantum efficiency and a very clear frequency response. Device performance was limited by the quality of the epitaxial layer. In the first set of devices, the carrier lifetime was too short to achieve collection of generated carriers in the germanium layer. In the second set of devices, the epitaxial layer had a high doping level, the consequence of which was a high device capacitance and short extent of the depletion region.

A germanium process flow was conceived that is essentially a folded version of past vertical device structures. This flow has some important advantages over the other published report of lateral germanium devices. First, it uses standard processing steps and lithography, most notably the metal fingers are sputtered and etched, rather than defined by a lift-off technique. Second, the polysilicon fingers are translucent to the wavelengths of interest, therefore gains in efficiency may be realised by using a thinner ohmic metal finger on top of the polysilicon electrode in order to

decrease the metal fill factor.

During the course of the germanium device fabrication, certain processing deficiencies were observed and explained. In particular, it was found that the mandated wafer cleaning process was not only ineffective, but in fact harmful to the growth substrates of the epitaxial wafers. This behaviour is examined and explained in full in appendix A.

Device simulation

All devices were simulated in MEDICI, and the simulated results were compared to measurements. The match for silicon devices, was excellent. All relevant device behaviour was confirmed *ab initio*, and investigated in depth.

Detectors on SOI substrates were investigated and compared with measured silicon device results. It was found that the fabricated silicon device bandwidths were dependent not on lateral diffusion as originally thought, but in fact on vertical diffusion and lateral drift, explaining the much higher than expected bandwidths for 3 and 4 μm spacing devices. This important discovery, enabling the decoupling of the absorption and carrier movement processes though a relatively fast pathway, led to the proposal of a novel SOI photodetector with buried p-n junction. These devices may be fabricated using higher doping levels than previously reported high-resistivity SOI devices, and exhibits a factor of two improvement in bandwidth and an approximately equal efficiency as compared to the high-resistivity SOI devices. The efficiency-bandwidth product may be improved further by optimisation of the electrode spacing and layer structure.

Simulations were performed on germanium epitaxial structures to determine the proper polarity of growth substrate. Silicon n-type substrates result in significantly higher bandwidths than p-type substrates, as a result of the p-n junction and band discontinuity. Additionally, it was found that the germanium on ultra-thin SOI structures proposed by other groups have a lower bandwidth than germanium grown on bulk n-type silicon due to the lack of buried heterojunction.

Finally, scaling of germanium devices was investigated to determine the maximum attainable bitrate for a given area. It was found that a project goal of a bitrate of 10 Gb/s is not achievable for a 200 μm diameter device. In order to reach 10 Gb/s, the device area diameter must be reduced to

approximately 150 μm . From a SNR analysis, it is found that such a device is well-suited for 10 Gb/s operation over a 200 μm core diameter polymer optical fibre.

8.2 Future Work

The future work to be carried out may be separated into two fields: process development and device development.

8.2.1 Process development

Much fundamental work is required before germanium may go from a classification of CMOS-compatible (a phrase that may be everything and nothing at once) to CMOS-manufacturable. Specifically, the following areas are in need of urgent development.

Wafer Cleaning A starting point for this discussion is appendix A. Standard RCA cleaning is incompatible with germanium, but some process is required prior to steps such as LPCVD polysilicon deposition or low-temperature oxide deposition. A full study is required of different cleaning chemistry, as was carried out in the original development of the RCA clean. Without a proven effective cleaning procedure, surface passivation will be difficult to achieve.

Surface Passivation The lateral geometry should have superior dark current characteristics compared to a vertical device since the current and electrical field paths avoid the highly defective sidewalls resulting for a patterned growth or etch step. However, an effective passivation material must be determined. This work applied silicon oxinitride and low-temperature oxide films suggested by previous reports [85]. Other work has suggested a native nitride or oxynitride film [97, 98]. A full study in this area is very important to the field of germanium photodetectors as well as to the field of germanium MOSFETs that has recently regained significant interest.

Electrode formation, metallisation Further work would be helpful in determining optimum polysilicon electrode structure (i.e. deposition temperature, pressure, implantation doping) and metallisation structure in order to understand and avoid the situation as observed in this

work. In fact, the problem of poor yield due to contact and metallisation problems is a recurring theme in the germanium photodetector field and a systematic study may improve results dramatically.

8.2.2 Device development

The fundamental work remaining to be done in the device development area should be focused at strategies to achieve 10 Gb/s performance with a 200 μm diameter device, which is not possible according to the simulations performed in this work. The principles of either forcing complete depletion of the inter-finger area or decoupling the vertical and horizontal carrier flow directions seem the most promising.

SOI with buried junction This structure, simulated in this work, has the potential to offer much higher performance than previously reported SOI photodetectors. Further device simulation and eventual fabrication are required to confirm this design.

Si_xGe_{1-x} epitaxial material The very high absorption coefficient of germanium is not strictly required for applications in the 850 nm or shorter wavelength regimes that are of greatest interest for POF applications. A limiting factor in the maximum achievable bandwidth for germanium devices is the lower saturation velocity and higher dielectric constant as compared with silicon. A silicon-germanium alloy, if an appropriate epitaxy technique is developed, may provide an acceptable absorption at a higher saturation velocity and lower dielectric constant. This would result in a higher ultimate bandwidth without significantly sacrificing efficiency.

Graded Si_xGe_{1-x} epitaxial material Following the techniques of SiGe HBT manufacturing, the active region of the device could be graded, resulting in an effective built-in field. This technique is also subject to the development of an appropriate epitaxy scheme.

8.2.3 Further work outside the scope of photodiode devices

Development of yellow or green high bandwidth emitters The development of yellow or green vertical emitters (RCLEDs or VCSELs) would enable the use of all-silicon integrated pho-

todetectors with very high sensitivity.

Three-dimensional integration There has been recent interest in the use of metallic (copper) wafer-wafer bonding in order to achieve three-dimensional integrated circuits. The integration of a photodetector layer (similar to the parallel flip-chip solutions discussed earlier) could result in very high-density, low-cost optoelectronic integrated circuits.

Appendix A

Silicon and Germanium Wafer Cleaning Techniques

An important element of modern integrated circuit fabrication technology is effective wafer cleaning; contamination of device surfaces is minimised thereby improving device characteristics and more importantly device reproducibility. The industrial standard, also used in the microsystem technology laboratories (MTL), is the so-called RCA clean, reported by Kern and Puotinen in 1970 [99]. The description of the process and principle of cleaning below is based on this citation. The precise cleaning procedure employed in the MTL laboratories is shown in Table A.1.

Step	Solution	Comment
1	5:1:1 DI H_2O:H_2O_2:NH_4OH	80°C bath, 10 minutes, "SC-1"
2	DI	Dump rinse
3	50:1 HF:DI H_2O	RT, 15sec-1 minute
4	DI	Dump rinse
5	6:1:1 DI H_2O:H_2O_2:HCl	80°C bath, 15 minutes, "SC-2"
6	DI	Dump rinse
7	Spin Rinse, Dry	

Table A.1: Standard Si RCA clean recipe

The first solution (SC-1) is designed to remove organic contaminants, oxidised by the peroxide. The alkalinity of the solution enhances solubility (i.e. by a saponification process) of the contaminants and their oxidation products. The NH_4^+ ions are also able to complex some metals such as

Ag, Ni, Co, and Cd (precisely ions thereof). The HF dip serves to remove an oxide layer that is formed during the SC-1 process, as will be discussed below. This dip is 1 min long according to the MTL standard operating procedure, but 15 seconds is a more appropriate duration since approximately 5 nm of silicon dioxide must be etched. The second solution (SC-2) removes heavy metals by an oxidation and complexation process. Metals atoms are oxidised to their ionic form, and then complex with either Cl^- (e.g. $CuCl_4{}^{2-}$) or H_2O (e.g. $Fe(H_2O)_6{}^{3+}$), thereby enhancing dissolution and preventing redeposition onto the wafer surface. A concentrated H_2SO_4:H_2O_2 solution (so-called piranha clean) has a similar cleaning behaviour, but presents more disposal problems than SC-2.

This cleaning procedure is incompatible with exposed germanium films. Germanium etchants typically include H_2O_2 as their primary component, with the etching reaction consisting of oxidation of Ge following by dissolution of the resulting oxide [100, 101, 102, 103]. As discussed in previous sections (chapter 4.3), this principle is used in the 6:1 H_2O_2:H_2O etch used to reduce the thickness of epi-grown layers , and also in the higher rate 1:3:4 HCl:H_2O_2:H_2O etch used to remove Ge from the back (unpolished) side of the Si substrates prior to annealing. The latter etch was studied in detail for both Ge and Ga, concluding that the etch mechanism for both materials is a two step process: 1) oxidation stage involving H_2O_2, 2) complexation/dissolution stage involving Cl^- and H_2O [104]. In the case of Ga, for a given molar HCl concentration, the maximum etch rate is achieved for a H_2O_2 concentration of 0.25 times the HCl molarity. Below this concentration, the etching is limited by the Ga oxidation reaction rate (i.e. insufficient H_2O_2), and above this concentration the etching is limited by passivation of the surface by Ga_2O_3 formation (i.e. excess H_2O_2). At $c_{H_2O_2} = 0$ the etch rate is negligible, and at $c_{H_2O_2} \geq 0.5 \times c_{HCl}$ the etch rate is also zero. This effect is much less pronounced for Ge etching due to the solubility of GeO_2 in water. At $c_{H_2O_2} = 0$ the etch rate is zero and rises with increasing peroxide concentration, reaching a maximum at $c_{H_2O_2} \approx 0.5 \times c_{HCl}$, and further decreasing slightly and settling at a plateau that is approximately 10% below the maximum rate. The non-passivating nature of the germanium oxide prevents the use of the SC-2 bath as described above, and likewise the SC-1 bath.

To address the need for wafer surface cleaning prior to high temperature processing steps, MTL mandated the use of a modified, so-called germanium RCA clean, as shown in Table A.2. This procedure uses the main components of the RCA SC-1 and SC-2 baths, but rearranges them

into 3 steps to avoid aggressive germanium etching. No systematic study of the effectiveness of this process has been performed. The method was inspired by published reports [105, 106, 107] investigating surface cleaning in preparation for epitaxy. The goal of these published methods was to form a smooth oxide cap on the Ge surface which was them removed by thermal adsorption in the respective growth systems, presumably also thereby mechanically removing organic or metallic contaminants on the surface.

Step	Solution	Comment
1	4:1 DI H_2O:NH_4OH	RT, 5 minutes
2	DI	Dump rinse
3	6:1 DI H_2O:H_2O_2	RT, 30 seconds, $\approx$ 35nm etch
4	DI	Dump rinse
5	50:1 HF:DI H_2O	RT, 15sec
6	DI	Dump rinse
7	4:1 DI H_2O:HCl	RT, 30 seconds
8	DI	Dump rinse
9	Spin Rinse, Dry	

Table A.2: Ge RCA clean recipe

With reference to the silicon RCA procedure, it is apparent that the so-called germanium RCA is an ineffective procedure. Steps 1 and 3 are meant to perform similarly to SC-1. However, due to the lack of oxidising agent, step 1 chemically has no effect on metal impurities, and likewise is not expected to have a significant impact on organic molecules. Step 3 will result in oxidation of some organic contaminants, but its main effect is an etching of the germanium surface. Step 7 is also ineffective as only oxidised metal atoms (i.e. ions) may form soluble complexes. Therefore, the only significant result of this cleaning procedure is an etching of the germanium surface in step 3. This may result in some contaminant removal, however there are no complexing agents in the bath to encourage dissolution and/or prevent redeposition onto the etched germanium surface.

In some cases it is desirable to remove epitaxial germanium from the back sides of the growth substrates prior to a high-temperature step (for example as described in section ?? prior to RTP where temperature is measured by a pyrometer calibrated to an unpolished silicon surface). In this case, step 1 of the germanium RCA procedure is very problematic. The concentrated H_2O:NH_4OH solution is an effective silicon etchant [108, 109], exhibiting properties similar to KOH (i.e. etch

rates dependent on crystal facet), which can result in a rough surface after treatment. An example of an unpolished silicon surface following a Ge RCA step 1 bath is shown in Figure A-1

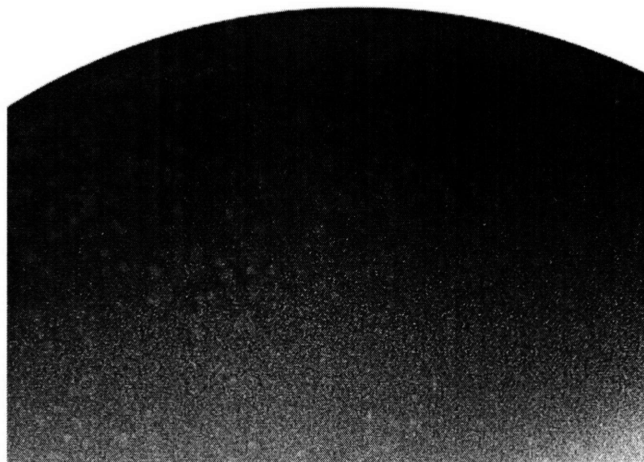
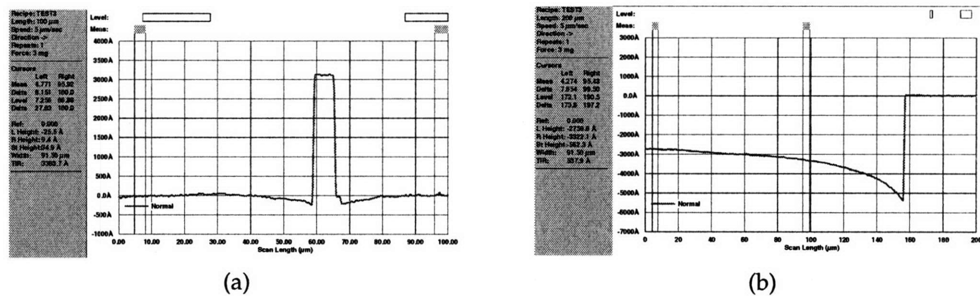


Figure A-1: The unpolished surface of a silicon wafer following step 1 of the Ge RCA cleaning procedure. Both surfaces of the silicon wafer are covered with rounded etched areas distributed over the surface

The addition of small amounts of hydrogen peroxide strongly affect the etching mechanism. As peroxide concentration rises from zero, the etch rate increases until reaching a critical concentration at which point the etch rate quickly falls to zero. The peroxide concentration is relatively small, specifically 3×10^{-3}M for a 1.9M NH_3 solution at 70°C. The mechanism for this behaviour is a competition between etching and oxidation of the silicon surface. At low concentrations, surface radical intermediates react with peroxide to form $Si(OH)_4$ (competing with a similar reaction involving OH^-, thereby enhancing the etch rate), while at higher concentrations the formation of SiO_2 dominates, resulting in a thin layer of surface passivation [110].

As is the case with other alkaline etches, the ammonium hydroxide etch mechanism includes an transfer of electrons between the electrolyte solution and the silicon conduction band. It has been shown that p+ doped silicon or a Si_xGe_{1-x} have much lower etch rates due to unfavourable band alignment[111]. Hence the exposed germanium surfaces are not attacked by ammonium hydroxide, but the unpolished silicon substrate backside are attacked. A natural experiment is to attempt a Ge RCA step 1 with the addition of 50 mL of H_2O_2 ($<$ 10 mM H_2O_2 concentration in the bath). In these conditions the silicon etch rate falls to approximately zero (as determined by profilometry). Ge-on-Si wafers with a patterned SiO_2 mask were also placed in the modified step 1

bath. After a BOE dip to remove the mask, it is visually apparent that significant etching has taken place ($\approx$ 60 nm/min), presumably due to the oxidation of germanium and the high solubility of GeO_2 (i.e. no surface passivation takes place). Interestingly, the etch rate is not uniform across the wafer and appears higher near larger mask features as shown in Figure A-2. Either the mask material, residual compounds in or near the mask, or wafer topology (perhaps an exposed crystal facet) appear to accelerate the etch process. This mechanism was not studied any further.



(a) (b)

Figure A-2: Profilometer traces of two different patterns on the same Ge-on-Si wafer following a modified Ge RCA step 1, followed by oxide mask removal. The jagged feature on the right falling edge of a) is due to a damaged profilometer tip.

In conclusion, the Ge RCA cleaning procedure is ineffective from a chemical point of view, despite consisting of qualitatively similar steps as the Si RCA cleaning procedure. The chemical reactions required for cleaning can not take place in the baths as formulated. Furthermore, step 1 may seriously damage exposed silicon surfaces on the wafers. An etching and/or roughening of the substrate backside may have a negative effect further in the process, particularly in process steps requiring a good seal between the wafer and wafer stage (e.g. He wafer cooling during dry etch processes). The addition of a small amount of hydrogen peroxide eliminates the silicon etching problem, but also results in a substantial germanium etch rate. In light of these experiences, the Ge RCA as described above should not be used. A reasonable alternative may be a series of short 6:1 DI H_2O:H_2O_2 dips and DI rinses as reported in [107], additionally followed each time by an aspiration of and remixing of the peroxide solution. This procedure will etch away some of the germanium film. Contaminants will not be removed chemically (i.e. by complexation), but there is a greater chance of mechanical removal as part of the etching process. Short dips and constants changes of the bath solution may also mechanically prevent redeposition.

Appendix B

Silicon and Germanium Device Detailed Process Description

B.1 Silicon Process

The detailed description of the silicon device process is shown below. The initial active region implant is performed by a two or three step implantation of boron at relatively low doses to achieve a relatively flat p-type profile approximately 1 μm deep.

Table B.1: Silicon lateral p-i-n detailed process description

Step	Description	Instrument, Process Recipe	Comment
0	Boron implantation	Innovion Inc.	2 or 3 step implant for isolated devices
1	RCA clean	rcaICL	Appendix A
2	Field oxide growth nominal 250 nm	Tube 5C-FieldOx, wet oxidation	1000°C, 30 min measured 240-250nm
3	Coat photoresist Expose photoresist Develop photoresist	coater6 i-stepper coater6	Alignment marks and Finger openings
4	Oxide etch	AME5000 (Baseline Ox New)	70 sec, nominally 2800 Å

Continued on Next Page...

Table B.1 – Continued

Step	Description	Process Recipe	Comment
5	Strip photoresist	asher-ICL	3 min
6	Coat photoresist Pattern photoresist Develop photoresist	coater6 i-stepper coater6	Phosphorus implant mask overexpose 15 ms
7	Phosphorus implantation	Innovion Inc.	Dose: 1e15 cm^{-2} Energy: 75 keV
8	Wafer clean	asher-ICL, premetal-piranha (2x)	
9	Coat photoresist Pattern photoresist Develop photoresist	coater6 i-stepper coater6	Boron implant mask overexpose 15 ms
10	Boron implantation	Innovion Inc.	Dose: 1e15 cm^{-2} Energy: 15 keV
11	Wafer clean	asher-ICL, premetal-piranha (2x)	
12	RCA clean	rcaICL	
	Dopant anneal	5B-Anneal	950°C, 30 min, N_2 ambient
13	Premetal wafer clean Metal deposition	premetal-piranha endura	 250-500 nm AlSi(2%)
14	Coat photoresist Pattern photoresist Develop photoresist	coater6 i-stepper coater6	Metal etch mask
15	Metal etch	rainbow 11032005	Etch stop monitor 25 sec for 250 nm
16	Strip photoresist	asher-ICL	
17	Metal anneal	TRL tubeA3	400°C, 30 minutes 95% N_2, 5% H_2 ambient
18	SiO_2 deposition	concept1	0.5 μm
19	Coat photoresist	coater6	Via etch mask

Continued on Next Page...

Table B.1 – Continued

Step	Description	Process Recipe	Comment
	Pattern photoresist Develop photoresist	i-stepper coater6	
20	SiO_2 etch	centura (wpgoxetch, 120 s oxygen clean)	High density plasma C_2F_6 chemistry
21	Strip photoresist	asher-ICL	
22	Metal deposition	endura	750 nm AlSi(2%)
13	Coat photoresist Pattern photoresist Develop photoresist	coater6 i-stepper coater6	Pad etch mask
24	Metal etch	rainbow 11032005	Etch stop monitor 70 sec for 750 nm
25	Strip photoresist	asher-ICL	

B.2 Germanium Process

The detailed description of the germanium device process flow is shown below. It may be separated into two sections: germanium growth and annealing, and device fabrication given an annealed wafer.

Table B.2: Epitaxial germanium growth and annealing process description

Step	Description	Instrument, Process Recipe	Comment
0	Starting material: prime grade Si wafer		
1	RCA clean	rcaICL	Appendix A
2	Germanium epitaxy	Sirius 6" UHVCVD Optional: *in situ* post-growth thermal anneal proceed to device fabrication flow, table B.3	Nominally 1 μm epilayer

Continued on Next Page...

Table B.2 – Continued

Step	Description	Process Recipe	Comment
3	Ge RCA clean	rcaICL	Appendix A
4	Cap dielectric deposition	DCVD	3000 Å HDP SiO_2
	alternative	Tube 6C-LTO	400°C, 2200 Å SiO_2
5	This step only if using LTO cap dielectric		
	Coat photoresist	coater6	
	Etch LTO	oxEtch-BOE	2 min
	Strip photoresist	asher-ICL	3 min
6	Etch backside germanium	TRL acidhood2	1:3:4 H_2O_2:HCl:H_2O
7	RCA clean	rcaICL	Front Ge protected
	Cyclic anneal	RTP	
		ge686ls: 650-850-650°C, 30 s at each temperature, 10 cycles	
8	Cap dielectric etch	oxEtch-BOE	3 min

Table B.3: Germanium lateral p-i-n detailed process description

Step	Description	Instrument, Process Recipe	Comment
0	Starting material: annealed Ge epilayer on Si		
1	Ge RCA clean	rcaICL	Appendix A
2	Passivation deposition	DCVD	2000 Å HDP SiON $n = 1.7$, NH_3 precursor
	alternative	Tube 6C-LTO	400°C, 2200 Å SiO_2
3	Coat photoresist	coater6	Alignment marks and
	Expose photoresist	i-stepper	Finger openings
	Develop photoresist	coater6	
4	Oxide etch	AME5000 (Baseline Ox New)	SiON: 45 sec LTO: 60 sec
5	Strip photoresist	asher-ICL	3 min

Continued on Next Page...

Table B.3 – Continued

Step	Description	Process Recipe	Comment
6	Ge RCA clean	rcaICL	Appendix A
7	Polysilicon deposition	TRL tubeB4	620°C, nominally 2000 Å measured 1600-2400 Å
8	Coat photoresist Pattern photoresist Develop photoresist	coater6 i-stepper coater6	Phosphorus implant mask overexpose 15 ms
9	Phosphorus implantation	Innovion Inc.	Dose: 5e15 cm^{-2} Energy: 90 keV
10	Wafer clean	asher-ICL, premetal-piranha (2x)	
11	Coat photoresist Pattern photoresist Develop photoresist	coater6 i-stepper coater6	Boron implant mask overexpose 15 ms
12	Boron implantation	Innovion Inc.	Dose: 2e15 cm^{-2} Energy: 25 keV
13	Wafer clean	asher-ICL, premetal-piranha (2x)	
14	RCA clean	TRL RCA	
	Dopant anneal	TRL tubeB3	650°C, 30 min, N_2 ambient
15	Coat photoresist Pattern photoresist Develop photoresist	coater6 i-stepper coater6	Polysilicon etch mask underexpose 15 ms
16	Polysilicon Etch	AME5000 Baseline Poly	Endpoint monitor ($\approx$ 60 s + 15 s overetch)
17	Premetal wafer clean Metal deposition	premetal-piranha endura	 500 nm AlSi(2%)
18	Coat photoresist Pattern photoresist Develop photoresist	coater6 i-stepper coater6	Metal etch mask
19	Metal etch	rainbow 11032005	Etch stop monitor 45 sec for 500 nm

Continued on Next Page...

Table B.3 – Continued

Step	Description	Process Recipe	Comment
20	Strip photoresist	asher-ICL	
21	Metal anneal	TRL tubeA3	400°C, 30 minutes 95% N_2, 5% H_2 ambient
22	SiO_2 deposition	concept1	0.5 μm
23	Coat photoresist Pattern photoresist Develop photoresist	coater6 i-stepper coater6	Via etch mask
24	SiO_2 etch	centura (wpgoxetch, 120 s oxygen clean)	High density plasma C_2F_6 chemistry
25	Strip photoresist	asher-ICL	
26	Metal deposition	endura	750 nm AlSi(2%)
27	Coat photoresist Pattern photoresist Develop photoresist	coater6 i-stepper coater6	Pad etch mask
28	Metal etch	rainbow 11032005	Etch stop monitor 70 sec for 750 nm
29	Strip photoresist	asher-ICL	

Bibliography

[1] J. L. Zerbe, C. W. Werner, V. Stojanovic, F. Chen, J. Wei, G. Tsang, D. Kim, W. F. Stonecypher, A. Ho, T. P. Thrush, R. T. Kollipara, M. A. Horowitz, and K. S. Donnelly, "Equalization and clock recovery for a 2.5-10-Gb/s 2-PAM/4-PAM backplane transceiver cell," *IEEE Journal of Solid-State Circuits*, vol. 38, pp. 2121–2130, 2003.

[2] R. King, R. Michalzik, R. Jaeger, K. J. Ebeling, R. Annen, and H. Melchior, "32-VCSEL channel CMOS-based transmitter module for Gb/s data rates," in *Vertical-Cavity Surface-Emitting Lasers V* (K. D. Choquette and C. Lei, eds.), vol. 4286, pp. 136–141, Proceedings of SPIE, 2001.

[3] R. Michalzik, R. King, F. Mederer, M. Kircherer, G. Giaretta, and K. J. Ebeling, "Short-wavelength vertical-cavity surface-emitting laser applications: from high-throughput multimode fiber links to two-dimensional interchip interconnections," *Optical Engineering*, vol. 40, pp. 1179–1185, July 2001.

[4] L. Schares, C. Schow, F. Doany, C. Schuster, J. Kash, D. Kuchta, P. Pepeljugoski, J. Schaub, J. Trewhella, C. Baks, R. John, L. Shan, S. Hegde, Y. Kwark, D. Rogers, F. Libsch, R. Budd, P. Chiniwalla, J. Rosner, C. Tsang, C. Patel, D. Kucharski, D. Guckenberger, R. Dangel, B. Offrein, M. Tan, G. Trott, A. Nystrom, A. Tandon, C. K. Lin, and D. Dolfi, ""Terabus" - a waveguide-based parallel optical interconnect for Tb/s-class on-board data transfers in computer systems," in *31st European Conference on Optical Communication*, vol. 3, pp. 369–372, 2005.

[5] T. Calvert, B. Corbett, and J. D. Lambkin, "80 degrees C continuous wave operation of AIGaInP based visible VCSEL," *Electronics Letters*, vol. 38, pp. 222–223, February 2002.

[6] J. D. Lambkin, T. Calvert, J. Woodhead, S. M. Pinches, J. Frost, J. Hosea, P. Van Daele, K. Vandeputte, A. Van Hove, A. Valster, J. G. McInerney, and P. A. Porta, "Development of a visible vcsel-to-plastic optical fibre module for use in high-speed optical data links," *Materials Science in Semiconductor Processing*, vol. 3, pp. 467–473, Oct-Dec 2000.

[7] J. Berwanger, M. Peller, and R. Griessbah, "Byteflight: A new high-performance data bus system for safety-related applications." White paper, BMW.

[8] P. Polishuk, "Automotive industry in Europe takes the lead in the introduction of optical data buses," *Wiring Harness News - Europe*, Nov/Dec 2001.

[9] J. Lambkin, "Private correspondance." Components for PMMA transceivers - Firecomms Inc.

[10] T. Ishigure, M. Sato, A. Kondo, and Y. Koike, "High-bandwidth graded-index polymer optical fiber with high-temperature stability," *Journal of Lightwave Technology*, vol. 20, pp. 1443–1448, August 2002.

[11] T. Ishigure, Y. Koike, and J. W. Fleming, "Optimum index profile of the perfluorinated polymer-based GI polymer optical fiber and its dispersion properties," *Journal of Lightwave Technology*, vol. 18, pp. 178–184, February 2000.

[12] Y. Koike, T. Ishigure, and E. Nihei, "High-bandwidth graded-index polymer optical fiber," *Journal of Lightwave Technology*, vol. 13, pp. 1475–1489, July 1995.

[13] Fuji Photo Film Co., Ltd, "Fujifilm develops new graded index plastic optical fiber "LUMISTAR-X" - enables high-speed transmission greater than 10 Gbps." Press Release, January 2006.

[14] M. Jutzi, K. Eve, W. Vogel, D. Wiegner, and M. Berroth, "Lateral PIN-photodetector in commercial CMOS technology operating at 1.25 Gbit/s and 850 nm," in *Proceedings 7th Workshop "Optics in Computing Technology"*, (Mannheim), September 2002.

[15] M. Jutzi, M. Grözing, E. Gaugler, W. Mazioschek, and M. Berroth, "2-Gb/s CMOS optical integrated receiver with a spatially modulated photodetector," *IEEE Photonics Technology Letters*, vol. 17, pp. 1268–1270, June 2005.

[16] M. Yang, K. Rim, D. L. Rogers, J. D. Schaub, J. J. Welser, D. M. Kuchta, D. C. Boyd, F. Rodier, P. A. Rabidoux, J. T. Marsh, A. D. Ticknor, Q. Yang, A. Upham, and S. C. Ramac, "A high-speed, high-sensitivity silicon lateral trench photodetector," *IEEE Electron Device Letters*, vol. 23, pp. 395–397, July 2002.

[17] M. Yang, J. Schaub, D. Rogers, M. Ritter, K. Rim, J. Welser, and B. Park, "High speed silicon lateral trench detector on SOI substrate," *IEDM Technical Digest*, pp. 547–550, 2001.

[18] C. Schow, R. Li, J. Schaub, and J. Campbell, "Design and implementation of high-speed planar Si photodiodes fabricated on SOI substrates," *IEEE Journal of Quantum Electronics*, vol. 35, pp. 1478–1482, October 1999.

[19] S. M. Csutak, J. D. Schaub, W. E. Wu, and J. C. Campbell, "High-speed monolithically integrated silicon optical receiver fabricated in 130-nm CMOS technology," *IEEE Photonics Technology Letters*, vol. 14, pp. 516–518, April 2002.

[20] R. Li, J. D. Schaub, S. M. Csutak, and J. C. Campbell, "A high-speed monolithic silicon photoreceiver fabricated on SOI," *IEEE Photonics Technology Letters*, vol. 12, pp. 1046–1048, August 2000.

[21] J. Schaub, R. Li, S. Csutak, and J. Campbell, "High-speed monolithic silicon photoreceivers on high resistivity and SOI substrates," *Journal of Lightwave Technology*, vol. 19, pp. 272–278, February 2001.

[22] M. Y. Liu, E. Chen, and S. Y. Chou, "140-GHz metal-semiconductor-metal photodetectors on silicon-on-insulator substrate with a scaled active layer," *Applied Physics Letters*, vol. 65, no. 7, pp. 887–888, 1994.

[23] M. Yamamoto, M. Kubo, and K. Nakao, "Si-OEIC with a built-in pin-photodiode," *IEEE Transactions on Electron Devices*, vol. 42, pp. 58–63, January 1995.

[24] R. Swoboda and H. Zimmermann, "A low-noise 1.8 Gbps bipolar OEIC," in *Proceedings of the 29th European Solid-state circuits Conference*, pp. 341–344, 2003.

[25] R. Swoboda and H. Zimmermann, "A low-noise monolithically integrated 1.5 Gb/s optical receiver in 0.6 μm BiCMOS technology," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 9, pp. 419–424, March/April 2003.

[26] M. Hein, M. Förtsch, and H. Zimmermann, "Low-power 300 Mbit/s OEIC with large-area photodiode," *Electronics Letters*, vol. 41, pp. 436–438, March 2005.

[27] J. Sturm, M. Leifhelm, H. Schatzmayr, S. Groiß, and H. Zimmermann, "Optical receiver IC for CD/DVD/blue-laser application," *IEEE Journal of Solid-State Circuits*, vol. 40, pp. 1406–1413, July 2005.

[28] R. Swoboda and H. Zimmermann, "2.5 Gbit/s silicon receiver OEIC with large diameter photodiode," *Electronics Letters*, vol. 40, pp. 505–507, April 2004.

[29] R. Swoboda, J. Knorr, and H. Zimmermann, "A 5-Gb/s OEIC with voltage-up-converter," *IEEE Journal of Solid-State Circuits*, vol. 40, pp. 1521–1526, July 2005.

[30] S. Luryi, A. Kastalsky, and J. C. Bean, "New infrared detector on a silicon chip," *IEEE Transactions on Electron Devices*, vol. 31, pp. 1135–1139, 1984.

[31] S. B. Samavedam, M. T. Currie, T. A. Langdo, and E. A. Fitzgerald, "High-quality germanium photodiodes integrated on silicon substrates using optimized relaxed graded buffers," *Applied Physics Letters*, vol. 73, pp. 2125–2127, October 1998.

[32] H. C. Luan, K. Wada, L. C. Kimerling, G. Masini, L. Colace, and G. Assanto, "High efficiency photodetectors based on high quality epitaxial germanium grown on silicon substrates," *Optical Materials*, vol. 17, pp. 71–73, Jun-Jul 2001.

[33] H.-C. Luan, *Ge photodetectors for Si microphotonics*. PhD thesis, Massachusetts Institute of Technology, 2001.

[34] J. F. Liu, D. D. Cannon, K. Wada, Y. Ishikawa, S. Jongthammanurak, D. T. Danielson, J. Michel, and L. C. Kimerling, "Tensile strained Ge p-i-n photodetectors on Si platform for C and L band telecommunications," *Applied Physics Letters*, vol. 87, p. 011110, July 2005.

[35] J. Liu, D. D. Cannon, K. Wada, Y. Ishikawa, S. Jongthammanurak, D. T. Danielson, J. Michel, and L. C. Kimerling, "Silicidation-induced band gap shrinkage in Ge epitaxial films on Si," *Applied Physics Letters*, vol. 84, no. 5, p. 660, 2004.

[36] O. I. Dosunmu, D. D. Cannon, M. K. Emsley, B. Ghyselen, J. Liu, L. C. Kimerling, and M. S. Ünlü, "Resonant cavity enhanced Ge photodetectors for 1550 nm operation on reflecting Si substrates," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 10, pp. 694–701, August 2004.

[37] O. I. Dosunmu, D. D. Cannon, M. K. Emsley, L. C. Kimerling, and M. S. Ünlü, "High-speed resonant cavity enhanced Ge photodetectors on reflecting Si substrates for 1550-nm operation," *IEEE Photonics Technology Letters*, vol. 17, pp. 175–177, January 2005.

[38] M. Jutzi, M. Berroth, G. Wöhl, M. Oehme, and E. Kasper, "Ge-on-Si vertical incidence photodiodes with 39-GHz bandwidth," *IEEE Photonics Technology Letters*, vol. 17, pp. 1510–1512, July 2005.

[39] J. Oh, S. Csutak, and J. Campbell, "High-speed interdigitated Ge PIN photodetectors," *IEEE Photonics Technology Letters*, vol. 14, pp. 369–371, March 2002.

[40] G. Dehlinger, S. J. Koester, J. D. Schaub, J. O. Chu, Q. C. Ouyang, and A. Grill, "High-speed germanium-on-SOI lateral PIN photodiodes," *IEEE Photonics Technology Letters*, vol. 16, pp. 2547–2549, November 2004.

[41] S. J. Koester, J. D. Schaub, G. Dehlinger, J. O. Chu, Q. C. Ouyang, and A. Grill, "High-efficiency, Ge-on-SOI lateral PIN photodiodes with 29 GHz bandwidth," in *Device Research Conference*, June 2004.

[42] S. J. Koester, G. Dehlinger, J. D. Schaub, J. O. Chu, Q. C. Ouyang, and A. Grill, "Germanium-on-insulator photodetectors," in *2nd International Conference on Group IV Photonics*, (Antwerp, Belgium), 21–23 September 2005.

[43] L. Schares, C. Schow, S. J. Koester, G. Dehlinger, R. John, and F. E. Doany, "A 17-Gb/s low-power optical receiver using a Ge-on-SOI photodiode with a 0.13-μm CMOS IC," in *OFC/NFOEC*, 2006.

[44] J. I. Pankove, *Optical Processes in Semiconductors*. Prentice-Hall, Inc., 1971.

[45] K. Ebeling, *Integrated Optoelectronics*. Springer, 1992.

[46] L. M. Giovane, *Strain-balanced silicon-germanium materials for near IR photodetection in silicon-based optical interconnects*. PhD thesis, Massachusetts Institute of Technology, 1998.

[47] C. G. Fonstad, *Microelectronic Devices and Circuits*. McGraw-Hill, Inc., 1994.

[48] S. Sze, *Physics of Semiconductor Devices*. Wiley-Interscience, 2nd ed., 1981.

[49] G. Lucovsky, R. Schwarz, and R. Emmons, "Transit-time considerations in p-i-n diodes," *Journal of Applied Physics*, vol. 35, pp. 622–628, March 1964.

[50] D. A. Tulchinsky, X. Li, N. Li, S. Demiguel, J. C. Campell, and K. J. Williams, "High-saturation current wide-bandwidth photodetectors," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 10, pp. 702–708, July/August 2004.

[51] M. Dentan and B. de Cremoux, "Numerical simulation of the nonlinear response of a p-i-n photodiode under high illumination," *Journal of Lightwave Technology*, vol. 8, pp. 1137–1144, August 1990.

[52] J. E. Bowers and C. A. Burrus, "Ultrawide-band long-wavelength p-i-n photodetectors," *Journal of Lightwave Technology*, vol. LT-5, pp. 1339–1350, October 1987.

[53] R. Sabella and S. Merli, "Analysis of InGaAs p-i-n photodiode frequency response," *IEEE Journal of Quantum Electronics*, vol. 29, pp. 906–915, March 1993.

[54] G. Wang, T. Tokumitsu, I. Hanawa, K. Sato, and M. Kobayashi, "Analysis of high speed p-i-n photodiodes S-parameters by a novel small-signal equivalent circuit model," *IEEE Microwave and Wireless Components Letters*, vol. 12, pp. 378–380, October 2002.

[55] Y.-G. Wey, K. Giboney, J. Bowers, M. Rodwell, P. Silvestre, P. Thiagarajan, and G. Robinson, "110-GHz GaInAs/InP double heterostructure p-i-n photodetectors," *Journal of Lightwave Technology*, vol. 13, pp. 1490–1499, July 1995.

[56] D. E. Sawyer and R. H. Rediker, "Narrow base germanium photodiodes," *Proceedings of the IRE*, vol. 46, pp. 1122–1130, June 1958.

[57] Ioffe Physico-Technical Insitute, "New semiconductor materials archive - physical properties of semiconductors."

[58] M. S. Tyagi and R. Van Overstraeten, "Minority carrier recombination in heavily-doped silicon," *Solid-State Electronics*, vol. 26, pp. 577–597, June 1983.

[59] J. Liu, J. Michel, W. Giziewicz, D. Pan, K. Wada, D. D. Cannon, S. Jongthammanurak, D. T. Danielson, L. C. Kimerling, J. Chen, F. O. Ilday, F. X. Kartner, and J. Yasaitis, "High-performance, tensile-strained Ge p-i-n photodetectors on a Si platform," *Applied Physics Letters*, vol. 87, no. 10, p. 103501, 2005.

[60] J. G. Graeme, *Photodiode Amplifiers: Op Amp Solutions*. McGraw-Hill, Inc., 1995.

[61] J. K. Roberge, *Operational Amplifiers*. John Wiley & Sons Inc., 1975.

[62] B. Zand, K. Phang, and D. A. Johns, "Transimpedance amplifier with differential photodiode current sensing," in *ISCAS*, pp. II–624–II–627, 1999.

[63] Analog Devices Inc., "ADN2821 transimpedance amplifier." Technical Data Sheet.

[64] Agilent Technologies Inc., "Agilent technologies S-parameter and TDR impedance measurement solution summary." Application Note, 2005.

[65] Agilent Technologies Inc., "High-precision time-domain reflectometry with the Agilent 86100 digital communications analyzer and Picosecond Pulse Labs 4020 source enhancement module." Application Note, 2003.

[66] W. Giziewicz, C. G. Fonstad, and S. Prasad, "High speed 0.9 μm lateral p-i-n photodetectors in a standard commercial GaAs VLSI process," *International Journal of High Speed Electronics and Systems*, vol. 14, pp. 714–719, September 2004.

[67] K. S. Giboney, R. L. Nagarajan, T. E. Reynolds, S. T. Allen, R. P. Mirin, M. J. W. Rodwell, and J. E. Bowers, "Travelling-wave photodetectors with 172-GHz bandwidth and 76-GHz bandwidth-efficiency product," *IEEE Photonics Technology Letters*, vol. 7, pp. 412–414, April 1995.

[68] G. Keiser, *Optical Communications Essentials*. McGraw-Hill Networking & Telecommunications, McGraw-Hill, Inc., 2003.

[69] J. R. Andrews, "RZ vs. NRZ." Picosecond Pulse Labs, Application Note AN-12, September 2001.

[70] D. K. Mynbaev and L. L. Scheiner, *Fiber-Optic Communications Technology*. Prentice-Hall, Inc., 2001.

[71] J. F. Ahadian, *Development of a Monolithic Very Large Scale Optoelectronic Integrated Circuit Technology*. PhD thesis, Massachusetts Institute of Technology, February 2000.

[72] J. Zerbe, "High performance link design: challenges and tradeoffs beyond 5 Gb/s," seminar series, MTL VLSI Seminar, 2005.

[73] Q. Xu, B. Schmidt, S. Pradhan, and M. Lipson, "Micrometre-scale silicon electro-optic modulator," *Nature*, vol. 435, no. 7040, pp. 325–327, 2005.

[74] L. A. Coldren and S. W. Corzine, *Diode Lasers and Photonic Integrated Circuits*. Wiley Series in microwave and optical engineering, Wiley-Interscience, 1995.

[75] J. Schaub, R. Li, C. Schow, J. Campbell, G. Neudeck, and J. Denton, "Resonant-cavity-enhanced high-speed Si photodiode grown by epitaxial lateral overgrowth," *IEEE Photonics Technology Letters*, vol. 11, pp. 1647–1649, December 1999.

[76] S. S. Gevorgian, T. Martinsson, P. L. J. Linnkr, and E. L. Kollberg, "CAD models for multilayered substrate interdigital capacitors," *IEEE Transactions on Microwave Theory and Techniques*, vol. 44, pp. 896–904, June 1996.

[77] C. Jacobini, C. Canali, G. Ottaviani, and A. Quaranta, "A review of some charge transport properties of silicon," *Solid-State Electronics*, vol. 20, no. 2, pp. 77–89, 1977.

[78] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*. Cambridge University Press, 1999.

[79] C. Jacobini, F. Nava, C. Canali, and G. Ottaviani, "Electron drift velocity and diffusivity in germanium," *Physical Review B*, vol. 24, no. 2, pp. 1014–1026, 1981.

[80] A. Yariv, *Optical Electronics in Modern Communications*. New York: Oxford University Press, 5 ed., 1997.

[81] R. T. Howe and C. G. Sodini, *Microelectronics: an Integrated Approach*. Prentice Hall Electronics and VLSI Series, Prentice-Hall, Inc., 1997.

[82] STMicroelectronics Users Reliability Engineering, "SURE 8 - quality and reliability program." Technical Booklet, 2000.

[83] D. D. Cannon, *Strain-engineered CMOS-compatible Ge photodetectors*. PhD thesis, Massachusetts Institute of Technology, 2003.

[84] O. O. Olubuyide, "Backside germanium removal." Private Correspondance, February 2006.

[85] C. W. Leitz, *High mobility strainged Si/Ge heterostructure MOSFETs: channel engineering and virtual substrate optimization*. PhD thesis, Massachusetts Institute of Technology, 2002.

[86] D. Cuttriss, "Relation between surface concentration of average conductivity in diffused layers in germanium," *Bell System Technical Journal*, vol. 40, pp. 509–521, March 1961.

[87] O. Nast, T. Puzzer, L. M. Koschier, A. B. Sproul, and S. R. Wenham, "Aluminum-induced crystallization of amorphous silicon on glass substrates above and below the eutectic temperature," *Applied Physics Letters*, vol. 73, pp. 3214–3216, November 1998.

[88] M. Gjukic, M. Buschbeck, R. Lechner, and M. Stutzman, "Aluminum-induced crystallization of amorphous silicon-germanium thin films," *Applied Physics Letters*, vol. 85, pp. 2134–2136, September 2004.

[89] A. J. McAlister and J. L. Murray, "The Al-Ge (aluminium-germanium) system," *Bulletin of Alloy Phase Diagrams*, vol. 5, no. 4, pp. 341–347, 1984.

[90] H. Choy, M. Kircherer, B. Ruedlinger, K. Sundarajan, J. Ahadian, J. Mikkelson, S. Prasad, and C. G. Fonstad, "High performance lateral p-i-n photodetectors realized in a standard commercial GaAs IC process," *Proceeding of SPIE 4746*, pp. 797–803, 2001.

[91] W. P. Giziewicz, H. K. H. Choy, C. G. Fonstad, and S. Prasad, "Measurement and modelling of high performance lateral p-i-n photodetectors," in *Proceedings of SPIE*, vol. 5445, pp. 114–119, 2004.

[92] W. Giziewicz, S. Prasad, and C. G. Fonstad, "Lateral p-i-n photodetectors fabricated in a standard commercial GaAs VLSI process," in *Sensors 2004*, (Vienna, Austria), October 2004.

[93] A. Fan, "Inter-metal level resistance problem." Private Correspondance, May 2006.

[94] C. Hilsum, "Simple empirical relationship between mobility and carrier concentration," *Electronics Letters*, vol. 10, no. 13, pp. 259–260, 1974.

[95] European Committee for Electrotechnical Standardization, "Safety of laser products." European Standard EN 60825-2, October 2004.

[96] J. Gowar, *Optical Communication Systems*. Series in Optoelectronics, Prentice-Hall, Inc., 1984.

[97] D. J. Hymes and J. J. Rosenberg, "Growth and materials characterization of native germanium oxynitride thin films on germanium," *Journal of the Electrochemical Society*, vol. 134, pp. 961–965, April 1988.

[98] J. J. Rosenberg and S. C. Martin, "Self-aligned germanium MOSFET's using a nitrided native oxide gate insulator," *IEEE Electron Device Letters*, vol. 9, pp. 639–640, December 1988.

[99] W. Kern and D. A. Puotinen, "Cleaning solutions based on hydrogen peroxide for use in silicon semiconductor technology," *RCA Review*, vol. 31, pp. 187–206, March 1970.

[100] J. Bloem and J. C. van Vessem, "Etching Ge with mixturs of HF-H_2O_2-H_2O," *Journal of the Electrochemical Society*, vol. 109, pp. 33–36, January 1962.

[101] P. R. Camp, "A study of the etching rate of single-crystal germanium," *Journal of the Electrochemical Society*, vol. 102, pp. 586–593, October 1955.

[102] S. K. Ghandhi and J. E. Ayers, "Chemical etching of germanium," *Journal of the Electrochemical Society*, vol. 135, pp. 2053–2054, August 1988.

[103] W. Kern, "Chemical etching of silicon, germanium, gallium arsenide, and gallium phosphide," *RCA Review*, vol. 39, pp. 278–308, June 1978.

[104] E. P. Veretenkin, S. M. Kireev, L. A. Nisel'son, and I. V. Podlipaeva, "Kinetics of the dissolution of gallium and germanium in hydrochloric acid-peroxide solutions," *Russian Journal of Physical Chemistry*, vol. 58, pp. 1821–1823, December 1984.

[105] T. Akane, H. Okumura, J. Tanaka, and S. Matsumoto, "New Ge substrate cleaning method for $Si_{1-x-y}Ge_xC_y$ MOMBE growth," *Thin solid films*, vol. 294, pp. 153–156, 1997.

[106] H. Okumura, T. Akane, and S. Matsumoto, "Carbon contamination free Ge(100) surface cleaning for MBE," *Applied Surface Science*, vol. 125, pp. 125–128, 1998.

[107] K. Prabhakarana, T. Ogino, R. Hull, J. Bean, and L. Peticolas, "An efficient method for cleaning Ge(100) surface," *Surface Science*, vol. 316, pp. L1031–L1033, 1994.

[108] U. Schnakenberg, W. Beneche, B. Löchel, S. Ullerich, and P. Lange, "NH_4OH-based etchants for silicon micromachining: influence of additives and stability of passivation layers," *Sensors and Actuators A*, vol. 25, pp. 1–7, Oct–Jan 1991.

[109] U. Schnakenberg, W. Beneche, and B. Löchel, "NH_4OH-based etchants for silicon micromachining," *Sensors and Actuators A*, vol. 23, pp. 1031–1035, April 1990.

[110] J. E. A. M. van den Meerakker and M. H. M. van der Straaten, "A mechanistic study of silicon etching in NH_3/H_2O_2 cleaning solutions," *Journal of the Electrochemical Society*, vol. 137, pp. 1239–1243, April 1990.

[111] K. C.-C. Wu, "Novel etch-stop materials for silicon micromachining," Master of Science Thesis, Massachusetts Institute of Technology, June 1997.