University of Sheffield

PhD Thesis

Low-Photon Detection Using InGaAs/InAlAs and InAs Avalanche Photodiodes

Author:
Vladimir Shulyak

Supervisor:
Prof. Jo Shien Ng

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy

in the

Department of Electronic and Electrical Engineering

April 2020

“Be your own spectator; seek your own applause.”

Seneca the Younger

Acknowledgements

I would like to give my sincerest thanks to my supervisor Prof. Jo Shien Ng for the opportunity to pursue my Ph.D. Your patience, guidance, and financial support have been invaluable, especially during the difficult moments in my research. I am grateful to Prof. Chee Hing Tan and Prof. John David for sharing advice and enlightening discussions. I am especially thankful to Dr Simon Dimler for the many hours spent advising me about measurements and debugging issues with equipment; this work would not have been possible without you.

I give thanks also to all the members of the IIG group, past and present, particularly Jon Petticrew, Dr Lucas Pinel, Dr Salman Abdullah, Leh Woon Lim, Elizabeth Stark, Dr Benjamin White, Jonathan Taylor-Mew, Yuting Ji, Dr Xinxin Zhou, Dr Xiao Meng, Stephen Doward, and Dr Ian Sandall for their friendship, and enriching my experience throughout my Ph.D.

I would also like to thank the staff at the University cleanroom facilities, in particular Dr Saurabh Kumar, Dr Kenneth Kennedy, Dr Timmothy Lewis-Roberts, and Dr Rob Airey for their many tips and tricks for device fabrication.

I greatly appreciate the support from the EPSRC, who provided invaluable funding to allow me to pursue this Ph.D.

Finally, I like to give a big thank you to my parents and friends. In particular I would like to thank Eve and Matt, without whom I would never have come so far, for their care and support during the hardest of times.

Abstract

Single-photon avalanche photodiodes (SPADs) based on the InGaAs/InAlAs material system are designed, fabricated, and characterised for 1550 nm light detection. The two designed wafers reduce the electric field across the InGaAs absorber to a minimum in order to minimise dark current. The first wafer is designed to punch-through at the point of high breakdown probability (above breakdown voltage), while the second is designed to punch-through just under breakdown voltage. The first wafer is found to be unsuitable for single-photon counting due to an uncharacteristically fast rise in dark count rate, likely caused by the onset of punch-through during breakdown. Low photon levels are detected using diodes fabricated from the second wafer, however the diodes were found to not fully punch-through, preventing single-photon counting. Peak laser pulse detection probabilities at 150 K were 73, 71, and 46 % for 100, 30, and 10 photons, respectively. At room temperature, pulse detection probabilities were 39, 35, and 30% for the respective photon levels. This informs future SPAD designs; crucially that full punch-through must occur before breakdown voltage.

A simulation model for the sensitivity of electron APDs (e-APDs) is developed and applied to InAs e-APD based optical receivers. The model simulates bit-error rate (BER), and captures the effects of inter-symbol interference (ISI), dark current, current impulse duration, avalanche gain, and amplifier noise. With a target BER of 10^{-12}, the receivers' sensitivities were -30.6, -22.7, and -19.2 dBm for 10, 25, and 40 Gb/s data rates. The simulated InAs APDs offer improvements over existing InAlAs APDs at 10 and 25 Gb/s, however SOA-PIN based receivers outperform both types of APD for 40 Gb/s for 1550 nm operation.

Utilising the newly developed e-APD sensitivity model, and a previously developed sensitivity model for standard APDs, simulations are performed comparing InAs, InAlAs, and InP based optical receivers. The simulations utilise a common parameter set where appropriate, allowing for a direct comparison between the three avalanche materials for high-speed operation. Simulated InAs APDs achieved the best performance for 10 Gb/s operation (-29.4 dBm for InAlAs), while the simulated InAlAs APDs were found to perform better at 25 and 40 Gb/s, achieving a sensitivity of -23.5 and -21.0 dBm, respectively. InP APDs showed sensitivities of -27.9, -22.5, and -19.9 dBm, for 10, 25, and 40 Gb/s operation, respectively. These simulations demonstrate significant performance benefits to replacing InP with InAlAs as an APD avalanche region. Additional simulations of InAs APDs were performed, exploring how to further optimise InAs based optical receivers.

Publications

Journal Papers

1. **V. Shulyak**, M. M. Hayat, and J. S. Ng, ‘Sensitivity Calculations of High-Speed Optical Receivers Based on Electron-APDs’, *Journal of Lightwave Technology*, vol. 38, no. 4, pp. 989–995, Feb. 2020.

Open Datasets and Software

1. **V. Shulyak**, M.M. Hayat, and J.S. Ng “Data for paper: “Sensitivity Calculations of High-Speed Optical Receivers based on Electron-APDs””, *Figshare Data*, October 2019, doi:10.15131/shef.data.9959468.
2. **V. Shulyak**, M.M. Hayat, and J.S. Ng “Jupyter Notebook for paper: “Sensitivity Calculations of High-Speed Optical Receivers based on Electron-APDs””, *Figshare Software*, October 2019, doi:10.15131/shef.data.10069391.v1.

Conferences

1. **V. Shulyak**, L.W. Lim, C.H. Tan, and J.S. Ng, “Sensitivity of High-Speed Optical Receivers with InAs Avalanche Photodiodes”, *UK Semiconductors*, Sheffield, UK, July 2017.
2. S. Abdullah, S.J. Dimler, **V. Shulyak**, J.D. Petticrew, J.S. Ng, C.H. Tan, “Excess Bias and Time Histogram Measurements of Dark Count Rate in 200 nm Thick Geiger mode $Al_{0.85}Ga_{0.15}As_{0.56}Sb_{0.44}$ Avalanche Photodiodes”, *Photonics by the Lake*, Windermere, UK, July 2018.

3. S. Abdullah, S.J. Dimler, J.S. Ng, **V. Shulyak**, J.D. Petticrew, and C.H. Tan, "Dark Counts and Their Timing Characteristics in $Al_{0.85}Ga_{0.15}As_{0.56}Sb_{0.44}$ Single Photon Avalanche Diodes", *UK Semiconductors*, Sheffield, UK, July 2018.

Contents

Acknowledgements iii

Abstract iii

Publications vi

Contents viii

List of Figures xiii

List of Tables xix

Abbreviations xxi

Physical Constants xxiii

Symbols xxv

1 Introduction 1
- 1.1 Development of Avalanche Breakdown 2
- 1.2 APD Materials 4
  - 1.2.1 Silicon and Germanium 4
  - 1.2.2 Group II-VI Alloys 5
  - 1.2.3 Group III-V Alloys 5
- 1.3 APD Wafer Structures 7
  - 1.3.1 PIN Junction 7
  - 1.3.2 Seperate Absorption-Multiplication Region 7
- 1.4 APD Device Structures 8
  - 1.4.1 Mesa 8
  - 1.4.2 Planar 9
  - 1.4.3 Waveguide 9
- 1.5 Few and Single Photon Detection 10
  - 1.5.1 Photomultiplier Tubes 10
  - 1.5.2 Superconducting Nanowires 11
  - 1.5.3 Transition Edge Sensors 12
  - 1.5.4 Geiger Mode APDs 12
  - 1.5.5 Comparison 14
- 1.6 High Speed Optical Receivers 15

1.6.1 PIN Photodiodes 15
1.6.2 Uni-Travelling-Carrier Photodiodes 16
1.6.3 High-speed APDs 17
1.6.4 Optical Preamplifier-PIN Detectors 18
1.7 Motivation 19
1.8 Thesis Organisation 20

2 Background Theory 23
2.1 Capacitance Modelling and Punch-through 23
2.2 Impact Ionisation and Breakdown Probability 25
2.3 Tunneling Current 30
2.4 Dark Counts 31
2.5 Circuit Quenching 31
2.6 3 dB Bandwidth and the Gain-Bandwidth Product 32
2.7 Bit-error Rate 33

3 Experimental Details 35
3.1 Standard Mesa Fabrication Procedure 35
3.2 Photolithography 38
3.2.1 BPRS200 38
3.2.2 SPR350 and Polymethylglutarimide (PMGI) 39
3.2.3 SPR220 40
3.2.4 Hexamethyldisilazane 41
3.3 Wet etching trials 41
3.3.1 Trial A: H_2SO_4:H_2O_2:H_2O (1:8:80) 41
3.3.2 Trial B: H_3PO_4:H_2O_2:H_2O (1:8:1) 42
3.3.3 Trial C: $C_6H_8O_7$:H_2O_2 (2:1) (InGaAs) and HCl:H_2O (3:1) (InAlAs) 44
3.3.4 Comparison 45
3.4 Anisotropic Etch Directions 47
3.5 Characterisation Techniques 49
3.5.1 IV 49
3.5.2 Responsivity and Gain 50
3.5.3 CV 51
3.5.4 Dark Count Rate 51
3.5.4.1 CQC Board 51
3.5.4.2 JANIS ST-500 Low Temperature Probe Station and LeCroy Oscilloscope 53
3.5.5 Photon Detection Efficiency 56
3.6 Double Mesa Design and Fabrication 58
3.6.1 Design and Simulation 58
3.6.2 Device Fabrication of Double Mesa Diodes 60
3.6.3 Results of Double Mesa Fabrication 61

4 InGaAs/InAlAs Based Single Photon Avalanche Photodiodes 65
4.1 Mask Design 65
4.2 Wafer Design and Considerations 68
4.2.1 Breakdown Probability Modelling 69

4.2.2 Tunneling Effects 70
4.2.3 Electric Field Modelling 71
4.3 Wafer Growth and Device Fabrication 74
4.4 Wafer Characterisation 76
4.4.1 IV Measurements 76
4.4.2 CV Measurements 79
4.4.3 Responsivity 81
4.5 DCR and Pulsed Photon Measurements 82
4.5.1 DCR Measurements 83
4.5.2 Afterpulsing 84
4.5.3 Temperature Dependence 85
4.5.4 Pulsed Laser Measurements 87
4.6 Discussion of Results 90
4.7 Conclusion 92

5 Optimising the Sensitivity of eAPD Based Optical Receivers 93
5.1 Introduction 93
5.2 Bit-error Rate Model for APDs 94
5.3 Modification of the Model for eAPD-based Optical Receivers 96
5.4 Model Parameters 99
5.5 Simulation Results 102
5.5.1 Effects of Gain, Bandwidth, and Dark Current 102
5.5.2 Sensitivity at Different Avalanche Widths and Bit Rates 104
5.5.3 Sensitivity with Different Target BERs 107
5.6 Conclusion 109

6 A Comparison of InAs, InAlAs, and InP APDs for High Speed Optical Receiever Applications 111
6.1 Introduction 111
6.2 Gaussian Approximation Model 112
6.3 Model Parameters 115
6.4 Results and Comparisons 116
6.4.1 Bandwidth 116
6.4.2 Optimising Width 118
6.4.3 Variable BER 118
6.5 Further Optimization of InAs 119
6.5.1 Varying Diameter 120
6.5.2 Tunneling Parameter 121
6.5.3 Bandgap and Temperature 122
6.6 Conclusion 124

7 Conclusion and Future Work 127
7.1 Conclusion 127
7.2 Suggestions for Future Work 128

A Mask Set: IIG Mesa 131

A.1 Mask 1 131
A.1.1 Frame 1: Top contact 132
A.1.2 Frame 2: Standard Mesa Etch 132
A.1.3 Frame 3: Deep Mesa Etch 133
A.1.4 Frame 4: Grid Contact 133
A.2 Mask 2 134
A.2.1 Frame 1: SU-8 135
A.2.2 Frame 2: Window Etch 135
A.2.3 Frame 3: AR Coating 136
A.2.4 Frame 4: Optical Mask 136

B Mask Set: IIG SPAD 137
B.1 Mask 1 137
B.1.1 Frame 1: Top contact 138
B.1.2 Frame 2: Standard Mesa Etch 138
B.1.3 Frame 3: LC Pad 139
B.1.4 Frame 4: Isolation Etch 139
B.2 Mask 2 140
B.2.1 Frame 1: Nitride Etch 141
B.2.2 Frame 2: SU-8 141
B.2.3 Frame 3: AR Coating 142
B.2.4 Frame 4: Bond pad 142

References 143

List of Figures

1.1 The guard ring diode structure used by Goetzberger *et al.* [4]. 3
1.2 Bandgap engineering [12]. 6
1.3 Comparison of absorption coefficients for Silicon, Germanium, and InGaAs for SWIR detection [19]. 6
1.4 Example of a mesa *nipip* SAMAPD structure. 8
1.5 An example of a planar APD structure. The P^+ implantation region is visible in red [26]. 9
1.6 A basic optical communication network consisting of a laser, optical fibre, photodiode, and TIA. 15
1.7 Comparison of InAlAs and InP based APDs and their gain-bandwidth products from [58]. 18

2.1 Example of extracting electric field data from a CV graph for a SAM structure. The voltage is increased in 1 volt steps and the corresponding field profile plotted. 24
2.2 An annotated electric field distribution for a depleted *pin* structure biased to an arbitrary voltage, with dopings N_1, N_2, and N_3 for the p^+, p^-, and n^+ regions. E_1 and E_2 are the electric field values at x_1 and x_2, respectively. 24
2.3 Plot of the implementation of McIntyre's 1973 breakdown probability equations. Solid lines represent P_e values, while dashed lines represent P_h values. These correspond to their ionisation coefficient ratio $k = \beta/\alpha$. $\delta = \int_0^w \alpha dx$. 28
2.4 Visualisation of hole and electron creation. Carriers are created at x' and ionise at x after travelling a distance Δx. 28
2.5 Spatial representation of the ionisation coefficients and their corresponding ionisation probabilities, moving away from carrier creation point x'. The probability of ionisation increases with distance travelled before peaking at approximately where α and β reach their constant values α_h and β_h [72]. 29
2.6 Band diagrams showing increasing voltage and the effects on the tunnelling barriers. The barrier width decreases with increased voltage. 30
2.7 Normalised 3dB bandwidth as it changes with gain. Bandwidth reduces with gain for APDs where $k \neq 0$, while stabilising for APDs where $k = 0$. The initial drop for $k = 0$ corresponds to an increase in excess noise factor, though this never increases above 2. Data is generated from models presented in later chapters. 33
2.8 An example of the overlap (highlighted in red) between a "0" and a "1" bit, centered at μ_0 and μ_1 with standard deviation σ_0 and σ_1, respectively. The decision threshold θ is selected as the level at which BER is minimised. 34

2.9 Bit-errors caused by a detector being too slow to reset after a detection event. The signal from “1” bits from previous detection windows extends into the next. 34

3.1 Example of a pattern (a) transferred to a sample coated with photoresist. The part shown in the optical microscope image (b) is used to determine if the exposure quality is good. The “fingers” (highlighted) are well defined, indicating the correct exposure and development time. 36
3.2 Lift off procedure. (a) is how the sample appears after removal from the metallisation chamber. Placing it in acetone will lift off the metal, leaving behind a pattern as shown in (b). 36
3.3 Wet etching. The profile of a mesa after wet etching depends heavily on the etchants used. In general, the angle is not perpendicular to the substrate. 37
3.4 A sample after undergoing the lower contact metallisation-liftoff procedure. Individual mesa devices with top contacts are visible, with each cell of devices surrounded by a metal grid contact (the lower contact). 37
3.5 Deposition and removal of dielectric from contact areas and mesa window. 38
3.6 PMGI-assisted lift-off. 40
3.7 SEM images of trial A. 43
3.8 SEM images of the sample from trial B. 44
3.9 SEM images of the sample from trial C. 46
3.10 SEM images of mesa sidewalls in the two cleave directions. (a) and (b) show the 011b etch direction. (c) and (d) show the 011 etch direction. (e) and (f) show the transition of mesa sidewall from 011b to 011. 48
3.11 (a) Linear and (b) semi-log plots of the IV relation for an ideal diode plotted with arbitrary units. The reverse saturation current is set to 3×10^{-9}. 50
3.12 Experimental setup for measurements of DCR for packaged devices. . . . 52
3.13 A demonstration of the biasing pulse, transients, and signal from the diode. 54
3.14 Traces of breakdown events recorded from the oscilloscope, with the transients indicated by the black line. 55
3.15 LeCroy filtering (see text for details). 55
3.16 Double mesa simulations using the structure given in table 3.5 with the top mesa edge 10 μm away from the outer mesa. 59
3.17 SEM images of a double mesa diode fabricated without using HMDS. Extreme undercutting of the photoresist has brought the wider mesa edge close to the top mesa, and led to a rough sidewall. 61
3.18 Microscope (top row) and SEM images of successfully fabricated double mesa diodes, coated with HMDS prior to the second etch step. For reference, the diodes shown in the top row have diameters of 70 and 120 μm, with the 70μm diode labelled. 62
3.19 Comparison of IV data from fabricated single (green) and double (red) mesa diodes. 63

4.1 Overview of the designed mask set. Dimensions of the whole mask cell are 1260×1360 μm. Some layers are omitted for clarity. (a) shows the overall design, while (b) shows the cell IDs. A close up of a 180×340 μm D30 diode cell is shown in (c) with various dimensions indicated. 66
4.2 IIG Mesa mask set. 67

4.3 (a) and (b): Simulated avalanche breakdown probabilities for perfect InAlAs *pin* structures with widths range from 0.5 μm to 1.5 μm in 0.1 μm steps. Solid lines are P_{be} and dashed lines are P_{bh}. (c) Breakdown voltages and fields extracted from (a) and (b) at $P_{\text{be}} = 0.01$. 69

4.4 Modelled tunneling current densities versus (a) reverse bias and (b) reverse electric field for InAlAs *pin* diodes with i region widths from 0.5 to 1.5 μm in 0.1 μm steps. (c) Tunneling current density versus avalanche width at $V_{\text{bd}} \times 1.1$ - a typical operating voltage for a SPAD. 70

4.5 Modelled tunneling current densities versus (a) reverse bias and (b) reverse electric field for InGaAs *pin* diodes with i region widths ranging from 0.5 μm to 1.5 μm in 0.1 μm steps. 71

4.6 (a) demonstrates the effects of varying N_c with a fixed $w_c = 69$ nm. (b) demonstrates the effects of varying w_c with $N_c = 4.1e17$ cm^{-3}. V_p and V_{bd} are extracted from the CV model. $V_{\text{p,start}}$ and $V_{\text{p,end}}$ represent where the absorption region starts to deplete and fully depletes, respectively. . . 72

4.7 Modelled CV of the (a) SF1318 and (b) SF1319 structures as designed. The two voltages indicated in (b) and (d) are the voltages at which the InGaAs absorber begins to deplete and fully depletes, respectively. 72

4.8 Extracting avalanche field from CV modelling. 73

4.9 (a) shows an image of a D25 diode from the IIG SPAD mask after the final fabrication step, (b) is a cross section diagram of the diodes, and (c) is an SEM image of a D60 diode cleaved as in the cross section diagram. . 75

4.10 IV measurements of SF1318 diodes before nitride deposition. 77

4.11 IV measurements of SF1319 diodes before nitride deposition. 77

4.12 Current density measurements of (a) SF1318 and (b) SF1319 diodes before (blue) and after (orange) nitride. 78

4.13 (a) SF1318 and (b) SF1319 temperature dependent IV measurements to determine breakdown voltage temperature dependence. 78

4.14 (a) CV data of SF1318 diodes with nominal diameters of 420 and 220 μm. There is an indication of the start of punchthrough, shown in the inset. The capacitance density data in (b) has a correction of -8 μm applied to the diameter. 79

4.15 (a) CV data of SF1319 diodes with diameters of 420 and 220 μm. A correction of -8 μm is applied to the diameter for the area calculation in (b). 79

4.16 Punch-through and breakdown probabilities predicted for SF1318 and SF1319 based on IV and CV measurements. Orange and green highlighted areas indicate the voltages across which the absorption regions of SF1318 and SF1319 are depleting, with an overlapping area in the centre. 80

4.17 SF1318 responsivity data for two D220 diodes, including the effects of avalanche gain. 81

4.18 Responsivity data for SF1319 obtained from two D220 diodes. (a) shows responsivity on a linear scale while (b) shows it on a log scale to provide more clarity in the low gain region. The responsivity shown includes gain, since it is difficult to define a region of gain 1. 82

4.19 Pulsed DCR for SF1318 diodes. 83

4.20 Pulsed DCR for SF1319 diodes. 83

4.21 Comparison of DCR for SF1318 and SF1319, with D20 and D25 diodes plotted together. 84

4.22 Double biasing pulse setup used for assessing afterpulsing. 84

4.23 Double pulsing DCR measurements for SF1318 and SF1319. Two 10 ns, 4 V pulses were applied with the seperation between the end of the first pulse and the start of the second pulse indicated on the x-axis. The DC reverse bias was set to 62.2 V and 64.0 V for SF1318 and SF1319 respectively. . . 85

4.24 Temperature dependence of DCR in a D20 SF1318 diode. The measurements were performed with a 1.5 ns wide 5 V pulse at a repetition frequency of 100 kHz. 86

4.25 Temperature dependence of DCR in a D20 SF1319 diode. The measurements were performed with a 5 ns wide 10 V pulse at a repetition frequency of 100 kHz. 86

4.26 The detection probability of (a) D30, (b) D25, and (c) D20 diodes for 1550 nm pulses with n photons per pulses as indicated in the legend. . . . 88

4.27 The detection probability of a D20 diode at (a) 295 K, (b) 250 K, (c) 200 K, (d) 150 K for different photon levels, with a 5 dB increase in attenuation between each level. 89

5.1 The calculated mean impulse current for an e-APD with $v_e = 5\times10^5\ ms^{-1}$, $v_h = 0.4\times10^5\ ms^{-1}$, and $w = 3\ \mu m$ reverse biased at 18 V. The important times are indicated, where the mean impulse changes shape. 97

5.2 Field-dependent electron drift velocity simulations of InAs [86], and experimental field-dependent hole drift velocity of GaAs [87]. The lines shown are the fittings obtained and used in the simulation of InAs. 100

5.3 Verification of the integral of equation (5.9) from $t = 0$ to $t = T_f$ using QUADPACK integration, against RPL simulations for a 3.0 μm InAs APD. 101

5.4 Varying θ for 14, 15, and 16 V at a fixed photon flux. 101

5.5 Flowchart for calculating sensitivity at a given avalanche region width w, voltage V, bit rate R_b, and target BER. In practice this is part of a larger process in which V, w, R_b, and target BER are also varied. 103

5.6 Gain, 3 dB bandwidth, dark current, and sensitivity simulations with $w = 3.0\ \mu m$. The optimal operating voltage is at -18 V for a 10 Gb/s bit rate and 1×10^{-12} BER. 104

5.7 Gain - bandwidth properties of three simulated InAs e-APDs. The bandwidth is largely independent of gain, other than an initial increase due to the field dependent drift velocities. 105

5.8 Sensitivity versus APD reverse bias for several avalanche widths. The data rate is 10 Gb/s, and the target BER is 1×10^{-12}. 106

5.9 Gain-optimised sensitivity simulations for a range of different avalanche widths, performed for 10, 25, and 40 Gb/s bit rates with a target BER of 10^{-12}. 106

5.10 Variable target BER versus optimum sensitivity for InAs APD-TIA combinations for (a) 1550 nm and (b) 1310 nm wavelength operation. 108

6.1 Block diagram of the simulation components and how they interact to calculate sensitivity. 116

6.2 A comparison of the gain-bandwidth properties of the three materials. An avalanche width of 0.11 μm is used for InP and InAlAs, while 1.5 μm is used for InAs. Due to the excess noise characteristics of InAs, it is able to overcome the gain-bandwidth limitation of regular APDs. In Figure (b) the solid lines represent 3dB bandwidth, and dashed lines represent the gain-bandwidth product. . . . 117
6.3 BER simulations for InAlAs and InP at different widths for a target BER of 1×10^{-12}. InAs results are also indicated at as green dots, though note that these are actually at their respective optimum widths of 3.0, 2.3, and 1.5 μm for 10, 25, and 40 Gbps respectively. . . . 118
6.4 Comparison of the optimal sensitvities of InP, InAlAs, and InAs diodes for different target BERs. InP $w = 0.14$, 0.11, and 0.10 μm. InAlAs $w = 0.12$, 0.10, and 0.09 μm. 10, 25, and 40 Gb/s are represented by solid, dashed, and dot-dashed lines, respectively. . . . 119
6.5 Sensitivity simulations with a fixed avalanche width $w = 1.5\mu m$, while varying diode diameter. . . . 120
6.6 Effects of varying σ_{tunn} for $w = 1$ μm on tunneling current. (a) shows raw tunneling current, (b) shows the reverse bias at which 1 μA of tunneling current is reached, where a significant impact on sensitivity begins to occur. 121
6.7 Sensitivity simulations with a fixed avalanche width $w = 1.5\mu m$, while varying σ_{tunn}. . . . 122
6.8 Effects of varying E_g for $w = 1.5$ μm on tunneling current. . . . 122
6.9 Sensitivity simulations with a fixed avalanche width $w = 1.5\mu m$. . . . 123
6.10 Comparison of InAs 40 Gbps sensitivity simulations at 77 K and 293 K. . 124

A.1 IIG mesa mask plate 1 . . . 131
A.2 Top contact deposition layer. . . . 132
A.3 Etching layer to define mesas. . . . 132
A.4 Etching layer to define mesas, designed for depths where undercutting would reach the top contacts if using the standard etch. . . . 133
A.5 Lower contact deposition layer. . . . 133
A.6 IIG mesa mask plate 2 . . . 134
A.7 Exposes SU-8, used to passivate the diodes. . . . 135
A.8 Etch the windows of the mesa diodes. For use with diodes where the top layer absorbs too much light. . . . 135
A.9 Remove AR coating nitride from all areas other than the mesa windows. . 136
A.10 Layer to deposit a metal optical mask, used to prevent side injection into the mesa sidewalls. . . . 136

B.1 IIG SPAD mask plate 1 . . . 137
B.2 Top contact deposition layer. . . . 138
B.3 Etching layer to define mesas. . . . 138
B.4 Lower contact bond pad deposition layer. . . . 139
B.5 Etching layer to isolate the diodes from each other, and to remove the contact layer from the top bond pad area. . . . 139
B.6 IIG SPAD mask plate 2. . . . 140
B.7 Remove dielectric from device windows, contacts, and between diode cells. 141
B.8 Exposes SU-8, passivating the diode sidewalls. . . . 141
B.9 Remove AR coating nitride from all areas other than the mesa windows. . 142

B.10 Deposition layer for the top contact bondpad. 142

List of Tables

1.1 Comparison of single-photon detectors. 14

3.1 Layer profile of wafer SF0935, used for the wet etching trials. The p+ and n+ regions are highlighted. 42
3.2 Comparison of etchants from the trials. 45
3.3 SF0751 concessionairy wafer structure. 47
3.4 Summary of the etch angles determined from the SEM images. Angles are taken from the horizontal plane. 49
3.5 SF0940 wafer structure, used for simulation and fabrication of double mesa APDs. 58

4.1 SF1318 and SF1319 wafer structures as designed (charge sheet doping is later confirmed through CV measurements, and is different than the design). 73
4.2 SF1318 and SF1319 wafer structures based on CV data fitting. 80

5.1 Model parameters for InAs e-APDs. 100

6.1 Parameters used to simulate InP and InAlAs APD sensitivity performance. 115

A.1 Overview of IIG mesa mask 1 layers. 131
A.2 Overview of IIG mesa mask 2 layers. 134

B.1 Overview of IIG SPAD mask 1 layers. 137
B.2 Overview of IIG mesa mask 2 layers. 140

Abbreviations

AC	Alternating Current
APD	Avalanche PhotoDiode
BER	Bit-Error Rate
CQC	Charge Quenching Circuit
CW	Continuos Wave
DC	Direct Current
DCR	Dark Count Rate
DUT	Device-Under-Test
e-APD	Electron Avalanche PhotoDiode
EDFA	Erbium-Doped Fiber Amplifier
EPSRC	Engineering and Physical Sciences Research Council
IIG	Impact-Ionisation Group
HMDS	HexaMethyldDiSilazane
ISI	Inter-Symbol Interference
LIA	Lock-In Amplifier
LIDAR	Light Detection And Ranging
LWIR	Long-Wave InfraRed
MCT	Mercury Cadmium Telluride
MWIR	Mid-Wave InfraRed
NEP	Noise Equivalent Power
NIM	Nuclear Instrumentation Module
NbN	Niobium Nitride
PECVD	Plasma-Enchanced Chemical Vapor Deposition
PMGI	PolyMethylGlutarImide
PMT	Photo-Multiplier Tube

TIA	Trans-Impedance Amplifier
RE-SPAD	Red-Enhanced SPAD
RIE	Reactive-Ion Etching
RPL	Random Path Length
RPM	Rotations Per Minute
SAMAPD	Seperate-Absorption-Multiplication Region APD
SEM	Scanning Electron Microscope
Super-K	Super-Kamiokande
SMU	Source-Measure Unit
SNR	Signal-to-Noise Ratio
SNSPD	Superconducting Nanowire Single Photon Detector
SOA	Semiconductor-Optical Amplifier
SPAD	Single-Photon APD
SPDE	Single-Photon Detection Efficiency
SWIR	Short-Wave InfraRed
TES	Transition-Edge Sensor
UTCPD	Uni-Travelling-Carrier PhotoDiode
UV	Ultra-Violet
QE	Quantum Efficiency
QKD	Quantum-Key Distribution

Physical Constants

Boltzmann Constant $k = 1.380\ 648\ 52 \times 10^{-23}\mathrm{m^2kgs^{-2}K^{-1}}$

Elementary Charge $q = 1.602\ 176\ 62 \times 10^{-19}\ \mathrm{C}$

Electron Mass $m_e = 9.109\ 383\ 56 \times 10^{-31}\ \mathrm{kg}$

Permittivity of Free Space $\varepsilon_0 = 8.854\ 187\ 81 \times 10^{-12}\ \mathrm{Fm^{-1}}$

Speed of Light $c = 2.997\ 924\ 58 \times 10^{8}\ \mathrm{ms^{-S}}$

Planck's Constant $h = 6.626\ 070\ 04 \times 10^{-32}\ \mathrm{kg/s}$

Symbols

A	area	m^{-2}
C	capacitance	F
d	deadspace	m
d_e	electron deadspace	m
d_h	hole deadspace	m
E	electric field	Vm^{-1}
E_g	band-gap	eV
E_{pulse}	pulse energy	eV
E_{photon}	photon energy	eV
f	repetition frequency	Hz
I	current	A
I_s	reverse saturation current	A
I_{tunn}	tunneling current	Am^{-1}
k	impact ionisation coefficient ratio	
m_e^*	effective electron mass	kg
n_i	ideality factor	
$\bar{n}$	average number of photons	
N	doping concentration	cm^{-1}
P	probability of a counting event	
P_d	dark count probability	
P_e	breakdown probability of electrons	
P_h	breakdown probability of holes	
P_{se}	electron survival probability	
P_{sh}	hole survival probability	
P_{pair}	combined breakdown probability	

P_{pde}	pulse detection probability	
V	bias voltage	V
V_t	thermal voltage	V
w_d	depletion width	m
α	electron ionisation coefficient	cm^{-1}
β	hole ionisation coefficient	cm^{-1}
ε_r	relative perimittivity	
ω	angular frequency	rads^{-1}
σ_0	photon flux	
σ_T	tunneling parameter	

Chapter 1

Introduction

Global demand for digital communication bandwidth has risen rapidly over the past two decades. To meet this demand, new and novel techniques have to be developed to efficiently and frequently increase capacity of communication systems. Optical fibre-based communication systems form part of this essential infrastructure. Avalanche photodiodes (APDs) have long been a vital component in long distance transmission of signals, being used to detect light transmitted through optical fibres. Thanks to their high internal gain, APDs increase the distance a signal is able to be transmitted before having to be retransmitted, and do so in an energy efficient manner by reducing the need for external amplifier circuitry. In particular InAlAs/InGaAs based APDs are a well established technology for the detection of near infra-red light in the optical window wavelength range (1310-1550 nm).

For high-speed detection, an emerging III-V based material, InAs, has been demonstrated to be of great interest. It has high electron drift velocities, and is one of the only materials able to be used in the fabrication of eAPDs (electron APDs), where only the electrons undergo impact ionisation. Such characteristics make it worth investigating for high speed optical communication applications.

APDs have evolved a great deal over the years from the discovery of the avalanche breakdown effect in a Si-Ge p-n junction by McKay in 1954 [1]. An APD at its most basic is a p-n junction operated in reverse bias. By operating the junction at a high reverse bias, it exhibits a signal gain effect. Past the junction breakdown voltage, an incident photon will cause a self-sustaining current to form, known as avalanche breakdown. This mechanism allows for large internal signal amplification, with a rapid increase of current from the nano-amp level up to hundreds of milli-amps. Due to this, APDs are detectors of choice for detecting weak signals where photon numbers are limited.

This chapter will explore the development of APDs from their beginnings in the 1960s, to more recent advances. The materials used and the development of the structures of the devices themselves will also be discussed. Reviews of single photon detection methods, and high speed optical receivers will also be given. More detailed theory behind their operation will be presented in chapter 2.

1.1 Development of Avalanche Breakdown

The avalanche breakdown effect in semiconductors was first noted by McKay in 1954. He observed a pulse-type noise from breakdown events which were originating from microplasmas in local regions of Si-Ge *pn* junctions, with the effect increasing with increasing reverse bias [2]. In 1960, Batdorf *et al.* studied this effect in greater detail by studying small Si p-n junctions [3]. At half the breakdown voltage they found that these microplasmas were occurring at field inhomogeneities and lattice inhomogeneities. By using p-n junctions without exposed edges and avoiding dislocations in the charge region, they reduced microplasma effects allowing the devices to operate at higher reverse biases. A more diffused, uniform light emission from the whole junction, termed a "macroplasma", was observed. It was shown that microplasmas didn't have to accompany the avalanche effect, but the presence of microplasmas does reduce the breakdown voltage [3].

Following the observations of these microplasmas in Si p-n junctions, Haitz *et al.* further studied how these microplasmas affected the breakdown voltage [1]. They targeted small areas of microplasmas using a light spot, and measured photocurrent as a function of reverse bias. At lower voltages the microplasmas had a negligible effect on the multiplication. They also found that the microplasmas were being caused by defects on the surface of the semiconductor rather than within the bulk structure. One of his key observations however, was that the breakdown voltage was not lowered by microplasma effects, contradicting Batdorf's previous conclusion. The main difference was the use of a

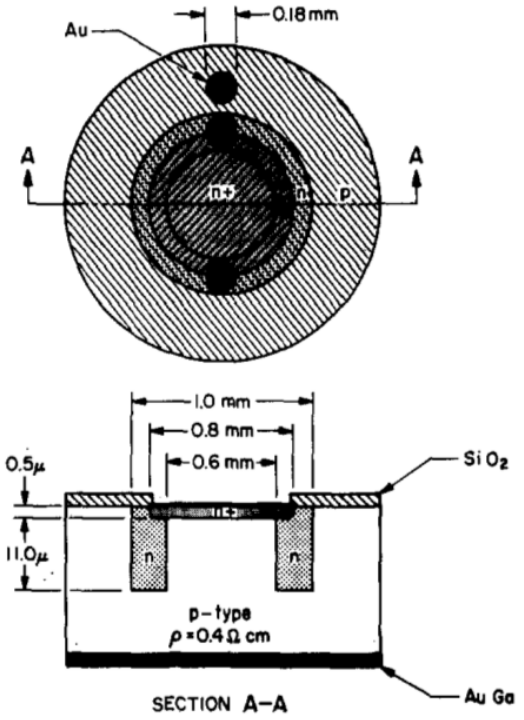


FIGURE 1.1: The guard ring diode structure used by Goetzberger *et al.* [4].

guard ring around the diode (figure 1.1), which eliminated most of the effects of surface breakdown allowing for plasma-free bulk breakdown [4].

Later, K. M. Johnson demonstrated the signal-to-noise ratio (SNR) enhancing effects of avalanche breakdown. They performed a detailed analysis of a planar Si photodiode, characterising the SNR bandwidth and Noise Equivalent Power (NEP) in avalanche conditions. They showed a 13 dB gain in SNR resulting from avalanche multiplication of the signal [5].

With the rapid development in the area by experimentalists, groups sought to theoretically model the avalanche effect to allow for a more quantitative description of the breakdown effect. In 1972 Oldham *et al.* formulated equations to describe the ionisation probability of carriers in a device [6], which referenced previous attempts by McIntyre to theoretically describe microplasma effects [7] and avalanche noise [8]. The two coupled differential equations, (3) and (4) in [6], could be used to model the breakdown probability of an arbitrary junction, given that the carrier ionisation coefficients were known for the material.

McIntyre took these equations and simplified them, providing an analytical solution for modelling breakdown probability [9]. He also formalised the gain theory of photodiodes, and determined the probabilities for non-detection and false signals which provided a way to calculate how many photons were required in a light pulse to guarantee detection. This theory could be used to determine the effectiveness of an APD at photon counting. The theories were confirmed experimentally by Conradi [10].

1.2 APD Materials

There are four main material systems used for fabricating APDs; silicon, germanium, II-VI, and III-V. For completeness, the former three will be discussed briefly though this project uses exclusively III-V alloys.

1.2.1 Silicon and Germanium

Silicon is the material of choice for most APD detection applications with wavelenghts below 1000-1100 nm. It is the cheapest semiconductor material with a large manufacturing base, and has many of the best properties. These include a low amount of noise, dark current, and small capacitance [11]. It is not effective at wavelenghts above 1000 nm due to its 1.1 eV indirect bandgap [12]. In addition, photon absoprtion in Silicon is significantly worse as the photon wavelength approaches that of the bandgap edge due to its indirect nature, since a phonon is required for the electron to be promoted. This is less likely to occur as photon energy decreases. This rules it out of being useful in the optical window of 1300 - 1580 nm, which is crucial for optical fibre communications.

Germanium is commonly employed as an absorption material for silicon based APDs. With an indirect bandgap of 0.62 eV and a direct bandgap of 0.80 eV [13], Germanium is able to absorb photons with wavelengths up to 1550 nm efficiently, and with reduced efficiency up 2000 nm. However, the growth of Germanium on Silicon provides a number of challenges to detector performance. Due to a lattice mismatch of around 4 % between Silicon and Germanium, dislocations can form due to strain in the lattice [13]. To overcome these issues, a SiGe buffer layer is commonly employed at the Si/Ge interface [14]. Additionally, due to the small bandgap, Germanium introduces a significant dark current to the device, which when coupled with the dislocations and defects generally leads to a higher dark current than comparable III-V systems [13].

1.2.2 Group II-VI Alloys

The most commonly used group II-VI alloy for APDs is HgCdTe (MCT). Due to applications in the defence industry, publicly available information on MCT APDs is more limited compared to III-V and Si/Ge APDs. MCT provides high gain, with low excess noise characteristics, owing in part to electron only impact ionisation [15]. By varying the concentration of Hg and Cd, the bandgap of MCT is tunable allowing for photon detection from short-wave infrared wavelenghts (SWIR) up to long wave infrared (LWIR) wavelengths of 30 μm [16]. MCT is also used in space applications, where a tolerance to high levels of radiation is required [17]. Due to a weak bond between the Hg and Te, the material is more tolerant of impacts from high-energy cosmic rays when compared to more strongly bonded materials [17].

The weak Hg-Te bond has disdvantages however, and can lead to non-uniformity across MCT wafers [17, 18]. At relatively low temperatures, the two disassociate which causes Hg vacancies to form [18]. This limits the industrial manufacturing base due to low yields, and drives up costs. Additionally, MCT suffers from high dark current characteristics, an order of magnitude worse than those observed in comparable III-V systems [17].

1.2.3 Group III-V Alloys

The group of III-V semiconducting alloys are a versatile alternative to germanium for detection in the optical communication window. Since they are alloys, bandgap engineering can be used through the mixing of different concentrations of elements from each group, to suit the desired application. Figure 1.2 shows how mixing the different alloys can produce new alloys with a desired bandgap [12]. Many of these alloys can be realised with established growth methods, allowing for fine tuning of the structure. This means that complex structures are possible, engineering unique layers into the bulk material from multiple different lattice matched alloys all into one diode depending on the application needed.

As demonstrated in figure 1.2, there are a range of III-V materials, many of which are in active use. For SWIR detection, InGaAs (with composition $In_{0.53}Ga_{0.47}As$) is the most commonly utilised absorber with a cutoff wavelength of 1.7 μm and has superior

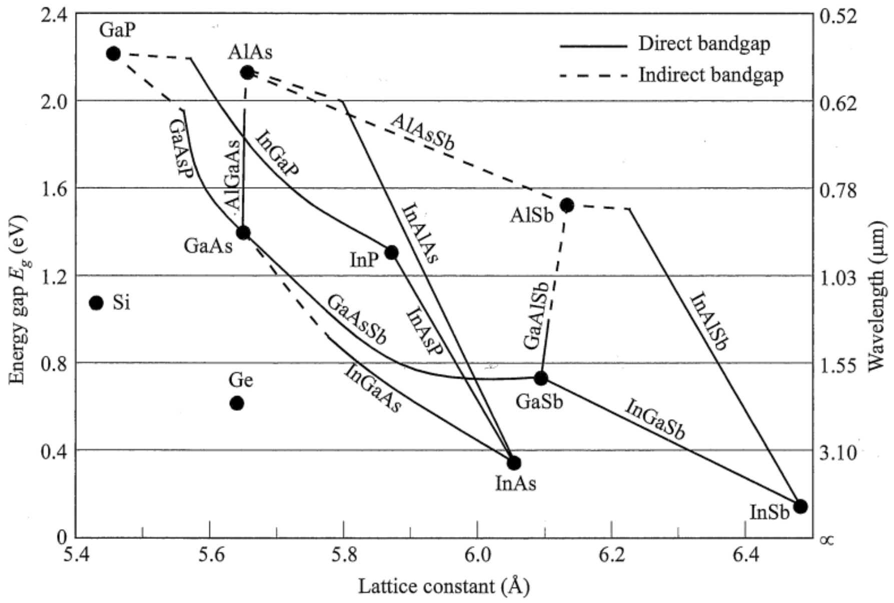


FIGURE 1.2: Bandgap engineering [12].

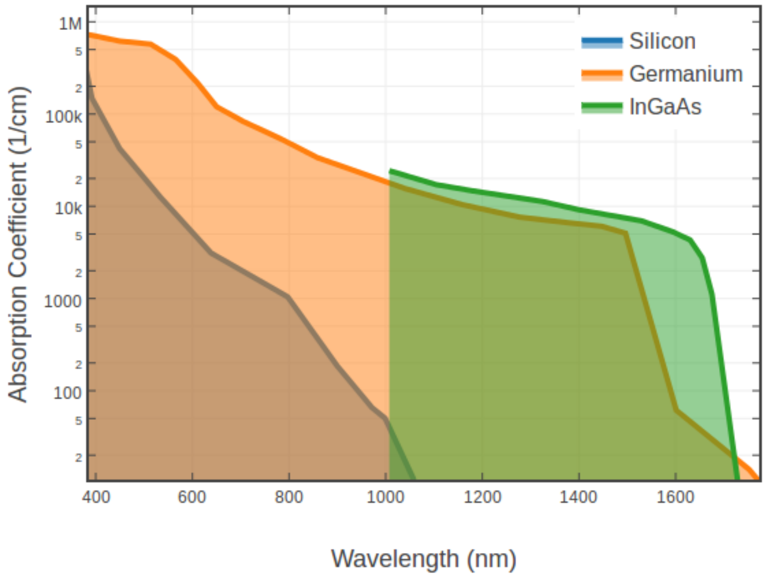


FIGURE 1.3: Comparison of absorption coefficients for Silicon, Germanium, and InGaAs for SWIR detection [19].

absorption characteristics to Ge (see figure 1.3 for comparison). For MWIR, InAs and InSb with cutoff wavelengths of 3.5 and 5.5 μm, respectively, are used [20]. InGaAs is commonly combined with materials InP or InAlAs to for APDs, due to lattice matching between these materials.

InP and InAlAs are widely utilised as avalanche materials for weak electrical signals, offering high gain with low dark current when compared to the more narrow bandgap materials such as InGaAs which can also provide high gain, but will be accompanied by high dark currents [21, 22]. InAlAs is gradually replacing InP in active use due to superior ionisation coefficients than the latter, with a greater disparity between electron and hole ionisation leading to reduced excess noise.

1.3 APD Wafer Structures

1.3.1 PIN Junction

The most basic structure an APD can have is the p-n junction, as in the first APD developed in the 1960s [5]. The intrinsic region acts as both the absorber and the multiplier, so the *pin* device is simple, and relatively easy to model. Although they are cheaper to manufacture than more complex structures, it is not easy to optimise the absorption characteristics independently of avalanche characteristics, and vice versa. These types of APD suffer from large dark current, which limits gain values due to an increasing dark current at higher operating voltages overpowering the signal. This occurs due to the quantum tunneling effects of high electric fields needed for breakdown [23].

Limiting factors such as these led to the development of more advanced structures - separate absorption-multiplication region APDs (SAMAPDs).

1.3.2 Seperate Absorption-Multiplication Region

SAMAPDs were first developed in the late 1970s by Nishida *et al.* who found that by separating the absorption region from the multiplication, the APDs achieved a lower dark current at high gains [21]. Absorbed photons generate electrons and holes in the absorption layer, which are swept into the avalanche region where impact ionisation occurs. The design in [21] was hampered by a build up of charge at the heterojunction interface between regions. Campbell *et al.* solved this issue by growing grading layers with an intermediate bandgap to smooth out the charge barrier at the interface [23]. Soon after, Capasso *et al.* showed that by tailoring the absoprtion region to have a low electric field relative to the multiplication region, it was possible to lower dark current, excess noise, and operating voltage [24]. These devices were superior to the old p-n junction APD designs in all ways except in complexity and cost.

1.4 APD Device Structures

1.4.1 Mesa

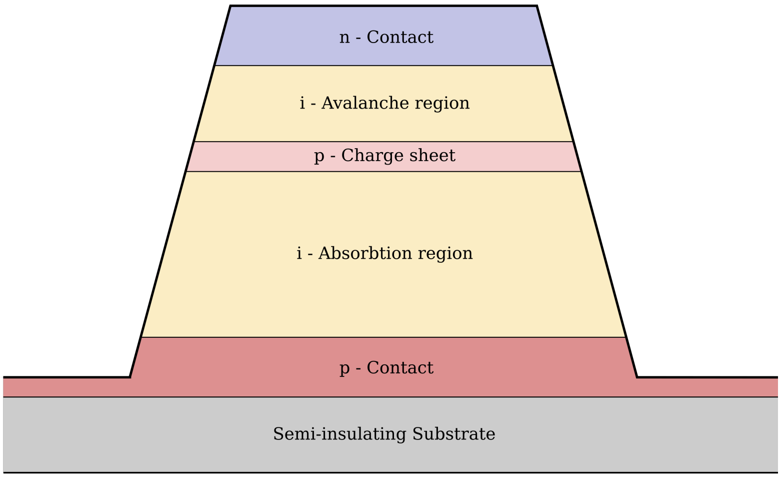


FIGURE 1.4: Example of a mesa *nipip* SAMAPD structure.

A mesa diode has an appearance like a geographical mesa, and is defined by etching the side walls of the device down to the bottom contact layers, as shown in figure 1.4. Mesa diode designs are commonly used for APDs made from III-V materials. This is due to the epitaxial growth of the layers allowing for fine control of each layer's doping, which reduces the need for extra dopant implantation steps. This means that it is relatively simple to create a device from a III-V wafer - isolating individual devices via etching.

This simplicity comes at a cost, however. Exposing the sidewalls of the device means that strong electric fields form there, resulting in surface and edge breakdown which increases dark current [25]. Additionally, the etching process can contaminate and damage the surface of the device, further reducing device performance [26].

More complicated structures have been devised to reduce these undesirable effects. Triple mesa InGaAs/InAlAs structures have been developed to suppress the electric field at the edges of the multiplication medium, eliminating surface recombination current as well as edge breakdown [27]. This requires a top contact layer with edges etched away from the second, edge-field buffer layer, which itself is etched with edges seperated from the third mesa below (which contains the avalanche and absorption layers), the electric field is mostly confined to the area defined by the top mesa. These structures, however, incur increased growth and fabrication complexity, and therefore cost.

1.4.2 Planar

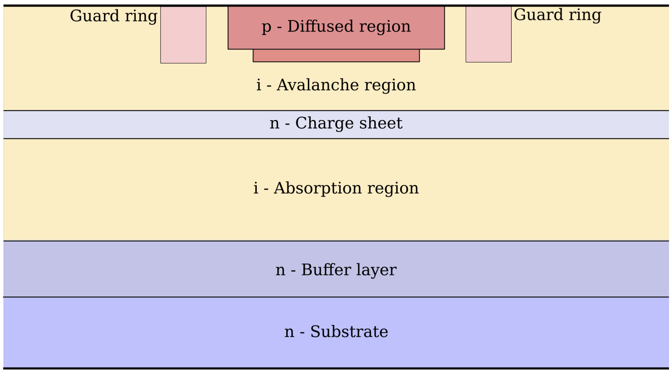


FIGURE 1.5: An example of a planar APD structure. The P^+ implantation region is visible in red [26].

Planar devices remove the need for etching, instead relying on localised dopant implantation or diffusion into a wafer's upper intrinsic region to define devices. This creates individual devices which are electrically isolated from each other by the intrinsic region.

Using planar fabrication it is possible to mass produce large quantities of uniform performance devices. The elimination of the exposed edges removes the issue of edge breakdown, allowing for higher operating voltages. This method, however, can require more fabrication steps which increases complexity and cost. An additional issue is that extra processing is needed to ensure there is no cross-talk between devices, since they are not physically isolated as in the mesa case [26, 28].

1.4.3 Waveguide

A waveguide is a structure on a semiconductor chip which allows for the routing of an optical signal along the chip to a target destination. Applications of waveguides are numerous, including the integration of lasers, receivers, and detectors onto a single chip in the context of optoelectronic circuitry [29]. Waveguides allow for the injection of light from the side of the target structure. For APDs, waveguides are utilised mainly for high-speed detection applications, where layer widths are small, to minimise the bandwidth limit imposed by carrier transit time. By using the waveguide to guide the light in through the side of the structure, the available absorption length increases without impacting detector speed.

In practice this is complicated by waveguide loss mechanisms, which include absoprtion, scattering, and leakage losses [29]. Absorption losses are mainly related to free carriers. In addition, absorption due to defects within the semiconductor can cause losses which can be exagerated in semi-insulating substrates. Dislocations in the crystal lattice can have a similar effect.

Also, scattering losses arise primarily due to rough boundaries between eplilayers. Leakage losses are those that result from bends in the waveguide due to a change in angle of incidence and depends on the refractive index difference between the semicondutor and surrounding media [29].

1.5 Few and Single Photon Detection

The detection of small amounts of photons is required for a range of important applications. These include areas such as quantum computing [30], quantum key distribution (QKD) [30], Raman spectroscopy, and a major recent area is light detection and ranging (LIDAR) for use in self-driving cars [31]. APDs operated in the Geiger mode are commonly used for such applications, though there are other devices available. This section explores those, including Geiger-mode APDs.

1.5.1 Photomultiplier Tubes

The first photomultiplier tube (PMT) was reported by Iams and Salzberg [32] in 1935, and allowed for the detection of the secondary emission of a photo-electron from the incidence of photons onto a large photocathode surface. This secondary photo-electron accelerates through a vacuum tube and passes through a cascade of dynodes which further stimulates the secondary emission of electrons. This was the first example of a detector capable of detecting light at a single-photon level, and the technology is still widely used in applications where photons need to be detected over a large area. For example, the Super-Kamiokande (Super-K) neutrino detector was a tank filled with water surrounded by PMTs for a combined mass of 50 000 tonnes, which would detect the weak Cherenkov light emitted as a neutrino passed through the water tank [33].

While PMTs are effective where large areas need to be covered, they are too bulky for use in most other applications. In addition, they require a large supply voltage in excess of 1 kV and are fragile, best demonstrated by the cascading failure of the Super-K detector where the implosion of one PMT resulted in the total loss of all the other 11,145 PMTs [33].

1.5.2 Superconducting Nanowires

Superconducting nanowire single-photon detectors (SNSPDs) utilise the Cooper pairs - bonded pairs of electrons in the superconductor - which flow through the nanowire without scattering, and therefore no resistance [34]. Cooper pairs have a binding energy of order meV. When a photon impacts a Cooper pair with an energy greater than the Cooper pair's binding energy, the pair is broken and superconductivity lost. This creates a localised hotspot, with a resistive barrier forming across the nanowire, causing a sudden current drop which can be detected by external readout circuitry, registering a photon count.

Niobium nitride (NbN) is the most used superconducting material for SNSPDs. This is due to a fast response time and a high (for superconductors) critical temperatue of 16 K [35]. The detectors themselves typically consist of a meandering nanowire which is thin (∼1-10 nm), narrow (∼100 nm), and lengths as long as 500 μm, though increased length limits detector performance due to an increased inductance increasing the recovery time [36].

The detection efficiency of NbN SNSPDs is relatively low, around 5 %, due to losses from coupling light into the small nanowire structure [36]. More complex structures are required to increase detection efficiency. Rosfjord *et al.* showed an NbN SNSPD with an SPDE of 57% at a 1550 nm wavelength [37], which was achieved through the combined use of a nanoelectrophotonic device consisting of an anti-reflection coating and an optical cavity around the SNSPD. The record SPDE attained at 1520 nm using an SNSPD was by Marsili *et al.*, who produced an amporphous tungsten silicide based SNSPD with a maximum SPDE of 93 %, though at an operating temperature of 2 K [38].

1.5.3 Transition Edge Sensors

Transition edge sensors (TESs) are highly sensitive detectors that can resolve the number of photons in a light pulse, down to the single photon level. Their operating principle is similar to that of SNSPDs, with a superconducting nanowire operated under bias below its critical temperature which is used to absorb the photons. The difference is that a circuit external to the nanowire is used to measure the temperature change that occurs when the superconductor transitions away from its superconducting state in the nanowire. The amount of thermal energy generated is proportional to the number of photons incident on the detector, allowing for the number of photons in the light pulse to be determined. A cold bath is connected to the absorber, which rapidly cools the detector back below its critical temperature.

TES detectors are the most sensitive single-photon detectors available, with record SPDEs reported of 95% and 98% at 1556 nm and 850 nm wavelengths, respectively for a titanium based TES [39]. This high SPDE comes at a great cost, with the detector requiring advanced cooling systems to operate below 300 mK. In addition, they suffer from high timing jitter on order of 100 ns, and require 4 μs to cool down after a detection event [40]. This places a limit on the data rate for the detection system (to $<$ 100 kHz).

1.5.4 Geiger Mode APDs

Geiger mode APDs - also known as single-photon avalanche photodiodes (SPADs) - are APDs that are operated above breakdown voltage. At these voltages gain becomes “infinite”, with any impact ionisation event causing a self-sustaining avalanche current which can be detected by external readout circuitry. To prevent catastrophic breakdown, where the diode is destroyed by this self-sustaining exponentially increasing current, external quenching circuitry is essential to drop the applied reverse bias and quench the avalanche current to a safe level.

Si based SPADs dominate the area of visible single photon detection. They offer high SPDEs of up to 70%, with extremely low dark counts as low as 10 Hz and large detection areas at room temperature [41, 42]. Si SPADs are able to achieve this due to their wide bandgap and highly developed manufacturing base which minimises crystal defects, with the added benefit of relatively low-costs. Their limitations appear at longer wavelengths, with a rapid fall-off in SPDE as photon energy approaches the bandgap (as discussed in 1.2.1). Red-enchanced SPADs (RE-SPADs) attempt to use relatively thick absorption layers to improve absorption efficiency, boosting SPDE from 15 to 40 % at 800 nm at the cost of increased timing jitter [42]. However, even RE-SPADs have sub 5% SPDE as wavelength passes 1000 nm [42].

Ge has been used as the infrared photon absoprtion layer in a SAM structure for Si based SPADs. These, however, suffer from high dark count rates from the Ge layer necessitating their cooling to operate effectively, though this lowers detection efficiency at 1550 nm due to an increased bandgap. Ge/Si SPADs have been reported with 14% SPDE at 200 K for 1310 nm operation, though even at this temperature they suffer from a high DCR of around 10^8 Hz [43]. In addition, the reported SPDE may be exagerated due to an average of 1 photon per pulse being used. Due to the Poissonian nature of the pulse photon distribution, some pulses would have contained more than 1 photon. 1550 nm detection has also been demonstrated using Ge/Si SPADs. They are these diodes suffered from very poor SPDE of around (0.15% at 125 K and 6 % overbias) [44]. This makes them unsuitable to applications such as LIDAR, which is heavily dependent on photon detection at wavelengths in the 1000 - 1550+ nm range due to the free-space nature. Longer wavelengths allow for higher laser powers to be used since the eye-safe power limit increases.

III-V SPADs are ideal for this application, since the absorber can be optimised for these wavelengths. Commercial InGaAsP/InGaAs based SPADs are available, with an SPDE of 20 % and low dark count rates of 1 kHz when operated at 218 K, although these suffer from a lack consistency in material quality, and afterpulsing effects at higher frequencies due to carrier traps in the crystal lattice [45]. These APDs utilised a planar structure, back illumination, and an anti-reflection (AR) coating in order to achieve this. InGaAs/InP based SPADs have been used for 3D LIDAR imaging at distances of up to 10 km using time-correlate single-photon counting, with eye-safe laser power at 1550 nm by Pawlikowska *et al.* [46]. The detector that they used was designed by Tosi *et al.*, and was operated at 225 K in a self-contained TO-8 package with a three-stage thermo-electric cooler, based on a commercial Princeton Lightwave SPAD (no longer available). The highest SPDE demonstrated in an InGaAs/InP based SPADs was 55% for 1550 nm operation, with 500 MHz count rate and 10% afterpulsing probability at room temperature [47]. A very narrow (360 ps) gate width was used to achieve this, allowing for high overbiases (20 %) to be reached [47].

InGaAs/InAlAs SPADs are a promising material system due to InAlAs having superior impact ionisation properties than InP. Recent InGaAs/InAlAs SPADs have been fabricated with SPDE of 26% for 1550 nm operation at 210 K, and 100 MHz dark count rate using a 1.2 ns gate [48]. The SPDE was limited by a higher than expected dark current arising from the InGaAs absorber [48]. As the InGaAs/InAlAs material system matures, this is likely to improve.

1.5.5 Comparison

Table 1.1: Comparison of single-photon detectors.

Detector	Material	Temp. (K)	λ (nm)	SPDE (%)	Rate (MHz)	DCR (Hz)	Ref.
SNSPD	NbN	4-10	1550	5	50	-	[36]
	NbN	2-4	1550	67	10	-	[37]
	WSi	2	1550	90	10	10	[38]
TES	Nb/Ti	0.3	844	98	0.05	0	[39]
SPAD	Si (RE)	295	800	40	10	25	[42]
	Si (Thick)	295	650	3	3	10^3	[42]
	Ge/Si	200	1310	14	1	10^9	[43]
	Ge/Si	125	1550	0.15	0.1	10^7	[44]
	InGaAs/InGaAsP	218	1550	20	0.1	10^7	[45]
	InGaAs/InP	295	1550	55	1000	10^8	[47]
	InGaAs/InAlAs	210	1550	26	5	10^8	[48]

Table 1.1 compares the various single photon detectors. Comparisons are made as closely as possible, with normalised dark count rate used for gated SPADs. For high rate 1550 nm single photon detection, SNSPDs are clearly the best performing devices. This performance comes at a cost, however. All the SNSPDs require cooling to below 4 K for effective operation, which is not practical in many applications [36–38]. This fact is more exagerated for TES detectors, which require cooling to below 0.5 K [39].

A SPAD is the only realistic detector for mobile low photon counting for applications such as LIDAR in self-driving cars, due to the expensive and bulky liquid helium cooling requirement for SNSPD operation. For SWIR detection, especially 1550 nm, III-V SPADs are the only effective SPAD material. InGaAs/InP is currently the best material system demonstrated for this application [47].

1.6 High Speed Optical Receivers

In recent years there has been a rapid growth in telecommunication technologies and methods in order to process and transmit the maximum amount of data possible in the shortest time possible. A long-haul optical communication network is composed of three main components, the electro-optic transmitter (commonly a laser-diode), the optical fibre to carry the signal, and the receiver. A crucial component of such receivers is almost universally transimpedance amplifiers (TIAs), with the actual photodetection component being variable.

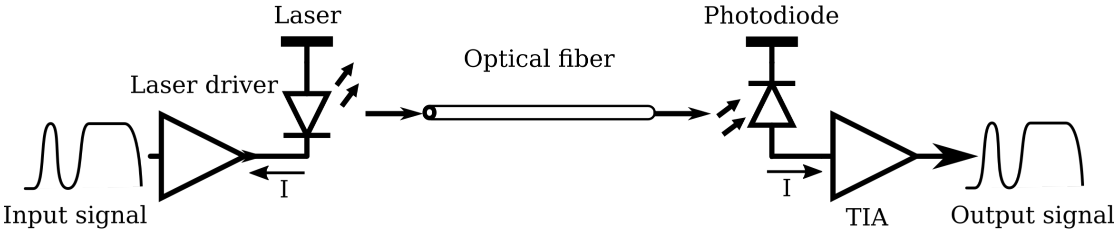


FIGURE 1.6: A basic optical communication network consisting of a laser, optical fibre, photodiode, and TIA.

A basic optical network is shown in figure 1.6. The laser driver provides a large current to the laser, amplifying the input signal and laser output which is coupled into the optical fiber. All detection systems require a component that is able to transfer the detection events from the detecting component efficiently. This is known as the front-end preamplifier, which is commonly a TIA in high-speed optical applications [49]. After exiting the optical fiber, the optical signal is coupled into the photodiode, the output of which is then amplified by the TIA and converted into a voltage. This allows for further processing of the signal by subsequent stages of the network.

1.6.1 PIN Photodiodes

The most basic photodetector consists of a *pin* photodiode operated at unity gain. Unity gain operation allows for photodetection with minimum noise, since there is no excess noise from gain mechanisms within the diode. A high-speed *pin* photodiode aims to maximise bandwidth by minimising the time it takes for a carrier generated in the depletion region to exit the region, while also minimising the time it takes for the detector to recover. The recovery time is dependent on the product of the parisitic junction capacitance and the load resistor, more generally known as the *RC* time constant. At the same time, high quantum efficiency (QE) requires wide depletion region. This leads to competing factors, the bandwidth reducing with increasing width, which must be balanced based on the target data transmission rate.

Additionally, with a higher optical power, the photodiode's bandwidth degrades due to carrier space charge effects arising from high current density in the junction exceeding the saturation current density. This is offset with an increased depletion width, again at the cost of an increased *RC* time constant. The balance between bandwidth and QE is a key limitation of *pin* photodiodes, and determines the sensitivity (optical power required for detection) of the detector.

Early state-of-the-art InGaAs *pin* photodiode - TIA based optical receivers, from the late 1980s, demonstrated bandwidths ranging from 10-16 GHz with sensitivities of around -19 dBm for 10 Gbps transmission at 1530 nm [50, 51]. Resonant cavity structures were utilised to increase the QE of InGaAs based photodiodes to 50 % while maintaining a bandwidth of 31 GHz [52].

1.6.2 Uni-Travelling-Carrier Photodiodes

Attempts to overcome space-charge limitations were made by groups utilising uni-travelling-carrier photodiodes (UTCPDs). As the name implies, only a single carrier type, the electron, is used in a UTCPD. This is achieved through the use of a wide-bandgap doped layer grown between the anode and a narrow bandgap absorption region, which blocks electron diffusion and drift [53]. The electrons are swept into a depleted carrier collecting layer, with holes swept out as conduction current. This reduces space-charge effects, with saturation current densities as high as 400 kA/cm^2 demonstrated [53]. Peak electrons velocity in materials used for UTCPDs are higher than that of holes, so the UTCPD has the added benefit of an increased bandwidth since only the faster electrons are active carriers [53]. When both holes and electrons are active, the slower of the two carriers acts to limit bandwidth.

High-bandwidth InGaAs UTCPDs have shown bandwidths as high as 310 GHz for 1550 nm detection. However, these devices do have very narrow absorption regions, leading to a low responsivity of 0.07 A/W [54]. InGaAsP UTCPDs with 1.04 A/W responsivity and 60 GHz have been demonstrated, achieved through the use of an evanescently coupled optical waveguide (to be discussed in section 1.6.4)[55]. With their high saturation current, UTCPDs are generally utilised for high-power, high-bandwidth applications.

1.6.3 High-speed APDs

Due to their high internal gain, APDs offer a improved sensitivity compared to standalone *pin* photodetectors since a lower optical power is needed to achieve the same output current. This makes APDs the photodetector of choice for applications where the optical transmission costs are significant, such as long-haul communication [56]. The gain-bandwidth product is an important measure for high-speed APDs, taking into account their signal gain effects in tandem with their bandwidth, and provides an important measure for comparing them. High-speed APDs utilise the SAM structure, allowing for the optimisation of absorption and multiplication independently. An APD's bandwidth is generally limited by the excess noise within the multiplication region, which increases with gain. This is due to the impact ionisation process, where average duration increases with gain depending on the impact ionisation coefficient ratio.

InP/InGaAs based SAMAPDs formed the majority of early high-speed APDs, with sensitivities as low as -38.1 dBm achieved for 1 Gbps transmission at 1550 nm as early as 1983 [57]. More recently InP/InGaAs APDs utilising waveguides have been used for 40 Gbps with a gain bandwidth product of 114 GHz at a gain of 9, allowing for a sensitivity of -19.0 dBm. Over the past decade InAlAs has begun to replace InP as the multiplication material, with InGaAs/InAlAs based APDs now dominating the area in terms of performance due to a more favourable impact ionisation coefficient ratio and gain-bandwidth product, as demonstrated in figure 1.7.

InAlAs/InGaAs mesa SAMAPDs have demonstrated record, for APDs, sensitivities of -21.8 dBm for 25 Gbps detection at 1310 nm, obtained through the use of a double mesa structure to keep the electric field away from the mesa sidewalls, minimising dark current [59]. In addition, a InAlAs/InGaAs SAMAPDs with -10.8 dBm sensitivity for

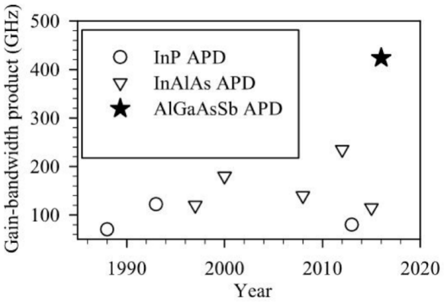


FIGURE 1.7: Comparison of InAlAs and InP based APDs and their gain-bandwidth products from [58].

50 Gbps operation was demonstrated using a similar structure, with a gain-bandwidth product of 141 GHz [60].

1.6.4 Optical Preamplifier-PIN Detectors

It is possible to amplify the optical signal prior to detection in the photodiode. With optical pre-amplification, a higher bandwidth photodiode can be used since the QE requirement is relaxed due to an increase in incident optical power. This is the basis of erbium-doped fiber amplifier (EDFA) and semiconductor-optical amplifier (SOA) based detectors. In an EDFA, the core of the fiber acts as an amplifier at 1550 nm when driven with a higher frequency, 1480 nm wavelength laser which stimulates the the dopant erbium atoms into a higher energy state. From this state the stimulated emission of photons by the signal photons causes signal gain at the signal wavelength [61]. A UTC/EDFA based receiver has shown record -28 dBm sensitivity at 40 Gbps, though the footprint is large and costs high which make it unsuitable for integration into long-haul networks [62, 63].

SOA-PIN detectors integrate the optical amplification component into the semiconductor itself, with a waveguide doubling as an optical amplifier. Utilising an InGaAs active layer for optical pre-amplification, SOA-PIN-TIA receivers have been demonstrated [64]. The SOA component exhibits high noise at frequencies below 40 GHz, which limits performance for 25 Gbps where SOA-PIN based receivers only match the performance of existing APD based receivers [65, 66]. These SOA-PIN based detectors however offer advantages over current APD based detectors for 40 Gbps operation, with sensitivities demonstrated as low as -23 dBm in a small integrated package [63]. SOA-PIN detectors are the main competitors to APDs for next generation optical communication networks.

1.7 Motivation

APDs are a mature technology which is perfect for application to single photon counting - essential for any quantum network. They have found heavy application in optical communications, especially over long distances thanks to their high gain and SNR increasing the distance required before a signal repeater is required [67]. III-V alloy APDs in particular are a good choice as they allow for fine-tuning of wavelength range by varying alloy composition. The main advantage of this is that they can be tuned to work with the near-IR wavelengths that silicon is unable to reach, mainly the optical window around 1500-1580 nm. This is vital for long distance fibre communication. They can be tailored to a particular specification which is important for working with the specialised technology required for quantum computing. Separate absorption and multiplication avalanche photodiodes (SAMAPD) take this concept to another level, reducing noise even further by separating the multiplication region from the absorption region.

InGaAs/InAlAs based SPADs are a promising technology which offers potential improvements over the current InGaAs/InP based SPADs due to better impact ionisation properties and a wider band-gap, and therefore reduced tunneling current [22]. InAlAs reaches higher breakdown probabilities at lower over bias ratios when compared to InP, potentially allowing for higher detection probabilies at reduced dark count rates [68].

Electron APDs (e-APDs) are based on materials where only electrons undergo impact ionisation. InAs in particular has shown promise as an effective e-APD material, with very low excess noise characteristics and high electron drift velocities, it is a potential candidate for high-speed optical receivers [69]. Modelling the theoretical potential of InAs based optical receivers is of great interest, since InAs has not been explored in this context before due to relatively high dark current [70]. Due to the unique impulse response shape of e-APDs, existing APD-based-receiver sensitivity models are not usable, so a new model has to be developed in order to explore the capabilties of InAs APDs in this area.

1.8 Thesis Organisation

The first half of this thesis is primarily focused on the design, fabrication, and characterisation of single photon APDs based on the InGaAs/InAlAs material system. The later chapters shift focus to modelling high-speed detector systems with a focus on developing a new theoretical model for determining the sensitivity of e-APDs, and exploring the theoretical posibility of using InAs APDs as the receiver component of a high-speed detector. An overview of each chapter is given below.

Chapter 2 provides background theory to aid in the understanding of core topics which include impact ionisation, breakdown probability, electric field modeling, dark current and dark count mechanisms, Geiger-mode APD circuit quenching, APD bandwidth, and bit error rates.

Chapter 3 introduces standard device fabrication procedures, and presents etching trials that were performed to optimise the fabrication process for InGaAs/InAlAs SPADs. Device characterisation techniques are explained. A double mesa fabrication trial is also presented.

Chapter 4 - two InGaAs/InAlAs based SAM APD wafers are designed and simulated, optimised for single-photon counting. Wafers grown based on the designs are fabricated, characterised, and then assessed for suitability to single-photon counting.

Chapter 5 shifts focus towards sensitivity optimisation of e-APD based high-speed optical receivers. A new sensitivity model is developed, allowing for the simulation of e-APD based optical receivers. This model is demonstrated using simulated InAs APDs, with the suitability of InAs based e-APDs to high-speed optical receiver applications assessed.

Chapter 6 uses the model developed in chapter 5 and a previously developed sensitivity model for standard APDs to compare the properties of e-APDs and APDs as part of high-speed optical receivers. Further simulations are performed on InAs, varying several parameters in order to explore how the sensitivity of InAs based receivers could be improved.

Chapter 7 concludes the thesis, and summarises the content. Future research work is suggested for improving the SPDE of InGaAs/InAlAs based SPADs, and designs suggested for InAs APDs optimised for high-speed detection.

Chapter 2

Background Theory

This section will discuss various theories used in the design and characterisation of APDs in this work.

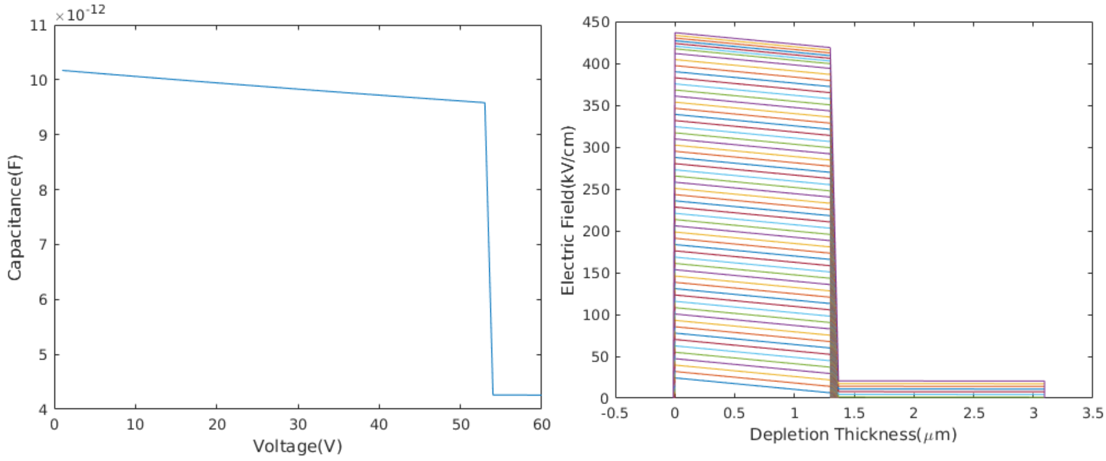
2.1 Capacitance Modelling and Punch-through

Punch-through is an event which occurs when a SAMAPD is reverse biased past a certain voltage. It corresponds to the point where the electric field in the depletion region "punches-through" into the absorption region. As voltage across a diode is increased, the electric field reaches further through the depletion region until it reaches the absorption region. If the device is being illumincated with light, there would be a clear rise in the photo-current at this voltage because the carriers originally generated through photon absorption in the absorption region can now drift into the multiplication region. This point corresponds to a sudden drop in capacitance across the diode, since depletion width has increased. An example of this is shown in figure 2.1, where at 54 V punch-through occurs.

Capacitance for a single region is described by the equation

$$C = \frac{\varepsilon_r \varepsilon_0 A}{w_d}, \tag{2.1}$$

where ε_r is relative perimittivity, A is diode area, and w_d is depletion width. The electric field profile within the depletion region for a given bias can be modelled using Poisson's



(a) (b)

FIGURE 2.1: Example of extracting electric field data from a CV graph for a SAM structure. The voltage is increased in 1 volt steps and the corresponding field profile plotted.

equation, which is given by

$$\frac{\partial E}{\partial x} = \frac{qN}{\varepsilon_r \varepsilon_0}, \tag{2.2}$$

where E is electric field, and N is the doping in the region.

An example electric field profile for a *pin* structure is given in figure 2.2. The gradients of the lines (eg. $\frac{qN_q}{\varepsilon_r\varepsilon_0}$) are given in the figure. For example, for the electric field gradient across the p^+ region $(dE/dx)_{p^+}$ is described by $qN_1/\varepsilon_r\varepsilon_0$. The area under the electric field profile is determined by the reverse voltage. As voltage is increased, the depletion widths x_1, x_2, and x_3 increase dependent on the doping concentration in each region, with the maximum width of x_2 defined by the i region width. With the depletion width of each stage of the electric field known, as well as the values of electric field, the structure can be fitted to by adjusting N_1, N_2, and N_3.

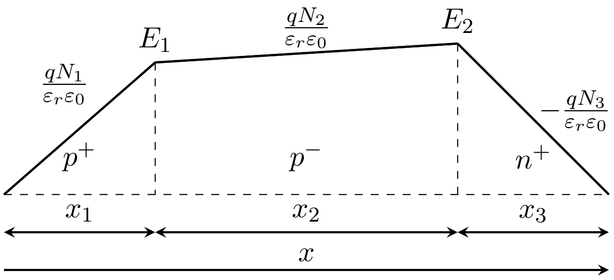


FIGURE 2.2: An annotated electric field distribution for a depleted *pin* structure biased to an arbitrary voltage, with dopings N_1, N_2, and N_3 for the p^+, p^-, and n^+ regions. E_1 and E_2 are the electric field values at x_1 and x_2, respectively.

This can be applied to a structure with m regions by using the generalised expressions for depletion width of the first region x_1, and m^{th} region x_m, which are given by

$$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \tag{2.3}$$

and

$$x_m = \frac{\sum_{k=1}^{m-1} N_k x_k}{N_m}, \tag{2.4}$$

where

$$a = \frac{N_1}{\varepsilon_1}\left(1 - \frac{\varepsilon_m N_1}{\varepsilon_1 N_m}\right), \tag{2.5}$$

$$b = \frac{2N_1}{\varepsilon_1}\left(x_k - \frac{\varepsilon_m N_k x_k}{\varepsilon_k N_m}\right) \tag{2.6}$$

and

$$c = \sum_{k=2}^{m-1}\left(\frac{N_k x_k^2}{\varepsilon_k}\right) + \sum_{k=2}^{m-2}\left(\frac{2N_k x_k}{\varepsilon_k}\left(\sum_{l=k+1}^{m-1} x_l\right)\right) - \frac{\varepsilon_m \left(\sum_{k=2}^{m-1}\left(\frac{N_k x_k}{\varepsilon_k}\right)\right)^2}{N_m} - \frac{2V}{q}, \tag{2.7}$$

where N_k and ϵ_k are the doping concentration and permittivity, respectively, within the k^{th} region, and V is the reverse bias voltage. The total depletion width w_d is given by

$$w_d = x_1 + \sum_{k=2}^{m-1} x_k + x_m. \tag{2.8}$$

Capacitance can then be calculated using

$$C = \frac{A}{\sum_{k=1}^{m} \frac{x_k}{\varepsilon_k}}, \tag{2.9}$$

which can be fitted to experimental measurements.

2.2 Impact Ionisation and Breakdown Probability

Impact ionisation is the mechanism by which avalanche breakdown occurs. When an APD is under high reverse electric field, an electron promoted into the conduction band will gain energy from the electric field. With enough energy, the electron may ionise another bound electron, promoting the second electron to the conduction band. When the effect becomes self sustaining, the APD has entered avalanche breakdown.

Several variables need to be considered in the calculation of breakdown probability of an APD. The initial electron that is triggering the breakdown must first cause an ionisation event. At least one of the resulting carriers needs to cause one ionisation. If more carriers are being created than those exiting the APD, there will be an increase in current. There is also a finite probability that carrier generation will not cease - this is the avalanche breakdown probability. For electrons this can be described by [6]

$$P_e(x+\Delta x) = P_e(x) + \underbrace{\alpha\Delta x\left[P_e(x)+P_h(x)-P_e(x)P_h(x)\right]}_{\substack{\text{Probability that in transit from x to } \Delta x \\ \text{ionization occurs which results in infinite descendents}}} - \overbrace{P_e(x)\alpha\Delta x\left[P_e(x)+P_h(x)-P_e(x)P_h(x)\right]}^{\text{Negative product of the first two terms}},$$

which can be represented in differential form as

$$\begin{aligned}\frac{dP_e}{dx} &= (1-P_e)\alpha\left[P_e+P_h-P_eP_h\right] \\ \frac{dP_h}{dx} &= -(1-P_h)\beta\underbrace{\left[P_e+P_h-P_eP_h\right]}_{P_{\text{pair}}}.\end{aligned} \tag{2.10}$$

α and β are the ionisation coefficients and P_{pair} is the probability that an electron-hole pair generated at x will initiate an avalanche, which combines P_e and P_h together - the probability of an electron, and the probability of a hole generated at x initiating an avalanche, respectively. These coupled differential equations for $\frac{dP_e}{dx}$ and $\frac{dP_h}{dx}$ can be solved using numerical integrations with boundary conditions $P_e(0) = 0$ and $P_h(w) = 0$ [6].

While equation (2.10) is valid, expressions for P_e and P_h that can be solved analytically are desirable. This was achieved in [71], which used the approximation $\alpha = k\beta$ so that $P_e(w)$ and $P_h(0)$ are expressed in terms of k and the term $\int_0^w(\alpha-\beta)dx$. This assumes k is close to unity and not $k \ll 1$ [71]. Re-writing P_{pair} as $P_{\text{pair}} = 1-(1-P_e)(1-P_h)$, differentiating, and integrating leads to:

$$\begin{aligned}\frac{dP_{\text{pair}}}{dx} &= (\alpha-\beta)P_{\text{pair}}(1-P_{\text{pair}}) \\ \frac{P_{\text{pair}}(x)}{1-P_{\text{pair}(x)}} &= \frac{P_{\text{pair}}(0)}{1-P_{\text{pair}}(0)}\underbrace{\exp\left[\int_0^x(\alpha-\beta)dx\right]}_{f(x)}.\end{aligned} \tag{2.11}$$

Using the boundary conditions from equation 2.10 it can be seen that $P_{\text{pair}}(w) = P_e(w)$ and $P_{\text{pair}}(0) = P_h(0)$, applying to (2.11) for $x = w$ gives the ratio between P_e and P_h:

$$\frac{P_e(w)}{1 - P_e(w)} = \frac{P_h(0)}{1 - P_h(0)} f(w) \tag{2.12}$$

This, on its own, is not useful. Solving equation (2.11) for P_{pair} gives

$$P_{\text{pair}}(x) = \frac{P_h(0) f(x)}{P_h(0) f(x) + 1 - P_h(0)}. \tag{2.13}$$

$\frac{dP_h}{dx}$ from (2.10) can be re-written in integral form:

$$1 - P_h(x) = [1 - P_h(0)] \exp\left[\int_0^x \beta(x') P_{\text{pair}}(x') dx'\right]. \tag{2.14}$$

Using $P_h(w) = 0$, $P_h(0)$ is given by

$$-\ln[1 - P_h(0)] = \int_0^w \beta(x) P_{\text{pair}}(x) dx \tag{2.15}$$

with P_{pair} defined in equation (2.13). This provides a single equation to determine breakdown probability, though at this point it still needs to be solved numerically. However, if $\beta = k\alpha$ is valid, equation (2.15) can be solved analytically since $df(x)/dx = (\alpha - \beta) f(x)$ can be integrated exactly:

$$-\ln[1 - P_h(0)] = \frac{k}{1-k} ln\left[P_h(0) \exp[(1-k)\delta] + 1 - P_h(0)\right], \tag{2.16}$$

where $\delta = \int_0^w \alpha dx$. Equation 2.12 can then be used to relate P_e and P_h to give:

$$1 - P_h(0) = [1 - P_e(w)]^k. \tag{2.17}$$

Using equations (2.16) and (2.17), $P_h(0)$ and $P_e(w)$ can be calculated as a function of k and δ, without needing to know the electric field distribution precisely. These are implemented and shown in figure 2.3, which plots breakdown probability versus δ. Since the ionisation coefficients, and hence δ, depend on electric field values, and consequently on voltage, effectively the graph can be viewed as being a function of voltage. A sharper gradient for both P_e and P_h means that they will reach their maximum values faster, allowing lower operating voltages.

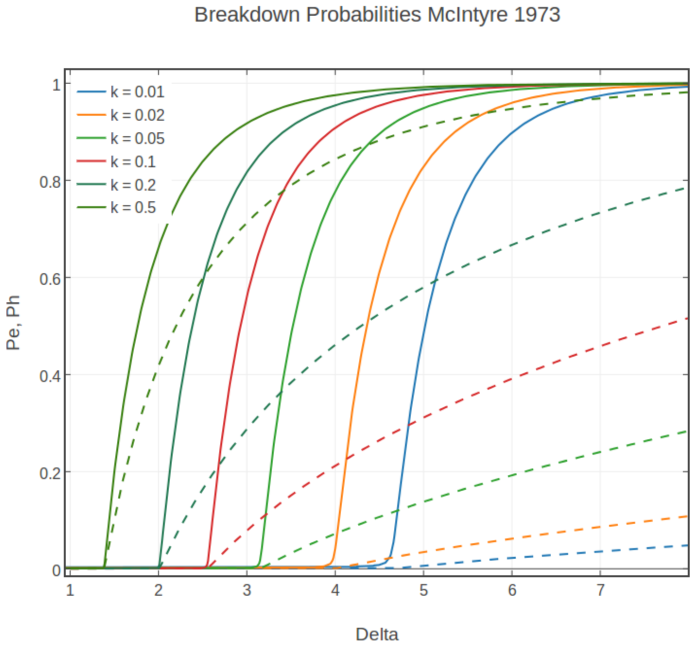
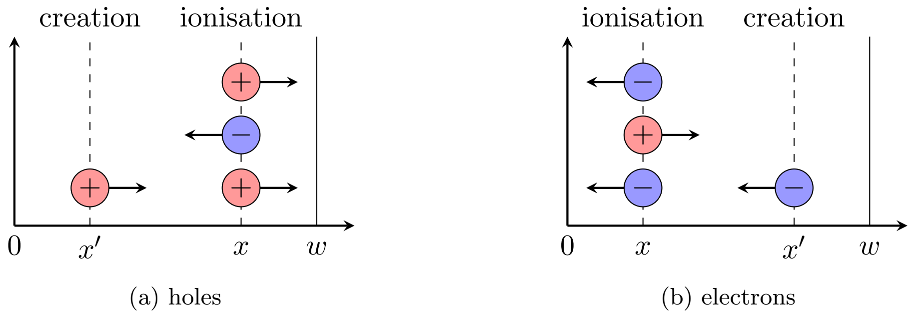


FIGURE 2.3: Plot of the implementation of McIntyre's 1973 breakdown probability equations. Solid lines represent P_e values, while dashed lines represent P_h values. These correspond to their ionisation coefficient ratio $k = \beta/\alpha$. $\delta = \int_0^w \alpha dx$.



(a) holes

(b) electrons

FIGURE 2.4: Visualisation of hole and electron creation. Carriers are created at x' and ionise at x after travelling a distance Δx.

While the previous equations are useful for modelling devices with a wide multiplication region, they are not able to accurately model what occurs in a device with a thin multiplication length. Non-local effects become significant in such regions, meaning that the carrier's history independence is no longer valid. A new variable must be considered; the dead space $d_{e,h}$.

$$d_e \approx \frac{1}{\alpha} \text{ and } d_h \approx \frac{1}{\beta} \tag{2.18}$$

d is defined as the length at which the ionisation coefficients reach 50% of their "constant" values. Over short distances the ionisation coefficients previously discussed are no longer valid and must be replaced with history dependent ionisation coefficients, defined as $\alpha(x'|x)$ for electrons and $\beta(x'|x)$ for holes as shown in figure 2.4. A charge carrier created x' is expected to ionise at x. The corresponding history-dependent ionising probability are given by

$$\begin{aligned} p_e(x'|x)dx &= \alpha(x'|x)\exp\left(-\int_x^{x'} \alpha(x'|x'')dx''\right)dx \\ &= \alpha(x'|x)P_{se}(x'|x)dx \end{aligned} \tag{2.19}$$

$$\begin{aligned} p_e(x'|x)dx &= \beta(x'|x)\exp\left(-\int_{x'}^{x} \beta(x'|x'')dx''\right)dx \\ &= \beta(x'|x)P_{sh}(x'|x)dx \end{aligned} \tag{2.20}$$

for electrons and holes respectively. $p_{e,h}$ are given by the ionisation coefficients multiplied by the newly defined term, survival probability $P_{s(e,h)}$. This is the probability that a carrier travelling from x' to x will survive without impact ionising. These are visualised in figure 2.5.

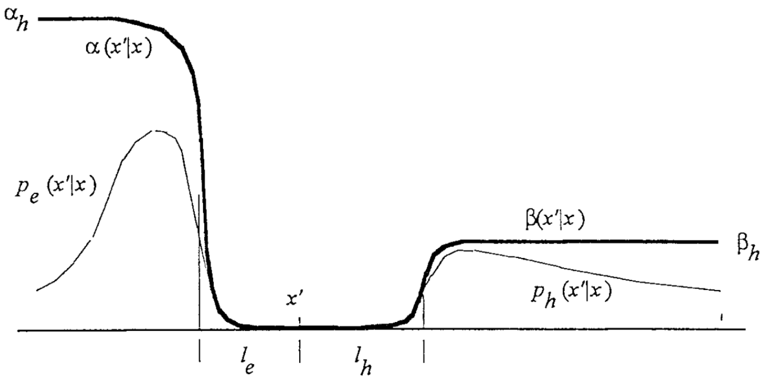


FIGURE 2.5: Spatial representation of the ionisation coefficients and their corresponding ionisation probabilities, moving away from carrier creation point x'. The probability of ionisation increases with distance travelled before peaking at approximately where α and β reach their constant values α_h and β_h [72].

Using the history-dependent ionisation probabilities, the history-dependent breakdown probabilities are given by

$$(1 - P_{\mathrm{e}}(x')) = P_{\mathrm{se}}(x'|0) + \int_0^{x'} p_e(x'|x)(1 - P_{\mathrm{e}}(x))^2(1 - P_{\mathrm{h}}(x))dx, \tag{2.21}$$

and

$$(1 - P_{\mathrm{h}}(x')) = P_{\mathrm{sh}}(x'|w) + \int_{x'}^{w} p_h(x'|x)(1 - P_{\mathrm{h}}(x))^2(1 - P_{\mathrm{e}}(x))dx. \tag{2.22}$$

These equations can be used to simulate any avalanche photodiode, regardless of depletion width and electric field.

2.3 Tunneling Current

Tunnelling occurs when an electron quantum tunnels through the barrier between valence and conduction bands. As reverse voltage is increased, the band structure is bent further with the separation barrier between the two bands reducing in thickness. This is shown in figure 2.6. This increases the probability that a carrier will tunnel through.

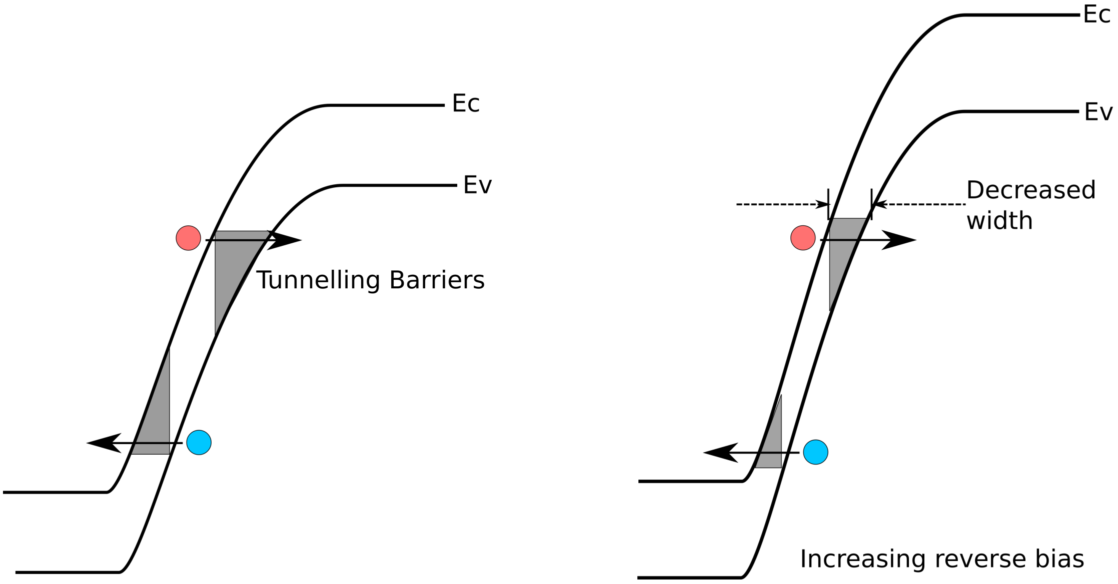


FIGURE 2.6: Band diagrams showing increasing voltage and the effects on the tunnelling barriers. The barrier width decreases with increased voltage.

The tunneling current for a direct bangap can be described using

$$I_{\text{tunn}} = \frac{(2m^*)^{0.5} q^3 EVA}{h^2 E_g^{0.5}} \exp\left(-\frac{2\pi\sigma_T (m^*)^{0.5} E_g^{1.5}}{qhE} \right), \tag{2.23}$$

where m^* is effective electron mass, E is electric field, V is voltage, A is diode area, E_g is bandgap, q is elementery charge, h is Planck's constant, and σ_T is the tunneling constant. The tunneling constant is a fitting parameter that encapsulates the barrier shape, and varies between different materials and doping distributions [73].

2.4 Dark Counts

In an APD which is being operated above breakdown voltage in Geiger mode, a dark count is an event which has been caused by noise rather than an incident photon. Dark counts have three main causes; thermal noise, afterpulsing, and tunnelling.

Thermal noise results from electrons being thermally excited from the valence band into the conduction band. This effect can be suppressed by cooling the APD to a lower temperature, however this can negatively impact the detection efficiency and maximum detection frequency via afterpulsing.

Afterpulsing is caused by carriers which get trapped within the trap centres of the device. These are defects which occur throughout the device and result from impurities during growth. After an avalanche event is detected, the circuit must be quenched to prevent damage to the device. After the circuit is quenched, these trapped carriers remain in the centres and limit how quickly the APD can be readied for the next detection event, since time must be given for them to disperse. This dispersal time will increase with a decreased temperature, so this afterpulsing effect has to be considered especially when operating at low temperatures.

2.5 Circuit Quenching

When the APD is operated in Geiger mode, it must have a circuit quenching mechanism to prevent a runaway avalanche effect from destroying the device. This quenching is achieved by using an external circuit which connects to the APD. There are three forms of quenching; passive, gated, and active.

Passive quenching uses a ballast resistor to dissipate the voltage across the APD after a breakdown event, reducing it back down to breakdown voltage. When the residual current from the avalanche event dissipates, the voltage rises back to the externally set DC bias. While simple, this method has the drawback of restricting the measurement repetition rate to around 100 ns (dependent on the size of the ballast resistor and the capacitance of the system) , which is inadequate for many applications [74].

Gated quenching uses short pulses, generally 10 ns or less, to bias the APD above breakdown voltage. These pulses are of such duration that the device is brought below breakdown voltage before damage can occur. This method is simple, however it relies on precise knowledge of when the photons are expected to arrive.

Active quenching is significantly more complex, with two main methods available. They both rely on having active logic to monitor the avalanche current, and act when this crosses a certain threshold. The first superimposes a bias of opposing polarity over the reverse bias, bringing the overall voltage below breakdown. The second uses passive quenching via a ballast resistor to reduce the current, which raises the voltage between the resistor and the diode. This is monitored by the logic circuit which then brings the voltage to the quenching voltage. The logic circuit often contains circuitry to actively recharge the APD, significantly reducing the dead time when compared to simple passive quenching.

2.6 3 dB Bandwidth and the Gain-Bandwidth Product

The 3 dB bandwidth of an APD is defined in terms of the fourier transform of the time domain mean impulse response function of the APD. The frequency at which the resulting frequency domain signal reaches 3 dB less than the maximum is defined as the 3 dB bandwidth. In addition to just the bandwidth, the inherent signal gain properties of an APD also provide an additional benefit to the bit-rate that the diode can realistically detect. This means that for an APD, the gain-bandwidth product is a more important measure than just the bandwidth itself, since the gain properties amplify the signal above the noise floor that would normally arise from just the bandwidth.

Generally for a regular APD the gain-bandwidth product is a mostly unchanging quantity. This is due to the gain-bandwidth product limitation that arrises due the inherent random, stochastic nature of the impact ionisation process within an APD. The theory for this limitation was extensively explored by Emmons [75].

For a given APD ($k \neq 0$) undergoing pure carrier injection, both the holes and the electrons are undergoing impact ionisation. As the gain increases within the diode (with increased reverse voltage), the overall probability that an avalanche event will be longer increases, since as gain increases the probability of a given carrier causing an impact ionisation event increases. In an avalanche medium with both holes and electrons having an equal probability to impact ionise, this provides the worst case in terms of effects on bandwidth ($k = 1$). The best case scenario is where only a single type of carrier undergoes impact ionisation ($k = 0$). Effects of varying k on bandwidth are plotted in figure 2.7. A diode exhibiting $k = 0$ has no negative impact from increased gain, thus overcoming the gain-bandwidth product limitation. In such a diode increased gain only provides benefits in terms of high-speed operation. An example of such APDs are e-APDS, where only the electrons undergo impact ionisation.

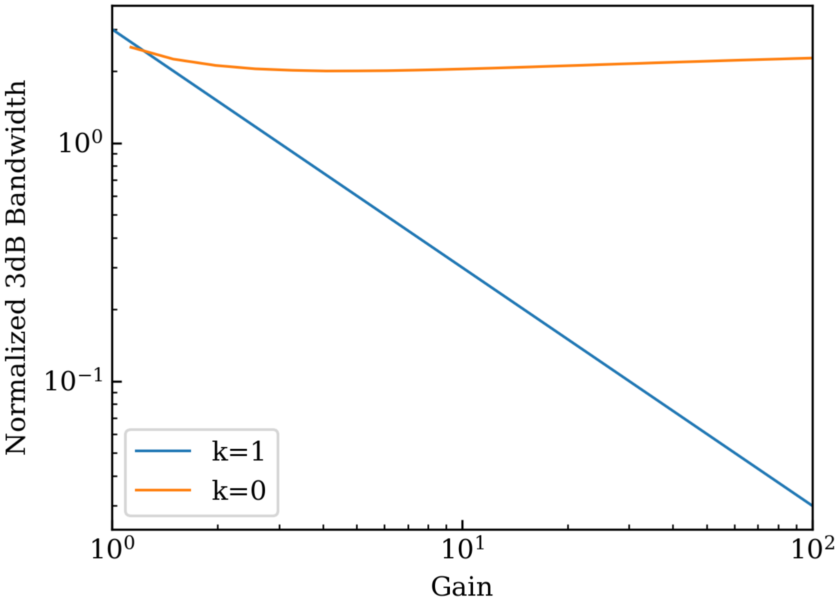


Figure 2.7: Normalised 3dB bandwidth as it changes with gain. Bandwidth reduces with gain for APDs where $k \neq 0$, while stabilising for APDs where $k = 0$. The initial drop for $k = 0$ corresponds to an increase in excess noise factor, though this never increases above 2. Data is generated from models presented in later chapters.

2.7 Bit-error Rate

In optical communcations, the bit-error rate (BER) is the probability that for a given detection threshold set between a “0” and a “1” bit, an incorrect measurement will be made. For a “0” or a “1” bit the measured signal generally has a gaussian distribution. A simplified schematic is demonstrated in figure 2.8. The mean value of the signal for the “0” bit is defined by the noise floor set by the combined noise level of the detection componenets in the optical receiver. The signal level of the “1” bit is defined by a combination of the optical signal strength and gain effects within the optical receiver. A weaker optical signal will result in a greater overlap between the distribution of the "0" and "1" bits, leading to a higher BER.

FIGURE 2.8: An example of the overlap (highlighted in red) between a "0" and a "1" bit, centered at μ_0 and μ_1 with standard deviation σ_0 and σ_1, respectively. The decision threshold θ is selected as the level at which BER is minimised.

The detector speed also has an impact on the probability of an error. If a detector is too slow for a given bit-rate, then "1" bits from previous detection windows can influence the next detection windows. This can lead to an error in the event detection window following a "1" if the proceeding bit is expected to be a "0". This is visualised in figure 2.9. Such errors limit the bit-rate that a given detector can operate at, since the detector will no longer be able to reset fast enough to detect reliably.

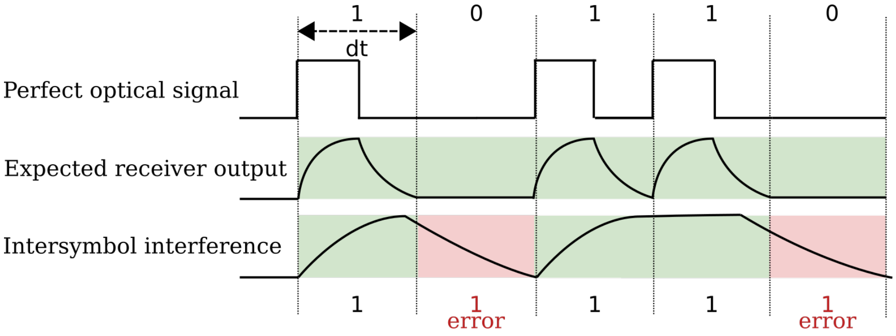


FIGURE 2.9: Bit-errors caused by a detector being too slow to reset after a detection event. The signal from "1" bits from previous detection windows extends into the next.

Chapter 3

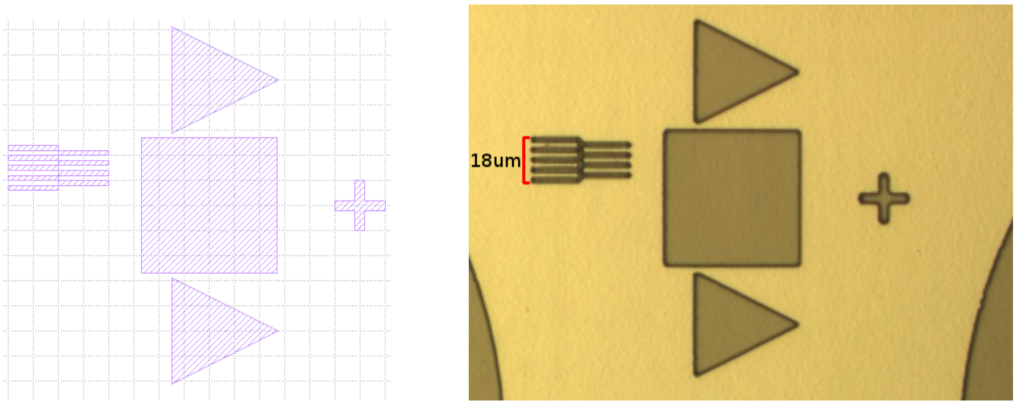
Experimental Details

This chapter will describe various experimental work that has been conducted as part of the thesis. It includes wafer structure design, device fabrication, and characterisation of SPADs.

3.1 Standard Mesa Fabrication Procedure

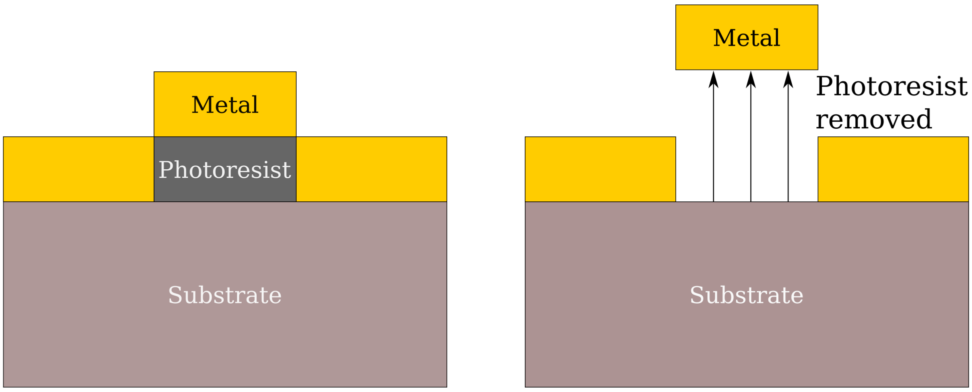
A device fabrication procedure for a basic SAMAPD will be described in this section. The first step of a fabrication run, after cleaving the sample from the wafer, is 3-stage cleaning. The sample is immersed in warmed n-butyl acetate, acetone, and isopropanol, in the given order. This removes any dirt or grease from the sample. Photo-lithography generally follows cleaning. This involves using a photo-resist spinner to evenly coat the sample in a photoresist - a photosensitive material which reacts strongly to exposure with certain wavelengths of light. The photoresist is exposed to UV light from a mask aligner loaded with a mask plate. This plate is engraved with a pattern that the light from the aligner shines through onto the sample, transferring this pattern to the photoresist on the surface. The sample is then submerged in a developer for 1 minute, which removes areas of photoresist that have been exposed. Figure 3.1 shows an example of a pattern that has been transferred to a substrate.

The next step is to metallise over the pattern to form the top contacts of the device. This is done by depositing metal on the sample, either in a metal evaporation chamber



(a) Pattern design file (b) Transferred pattern

FIGURE 3.1: Example of a pattern (a) transferred to a sample coated with photoresist. The part shown in the optical microscope image (b) is used to determine if the exposure quality is good. The "fingers" (highlighted) are well defined, indicating the correct exposure and development time.



(a) Metal deposited over substrate and photoresist.

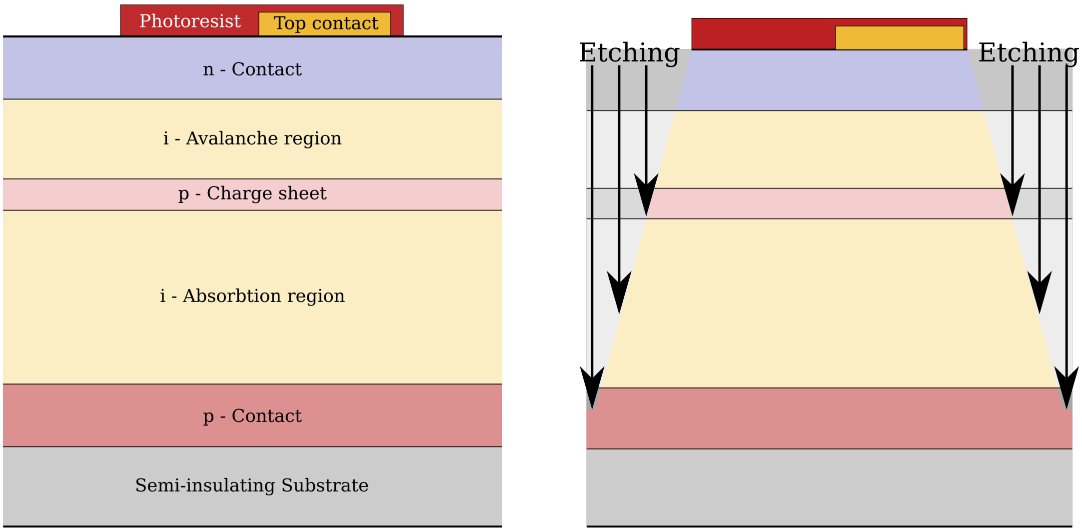
(b) Removal of photoresist with acetone lifts off metal.

FIGURE 3.2: Lift off procedure. (a) is how the sample appears after removal from the metallisation chamber. Placing it in acetone will lift off the metal, leaving behind a pattern as shown in (b).

or a metal sputterer. Following metal deposition, the sample is placed in acetone to lift off metal in areas with photoresist. The procedure is demonstrated in figure 3.2.

Following lift off, the sample is then re-patterned with a mesa etch mask. Trenches are etched around the top contacts using either wet (acid) etching or dry (plasma) etching. This effectively creates individually defined devices on the bulk sample. The aim is to etch down through layers of semiconductor to reach the lower charge layer of the sample, as well as to isolate devices.

The sample is then patterned with a lower contact mask layer, and the metallisation-liftoff procedure is repeated. Depending on the requirements for the diodes, this can



(a) Pre-etch (b) Post-etch

FIGURE 3.3: Wet etching. The profile of a mesa after wet etching depends heavily on the etchants used. In general, the angle is not perpendicular to the substrate.

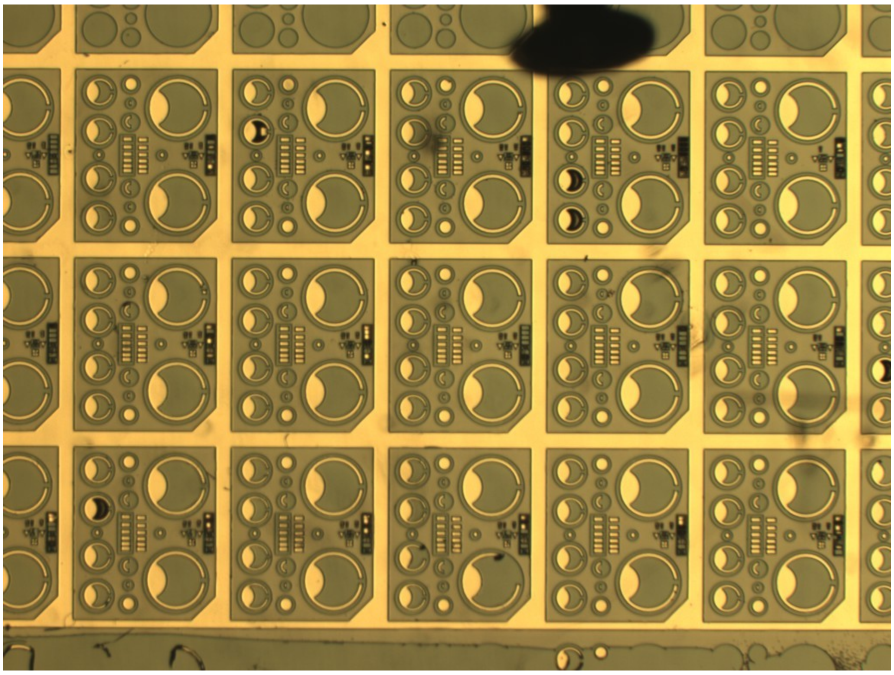
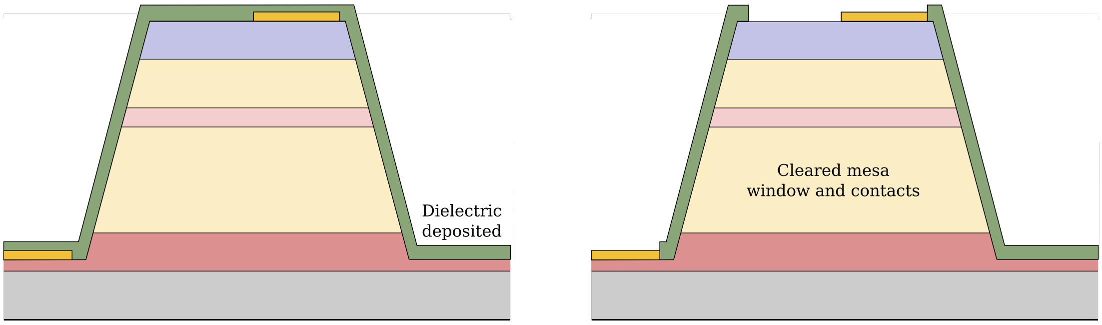


FIGURE 3.4: A sample after undergoing the lower contact metallisation-liftoff procedure. Individual mesa devices with top contacts are visible, with each cell of devices surrounded by a metal grid contact (the lower contact).

often be the final step. The resulting devices and lower contact from such a fabrication run can be seen in figure 3.4. These diodes, ranging from 60 μm to 420 μm diameters, are now ready for electrical characterisation, as well as packaging.

Fabrication of smaller diodes generally requires more steps than this, however. A contact area must be approximately 100×100 μm for devices to be reliably bonded to, so if a diode's diameter doesn't allow for such a contact then bond-pads must be used. In order to not short-circuit the contact layers with the bond-pad tracks running down the sidewall, it must first be passivated with a dielectric nitride layer. The thickness of this layer is determined by the deposition temperature (and therefore breakdown voltage) of the nitride and the required operating voltage, since it must be thick enough to prevent breakdown between the bondpad track and the mesa sidewall.



(a) (b)

FIGURE 3.5: Deposition and removal of dielectric from contact areas and mesa window.

With the nitride deposited, it must be cleared away from contact areas and the mesa window. To achieve this a nitride etch mask is patterned onto it. This is subsequently eteched using reactive ion etching (RIE). The bondpad layer is then deposited over the sample, linking to the mesa top contact.

3.2 Photolithography

Several photoresists and chemicals are used for the photolithography stage of fabrication. This section explores the differences and reasoning for their use. For consistency all spin coating is performed at 4000 RPM for 30 s, so where thickness is mentioned it is based on these parameters.

3.2.1 BPRS200

The BPRS200 positive photoresist was widely used within the group for various stages of fabrication, including metal contact deposition, etching, and nitride removal. With a spin thickness of ∼2-2.5 μm, it is thick enough to be used on its own for metallisation while not causing issues with definition of the finer metal patterns on the surface. Additionally, the viscosity of the photoresist provides a highly even coating when spin coated, which is important for etching where accurate surface profiling of the photoresist is important. The developer used is diluted MF26A:H_2O (0.7:1). Unfortunately, due to issues with securing supply of the photoresist, it is being phased out, so alternative photoresists must be used.

3.2.2 SPR350 and Polymethylglutarimide (PMGI)

A suitable replacement for BPRS200 for etching purposes is the SPR350 photoresist. Its low viscosity means that it spins evenly to a thickness of $\sim 1.3\ \mu m$, with minimal edge bead buildup at the sample perimeter. The small thickness and lack of edge beads is advantageous during mask alignment, enabling minimal seperation between the mask plate and the semiconductor surface (crucial for defining highly detailed pattern features). The dehydration bake time is 1 minute, and the developer solution is MF26A:H_2O (0.7:1) with post exposure development taking 1 minute.

At $\sim$1.3 μm thick, using SPR350 is unsuitable for metallisation, because metal lift-off can be difficult (especially when metal has been deposited using a sputterer with a rotating platform). There is insufficient surface area for acetone to attack and remove the resist, leading to a need of extensive use of an ultrasonic sonicator. This is undesirable since using a sonicator for extended periods of time damages the sample and can lift off metal in areas where it is required.

An additional resist called polymethylglutarimide (PMGI) is required to be used in a bi-layer stack for metallisation using SPR350 to be effective. The PMGI+SPR350 procedure is as follows:

1. Spin PMGI onto the sample
2. Bake the sample for 5 minutes at 180°C for 5 minutes
3. Spin SPR350 onto the sample
4. Bake the sample for 1 minute at 100°C
5. Expose the desired pattern
6. Develop in MF26A:H_2O (0.7:1) for 1 minute

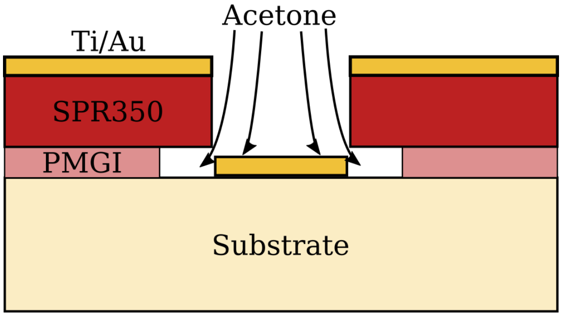


Figure 3.6: PMGI-assisted lift-off.

After developing the resists, the SPR350 will be removed in the exposed areas as usual, while the PMGI is attacked by the developer a greater amount. The resulting profile is shown in figure 3.6, with the SPR350 overhanging the PMGI. Following metal deposition, the PMGI provides a channel for the acetone to more readily remove the SPR350, greatly easing the lift-off process and eliminating the need for sonicator action. Following removal of the SPR350 in acetone, the sample must be submerged in MF26A developer to remove the PMGI.

3.2.3 SPR220

SPR220 is a relative viscous photoresist, spin-coating to a thickness of approximately 5 μm. It is used in applicactions where SPR350 is found to be insufficient, mainly in processes where the photoresist is actively etched along with the sample such as RIE/ ICP etching, or where SPR350 does not effectively coat the surface features of a sample. Such a situation usually arises when there are many fine, deep features in excess of 4 μm etched into the surface. The dehydration bake time for SPR220 is 1 minute 45 seconds, and the post-exposure development time is 1 minute 10 seconds.

The high viscosity and thickness of SPR220 causes issues during spin coating and exposure. Large edge-beads form, preventing close contact with mask plate, and hence causing poor definition of exposed features. Removal methods include mechanical removal, disolving in acetone using cotton buds, or exposing and developing the edge-bead areas (the most effective, though time consuming method). In addition, the thickness that the resist spins to can vary by several hundred nanometers across the surface, so great care must be taken when etching and using surface profiling to account for this.

For metallisation purposes it was found that SPR220 can effectively be used alone, without PMGI. Indeed, for samples with deep mesas ($> 4\ \mu m$) it was found to be the most effective method for bondpad/ track deposition to the tops of mesas. PMGI+SPR350 proved to be unable to effectively coat the tops of and sides of such mesas, which led to great difficulty during lift-off.

3.2.4 Hexamethyldisilazane

Hexamethyldisilazane (HMDS) is an adhesion promoter. It is used in processes where there is risk of the photoresist pealing away, such as during etching. Such an effect can be extremely detrimental to the etch profile, causing notches and rough edges to form. This is shown and discussed in more detail in section 3.6.2. Additionally, after a sample has been coated in silicon nitride or silicon oxide, the adhesion of photoresists during spin-coating is reduced. It is important to always use HMDS after a sample has either of these coatings applied, since without it the resist thickness and therefore the UV lithography exposure duration required can be highly inconsistent.

3.3 Wet etching trials

Three different trials were conducted to etch wafer SF0935 (wafer structure details in table 3.1), with images captured before and after the etches, including SEM of cleaved samples for detailed etch profile observation.

3.3.1 Trial A: H_2SO_4:H_2O_2:H_2O (1:8:80)

The first trial uses the sulfuric acid etch that has previously been used [48]. This was to allow a detailed comparison with the other etching trials. The sample was etched initially for 1 minute, after which the etch rate was characterised using a Dektak surface profiler. This showed an etch rate of 680nm/min.

With additional etches, one of 3m30s and the final etch lasting 1m45s, the total average etch depth was 3110nm, with a total etch time of 6m15s this means an average etch rate 518nm/min. This indicates that the etching solution lost potency over the duration of the etch. There was an extended period of time in-between etches while

Doping type	Material	Doping density (cm^{-3})	Thickness (nm)
N	$In_{0.53}Ga_{0.47}As$	5.0E18	100.0
N	$In_{0.52}Al_{0.48}As$	2.0E18	300.0
Undoped	$In_{0.52}Al_{0.48}As$	1.0E15	1000.0
P	$In_{0.52}Al_{0.48}As$	4.2E17	69.0
Undoped	$In_{0.52}Al_{0.48}As$	1.0E14	25.0
Undoped	$Al_{0.15}Ga_{0.32}In_{0.53}As$	1.0E14	25.0
Undoped	$Al_{0.29}Ga_{0.18}In_{0.53}As$	1.0E14	25.0
Undoped	$In_{0.53}Ga_{0.47}As$	1.0E14	1500.0
Undoped	$Al_{0.15}Ga_{0.32}In_{0.53}As$	1.0E14	25.0
Undoped	$Al_{0.29}Ga_{0.18}In_{0.53}As$	1.0E14	25.0
Undoped	$In_{0.52}Al_{0.48}As$	1.0E14	100.0
P	$In_{0.52}Al_{0.48}As$	2.0E18	100.0
P	$In_{0.53}Ga_{0.47}As$	5.0E18	1000.0
-	InP Substrate	-	-

Table 3.1: Layer profile of wafer SF0935, used for the wet etching trials. The p+ and n+ regions are highlighted.

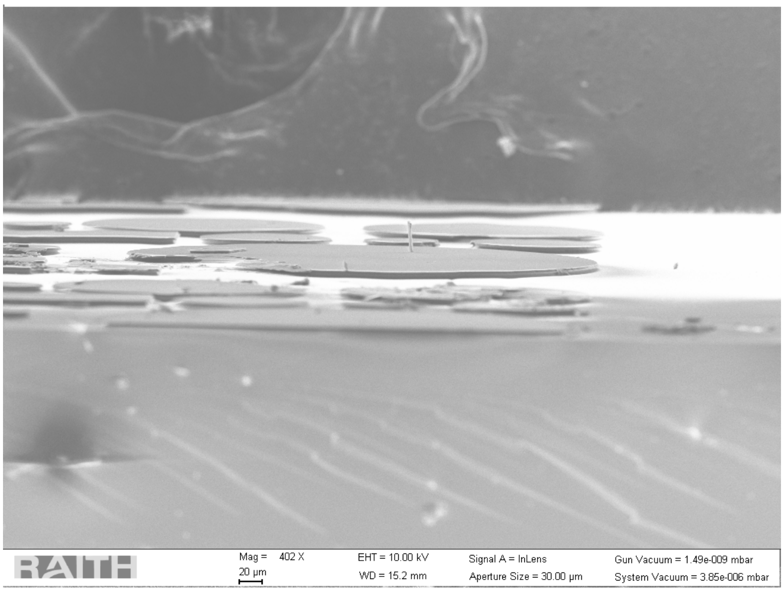
images were captured and the surface profiler was used. Accounting for time taken for observations, the entire etch process took over 1 hour. This is likely to result in a less effective etching solution.

Figure 3.7 shows the etching profiles obtained using SEM. The photoresist was present on the sample for quantifying the undercutting effects of the etch. The etchant undercut the photoresist by a maximum of 6.3 μm, with the etching angle appearing to depend on the material. The top layers (1.4 μm) are mostly InAlAs, and the bottom layers mostly InGaAs. There is a clear change of etch angle corresponding to InAlAs-InGaAs interface.

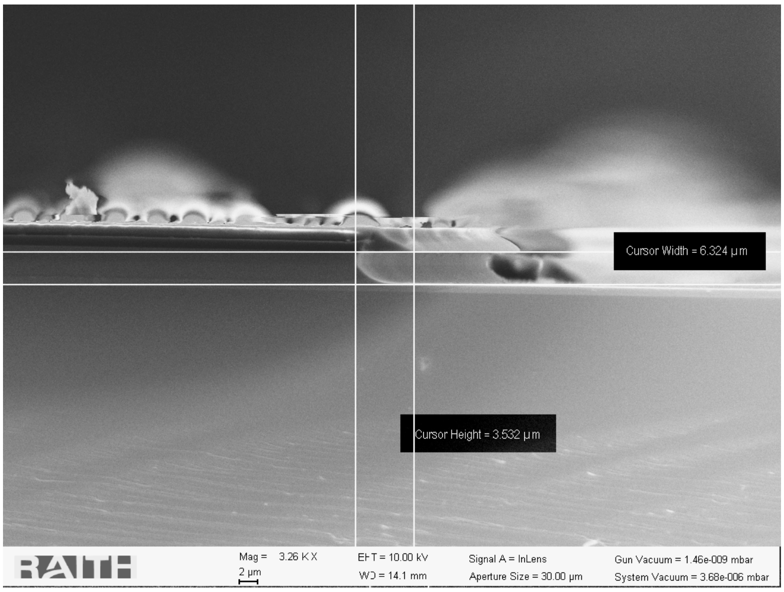
3.3.2 Trial B: H_3PO_4:H_2O_2:H_2O (1:8:1)

The etchant used in this trial was expected to etch more quickly compared to trial A. After etching for 1 minute, the samples surface profile confirmed this with an etch rate measured of 1350 nm/min. After a second 1 minute etch the depth had increased by an additional 1300 nm, with total etch depth of 2650 nm. This etchant would be especially useful if thick layers of InAlAs/InGaAs need to be etched, since it is rapid and doesn't appear to be hindered by intermediate layers in-between.

The etch profiles seen in figures 3.8a and 3.8b also showed favourable results, indicating a smooth, consistent etch angle independent of layer composition. Unfortunately it is not



(a) Zoomed out side view of the etched devices. The mesas and general undercutting of the photoresist is visible.



(b) Etching profile with undercutting and etch depth measurements. The photoresist is overhanging the mesa where the acid has etched under it. The halo like scattering effects are due to the photoresist.

Figure 3.7: SEM images of trial A.

clear if this profile is representative of the whole sample, as the sample shattered during cleaving prior to measurement so only a limited amount of mesas were observable.

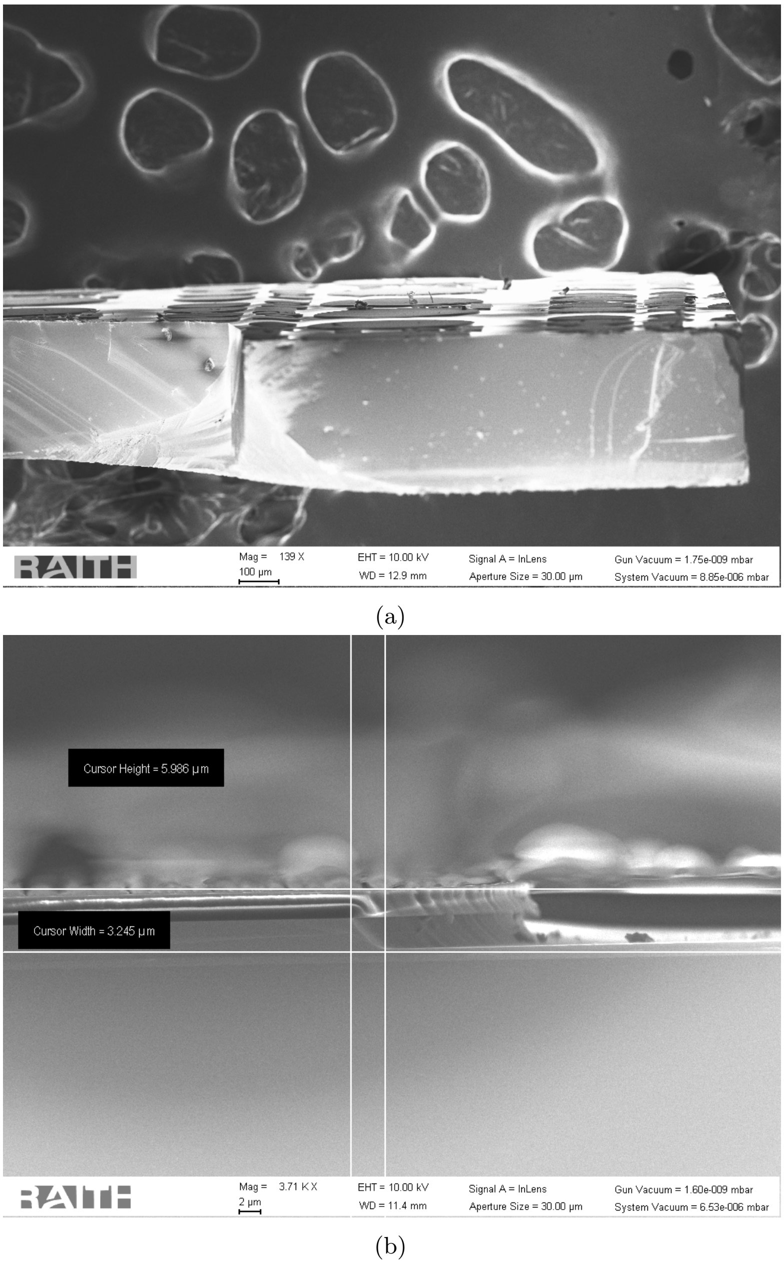


FIGURE 3.8: SEM images of the sample from trial B.

3.3.3 Trial C: $C_6H_8O_7$:H_2O_2 (2:1) (InGaAs) and HCl:H_2O (3:1) (InAlAs)

Unlike the previous trials, this trial used two selective etchants. Etchant 1 - $C_6H_8O_7$:H_2O_2 (2:1) - etched InGaAs, while etchant 2 - HCl:H_2O (3:1) - etched InAlAs. This provides greater control over the etch, as etching will essentially stop when it reaches the appropriate layer, reducing the risk of over-etching.

The sample was spun with SPR350 and exposed with the mesa etch mask. The etch was performed in several stages to determine the selectivity of the etch. The sample was submerged in etchant 1 for 45s to remove the 100 nm InGaAs contact layer. Surface profiling showed an average etch depth of 90 nm. This provided an opportunity to test the selectivity of etchant 2. Etchant 2 was used for 2min15s - the time expected to etch to the next InGaAs layer. Surface profiling after this etch showed an average etch depth of 0, though the surface had been roughened, as shown in figure 3.9b. As expected, etchant 2 was unable to remove the remaining 10 nm of InGaAs. The sample was placed back in etchant 1 for 30 seconds, with surface profiling showed an average etch of 9 nm.

Finally, the sample was etched with etchant 2 for 3 minutes to etch the layers InAlAs. An average etch depth of 1430 nm indicated that it had stopped etching at the InAlAs/InGaAs interface. The surface roughness was observed to have reduced. The etching was terminated here, and the sample cleaved for observation with an SEM. See figures 3.9a and 3.9b for profiles.

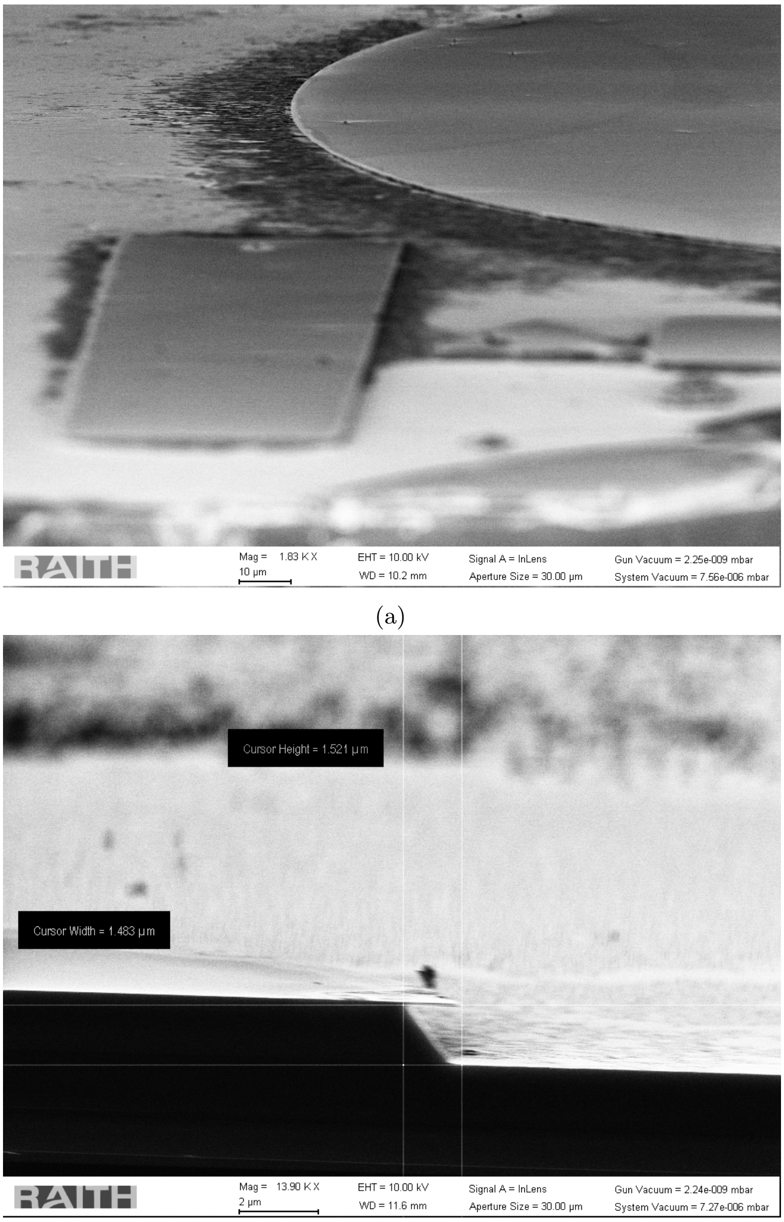
The most striking feature of the SEM images is the rough surface around the mesas. This is suspected to be the quaternary structures in-between the InAlAs and InGaAs layers which have not been fully etched away. It seems that the selective etchants struggle with penetrating through these layers. Additionally the selective etching has created an overhanging InGaAs layer on top of the under-etched InAlAs layer. This has occurred since the InAlAs has been etched for longer selectively, undercutting the unaffected InGaAs layer on top. This could cause problems with the device performance due to increased surface around the device edges.

3.3.4 Comparison

Trial	Etchants	Duration (s)	Depth (nm)	Rate (nm/s)
A	H_2SO_4:H_2O_2:H_2O (1:8:80)	375	3110	518
B	H_3PO_4:H_2O_2:H_2O (1:8:1)	120	2650	1325
C	$C_6H_8O_7$:H_2O_2 (2:1)	45	90	120
	HCl:H_2O (3:1)	180	1430	480

Table 3.2: Comparison of etchants from the trials.

A comparison of the etches can be found in table 3.2. Etchant B proved to be the fastest etch of the three. This means that it is a good choice for applications that do not require a lot of etch depth accuracy. Indeed, for most wafers to be used for this project



(a)

(b)

FIGURE 3.9: SEM images of the sample from trial C.

the full etch would only take 3 minutes. This means that it is better for general etching than etchant 1 that was previously used.

For a more controled etch, A and C are more suited. They are useful if a more complex and accurate etch process is required, such as stopping at the thin, central P doped layer. In such a case etch C would offer slightly more control due to the slower etch rate for InAlAs, but being selective it allows for wafers to be designed with this in mind. Stopping layers can be engineered for etch C to ensure that over-etching does not occur past the desired point without switching to the second compound. Currently,

Doping type	**Material**	**Doping density (cm^{-3})**	**Thickness (nm)**
P	$In_{0.53}Ga_{0.47}As$	2.0E18	200
Undoped	$In_{0.53}Ga_{0.47}As$	0	2000
Undoped	$In_{0.52}Al_{0.48}As$	0	25
P	$In_{0.52}Al_{0.48}As$	1.6E17	175
Undoped	$In_{0.52}Al_{0.48}As$	0	1000
N	$In_{0.53}Ga_{0.47}As$	5.0E18	1000
-	InP Substrate	-	-

Table 3.3: SF0751 concessionairy wafer structure.

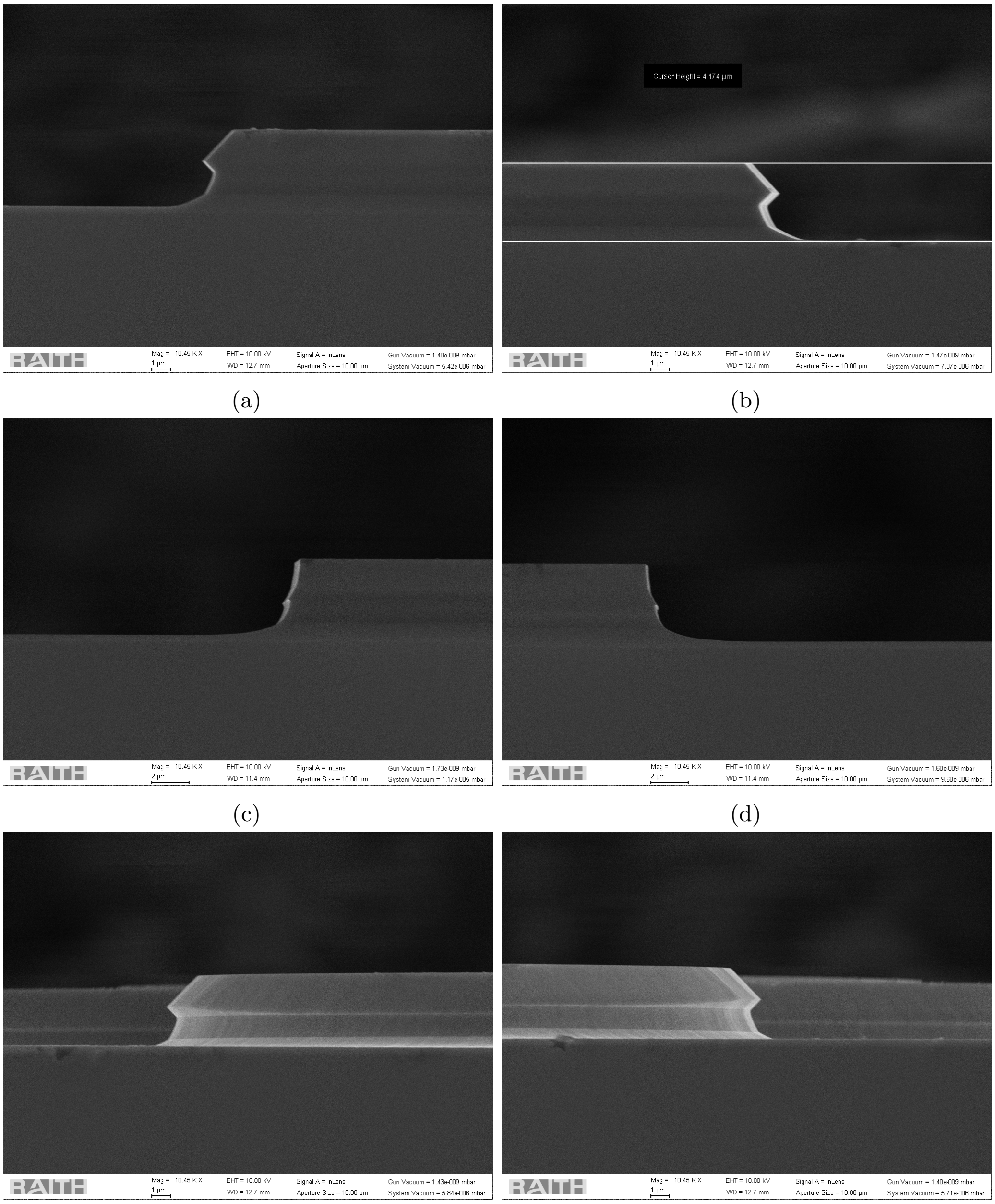
there is no significant advantage to etch C over etch A with the current wafers, however designs of further wafers could take advantage of the material selectivity. Etch A is selected as the main etchant for all further fabrication used in this work, as it offers the best tradeoff between etch profile and control.

3.4 Anisotropic Etch Directions

Tests were performed to evaluate etch profiles of SAMAPD structures when using the H_2SO_4:H_2O_2:H_2O (1:8:80) etching solution. Such an investigation is important as it determines which is the best direction to deposit the bond pad tracks in.

The wafer structure of the sample used in these trials is shown in table 3.3, and was selected due to its lack of quaternery layers. An important factor to note is that the wafer is p on n, unlike the majority of the wafers in this work so the etch profiles would be reversed between such wafers. The trial started with the sample having a 3-stage clean, before being patterned using HMDS+SPR350 with the deep mesa etch photolithography mask. Care was taken to keep track of the wafer major and minor flat directions. The etching took 7 minutes, resulting in an etch depth of 4.2 μm. The photoresist was then removed by acetone. Finally, the sample was cleaved down the centre of the mesas along perpendicular directions. SEM imaging of the cleaved samples captured the etch profiles along the different directions, as shown in figure 3.10.

A summary of the etch profiles is shown in table 3.4. It is found that while the InAlAs etch profile is largely independent of the lattice plane, InGaAs has significant difference with etch angles ranging from 47 to 77 degrees. The 011b direction is determined to be most suitable for bond-pad deposition, since it offers the smoothest profile - minimising



(a) (b)

(c) (d)

(e) (f)

FIGURE 3.10: SEM images of mesa sidewalls in the two cleave directions. (a) and (b) show the 011b etch direction. (c) and (d) show the 011 etch direction. (e) and (f) show the transition of mesa sidewall from 011b to 011.

the chance of bond-pad breakage. Additionally the flatter surface means that nitride will deposit more effectively in this direction.

An interesting observation to note about the etch profile has sharp edges within the InGaAs layer close to the InGaAs/ InAlAs interface. While this is not expected for n on p wafers due to the reversal of materials, for a p on n wafer the performance of the diodes could be seriously hampered due to the electric field hotspot that is likely occur

Material	**011b**	**011**
InGaAs	47^{o}	77^{o}
InAlAs	70^{o}	65^{o}

Table 3.4: Summary of the etch angles determined from the SEM images. Angles are taken from the horizontal plane.

here. This suggests that this etchant would be unsuitable for fabrication of a p on n SAMAPD.

3.5 Characterisation Techniques

After a successful fabrication run, the resulting devices must be tested and characterised to determine their quality. There are several different characterisation steps that the devices on a sample will go through, with some depending on the success of the previous step. The first step in any device characterisation is a room temperature IV trace.

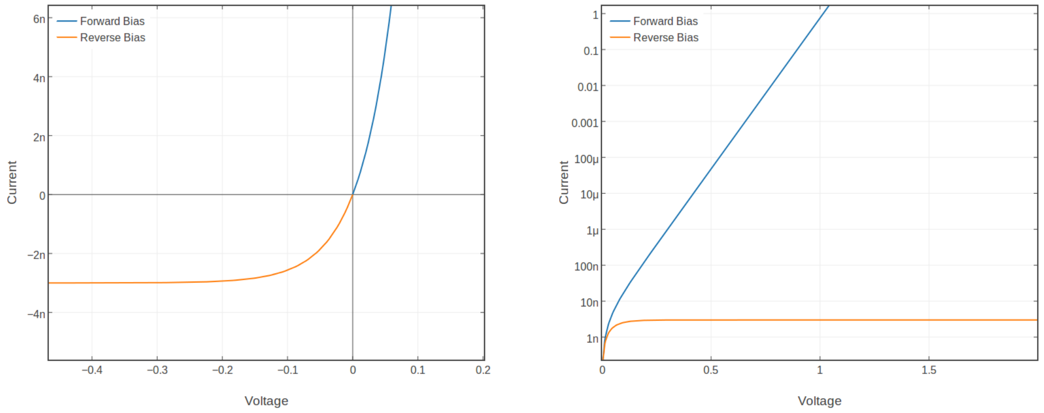
3.5.1 IV

The experimental set-up for measuring IV characteristics is relatively basic. The diode p and n contacts are probed and connected to a source-measure unit (SMU). The SMU applies a DC bias for forward and reverse bias measurements. The forward bias sweep is generally used to verify that good connections have been made, with all diodes expected to have the same current densities at low forward biases.

The reverse bias sweep gives valuable information on the quality of the diode. Deviation from the expected "ideal" dark current indicates undesirable mechanisms causing leakage current. These can include generation-recombination and surface effects such as surface and edge breakdown [12]. The severity of these will influence the decision as to whether it is worth pursuing further characterisation of the diode. The ideal I-V relation for a diode is given by the Shockley diode equation:

$$I = I_s(\exp(V/n_i V_T) - 1), \tag{3.1}$$

where I_s is the reverse saturation current, n_i is the ideality factor, and $V_T = kT/q$ is the thermal voltage. This is plotted in figure 3.11.



(a)

(b)

FIGURE 3.11: (a) Linear and (b) semi-log plots of the IV relation for an ideal diode plotted with arbitrary units. The reverse saturation current is set to 3×10^{-9}.

Preliminary reverse breakdown voltage can be obtained from reverse IV measurement, though this is not modelled by equation 3.1. Additionally, illuminating the diode with white light allows for punch-through voltage to be determined.

3.5.2 Responsivity and Gain

Responsivity is measured by focusing a laser source onto the optical window of a reverse biased diode. Prior to measurement, the power of the laser is characterised using an optical power meter. An SMU is used to bias the diode and measure the dark current while the laser is off, and the photocurrent generated while the diode is illuminated by the laser. The dark current is subtracted from the photocurrent, and the result divided by laser power to obtain the photo-response of the diode.

Gain is extracted from the data by selecting the photocurrent at the first voltage point where the diode has fully depleted as the reference photocurrent, which is expected to have a gain of one. All subsequent photocurrents measuremented after this point are divided by this value, providing gain. It is important to use large diodes to ensure that all of the laser is focused into the optical window.

For diodes with high dark current or weak photoresponse, it is necessary to use a lock-in amplifier (LIA) for a phase-sensitive measurement. As before, the diode is illuminated with a laser, however the laser light is passed through an optical chopper which modulates the laser. The LIA then measures the relative voltage drop across a load resistor connected in series with the diode. The reference frequency from the optical chopper is supplied to the LIA, so that the LIA only measures the optical signal at the reference frequency. This allows it to differentiate the modulated photocurrent from the dark currents.

3.5.3 CV

CV measurements are carried out using an LCR meter. The p and n contacts of the device are connected to the LCR meter, which then applies a small AC voltage superimposed onto a DC voltage which is adjusted. The device is thus reverse-biased for the capacitance measurement.

CV data can reveal detailed information on the structure of the APD. Sudden increase in the depletion region width is often important for a SAMAPD, since it produces a sudden decrease in capacitance. This corresponds to the punch-through voltage, which may be more easily observed in CV compared to IV data. Additionally, by solving Poisson's equation (see section 2.1) it is possible to fit the data and estimate the electric field profile across the device, and consequently the device structure and doping levels. This is useful as it verifies whether the structure of the device is as expected, or if there have been issues during wafer growth.

3.5.4 Dark Count Rate

For single photon counting, the dark count rate is one of the most important characteristics of a SPAD. Minimising this is a central aim of this project, so it is important to accurately characterise it.

3.5.4.1 CQC Board

The experimental setup, shown schematically in figure 3.12, was used to test the DCR packaged devices. The device is placed in the DUT circuit which contains the CQC (designed and constructed by Dr. S. Dimler [74]).

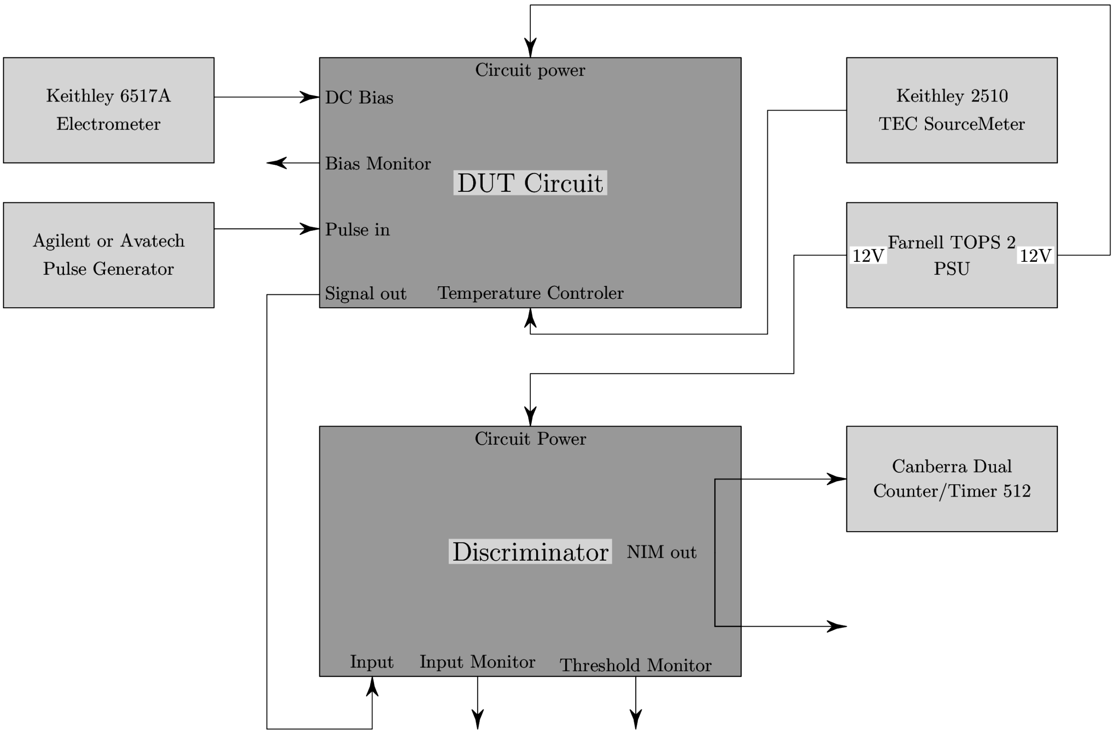


FIGURE 3.12: Experimental setup for measurements of DCR for packaged devices.

The measurement process is as follows. The temperature controller is set to maintain the device at the desired temperature, in a range of 0-30 °C. It is important to wait for temperature to stabilise, as the breakdown voltage of the device can depend heavily on the temperature. Once stable, a DC bias is applied to the circuit which is overlayed with a 20ns square-wave pulse of around 5 V. The output of the DUT circuit is fed through the discriminator, where the input can be monitored using an oscilloscope.

As the DC bias is stepped up, the oscilloscope is used to monitor avalanche events, which appear as sharp, sudden peaks above the noise floor. Once these appear, the device is at breakdown voltage. The discriminator threshold is adjusted such that it is above any noise. The NIM output of the discriminator is fed to a counter to measure the DCR. The DCR is then measured as a function of bias over the breakdown voltage.

The DCR is defined as

$$\mathrm{DCR} = \frac{N^{\underline{o}}\ \text{dark counts}}{\text{duration} \cdot \tau f}, \tag{3.2}$$

where τ is on-time, and f is frequency. The dark count probability is

$$P_d = \frac{N^{\underline{o}}\ \text{dark counts}}{f}. \tag{3.3}$$

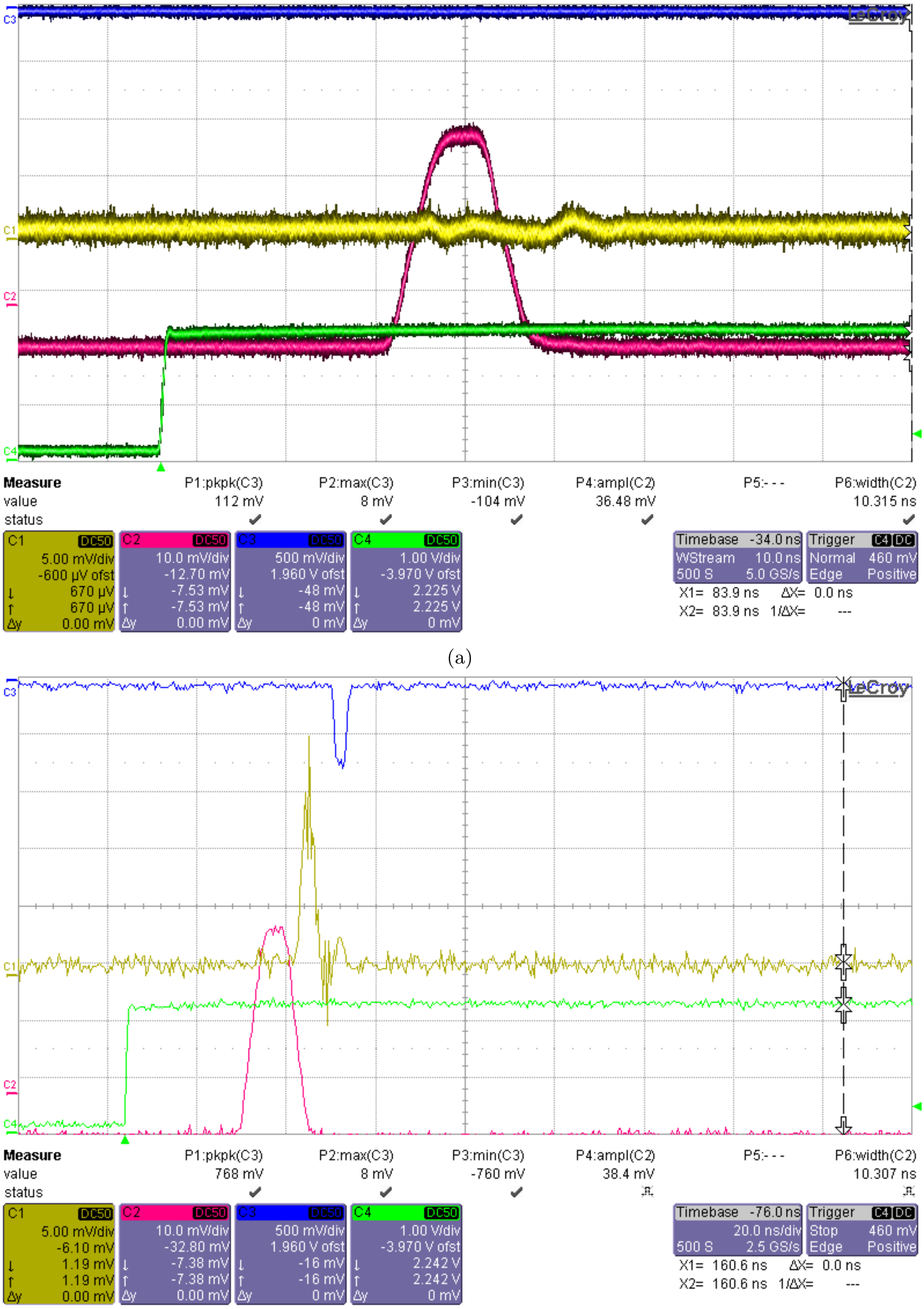
Figure 3.13 shows examples of the signals observed during the operation of the CQC board with a SPAD operated in gated mode. In figure 3.13a there is only the 5 V biasing pulse (C2) applied, with no DC offset. The small transients as a result of the pulse can be seen in channel C1 - small due to the differential amplifier component of the circuit, with the variable capacitor adjusted to match the capacitance of the diode. Channel C4 is the gate output of the pulse generator. A demonstration of the circuit with a DC bias applied to a combined pulse and DC reverse voltage above breakdown is shown in figure 3.13b. Channel C1 shows an avalanche breakdown event occuring and being detected by the discriminator circuit, with the NIM output trigger appearing in channel C3. The NIM output is input into the Canberra counter, which is able to count the number of events that have occured within a given time period. Using this result, it is possible to calculate the normalised dark count rate and the dark count probability using equations 3.2 and 3.3.

3.5.4.2 JANIS ST-500 Low Temperature Probe Station and LeCroy Oscilloscope

For low temperature measurements, the JANIS ST-500 probe station is required. The JANIS provides an enclosed, temperature controlled vacuum chamber with probes inside able to be manipulated externally. This allows for the measurement of multiple devices when the chamber is evacuated and cooled down.

Due to the limited space within the chamber, it is not possible to have the differential amplifier used in the CQC circuit close to the diode. Indeed, due to the inherent capacitance and inductance of the chamber and cabelling itself, large transients are generally observed when pulses are applied to a diode within the chamber. The problem is demonstrated in figure 3.14. For a wide pulse, as in figure 3.14a, the avalanche breakdown events mostly reach above the transients. With a reduced pulse width, shown in figure 3.14b, breakdown events are not able to reach above the transients since they are unable to build up enough charge. A new method of counting is required, since the size of these transients prohibits the use of a simple threshold circuit due to the majority of avalanche events not reaching a level higher than the transients.

A method to use a LeCroy Waverunner oscilloscope to filter and count events was devised. Figure 3.15 shows an example of this method in operation and detecting a



Measure	P1:pkpk(C3)	P2:max(C3)	P3:min(C3)	P4:ampl(C2)	P5:- - -	P6:width(C2)
value	112 mV	8 mV	-104 mV	36.48 mV		10.315 ns
status	✓	✓	✓	✓		✓

C1 DC50	C2 DC50	C3 DC50	C4 DC50
5.00 mV/div	10.0 mV/div	500 mV/div	1.00 V/div
-600 µV ofst	-12.70 mV	1.960 V ofst	-3.970 V ofst
↓ 670 µV	↓ -7.53 mV	↓ -48 mV	↓ 2.225 V
↑ 670 µV	↑ -7.53 mV	↑ -48 mV	↑ 2.225 V
Δy 0.00 mV	Δy 0.00 mV	Δy 0 mV	Δy 0 mV

Timebase	-34.0 ns	Trigger	C4 DC
WStream	10.0 ns	Normal	460 mV
500 S	5.0 GS/s	Edge	Positive

X1= 83.9 ns ΔX= 0.0 ns
X2= 83.9 ns 1/ΔX= ---

(a)

Measure	P1:pkpk(C3)	P2:max(C3)	P3:min(C3)	P4:ampl(C2)	P5:- - -	P6:width(C2)
value	768 mV	8 mV	-760 mV	38.4 mV		10.307 ns
status	✓	✓	✓	⎍		⎍

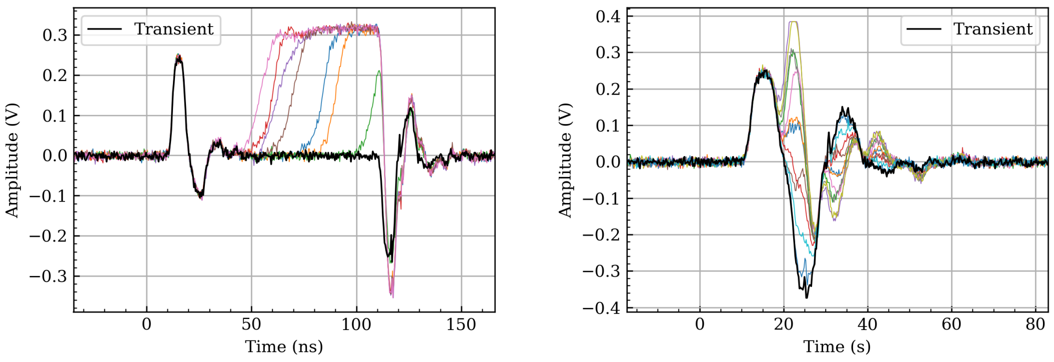
C1 DC50	C2 DC50	C3 DC50	C4 DC50
5.00 mV/div	10.0 mV/div	500 mV/div	1.00 V/div
-6.10 mV	-32.80 mV	1.960 V ofst	-3.970 V ofst
↓ 1.19 mV	↓ -7.38 mV	↓ -16 mV	↓ 2.242 V
↑ 1.19 mV	↑ -7.38 mV	↑ -16 mV	↑ 2.242 V
Δy 0.00 mV	Δy 0.00 mV	Δy 0 mV	Δy 0 mV

Timebase	-76.0 ns	Trigger	C4 DC
	20.0 ns/div	Stop	460 mV
500 S	2.5 GS/s	Edge	Positive

X1= 160.6 ns ΔX= 0.0 ns
X2= 160.6 ns 1/ΔX= ---

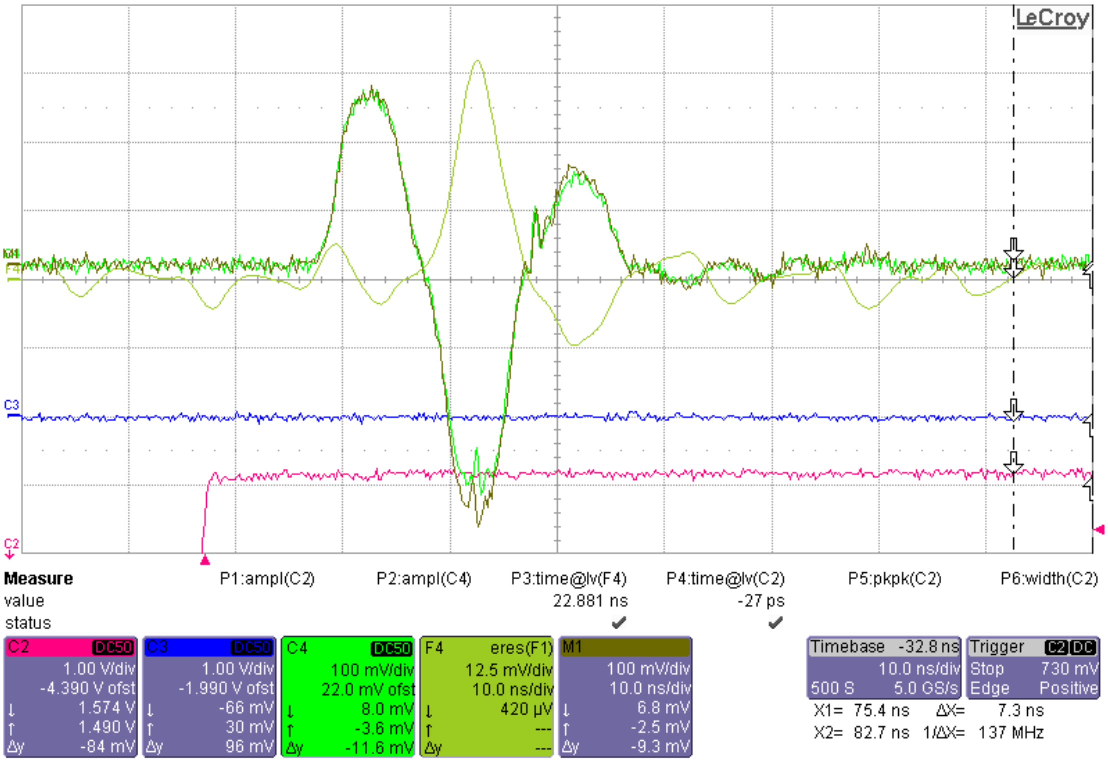
(b)

FIGURE 3.13: A demonstration of the biasing pulse (5 V, 10 ns), transients, and signal from the diode. C1 is the signal monitor from the discriminator, C2 is the pulse bias monitor, C3 is the NIM output from the discriminator, and C4 is the trigger output from the pulse generator. (a) shows the transient resulting from the biasing pulse. It is minimised by the differential amplifier and variable capacitor built into the CQC board. (b) shows an observed breakdown event on C1. Note the NIM trigger signal on C3 from the discriminator. There is a delay between events due to different signal path lengths.



(a) (b)

FIGURE 3.14: Traces of breakdown events recorded from the oscilloscope, with the transients indicated by the black line.



```
Measure      P1:ampl(C2)   P2:ampl(C4)   P3:time@lv(F4)   P4:time@lv(C2)   P5:pkpk(C2)   P6:width(C2)
value                                      22.881 ns           -27 ps
status                                         ✔                  ✔

C2        DC50   C3        DC50   C4        DC50   F4      eres(F1)   M1                  Timebase  -32.8 ns   Trigger    C2 DC
     1.00 V/div       1.00 V/div      100 mV/div       12.5 mV/div        100 mV/div          10.0 ns/div   Stop     730 mV
  -4.390 V ofst    -1.990 V ofst     22.0 mV ofst      10.0 ns/div        10.0 ns/div   500 S    5.0 GS/s   Edge   Positive
↓      1.574 V  ↓        -66 mV   ↓      8.0 mV   ↓       420 µV   ↓          6.8 mV     X1=  75.4 ns    ΔX=     7.3 ns
↑      1.490 V  ↑         30 mV   ↑     -3.6 mV   ↑          ---   ↑         -2.5 mV     X2=  82.7 ns  1/ΔX=   137 MHz
Δy      -84 mV  Δy        96 mV   Δy   -11.6 mV   Δy         ---   Δy        -9.3 mV
```

FIGURE 3.15: LeCroy filtering (see text for details).

small avalanche event. The oscilloscope is able to apply mathematical operations on the signal from the measured channels. Before the diode is placed into the SPAD gated operating regime (with biasing pulses taking overall reverse bias above breakdown), the transient signal from channel C4 is recorded and stored into memory bank M1. This stored signal is subtracted from the channel C4 signal, and a smoothing function applied to the result to reduce the effects of noise. The resulting, filtered signal can be seen in F4. The histogram function of the oscilloscope (not shown) is then used and set to trigger on channel C2 (the gate output of the pulse generator), and at a certain threshold to trigger on F4. Thus, based on the number of events that are recorded from channel C2 and F4, the probability of an event can be determined as

$$P = \frac{\text{channel F4 counts}}{\text{channel C2 counts}}. \tag{3.4}$$

Depending on if photons are present or not during the measurement, the dark count or photon detection probabilities can be determined.

Due to the speed limitation of the oscilloscope, this method requires approximately 45 seconds to 1 minute to measure 1000 pulses. It is therefore only utilised when absolutely required as here, since this greatly limits measurement accuracy at lower count rates due to time constraints. Additionally, software was developed to control the oscilloscope since the histograms are only able to count up to 1000 events before a reset is required. This allowed for easier extended measurements to 10000 counts per voltage.

3.5.5 Photon Detection Efficiency

For photon detection efficiency (PDE) measurement, the CQC board can be used for packaged devices, and the low temperature probe station for unpackaged devices. The CQC board is enclosed in a large black box which is fitted with optics mounted on electronic micro-positioners to couple light into the diode, and a peltier inside for temperature stabilisation. Using the low temperature probe station allows for positioning of a micrometer adjusted single-mode optical fiber which couples light into the chamber from an outer fiber connection.

In both cases, a CW 1550 nm laser with a known power is used to initially align the optics to the DUT, which is reverse biased above punch-through. For a diode with known responsivity, it is possible to calculate the amount of the laser spot coupled into the optical window. The loss is recorded as μ_{coupling} (dB), calculated using

$$\mu_{\text{coupling}} = 10 \cdot \log_{10}\left(\frac{P_{\max} \cdot R}{I_p}\right), \tag{3.5}$$

where $P_{\max}$ is the maximum measured power at the end of the optical fiber, R is diode responsivity, and I_p is measured photocurrent.

A 1550 nm picosecond pulsed laser (Picosecond Laser Diode Systems PiL155) is used to provide the laser pulses for photon counting measurements, triggered by the pulse generator and synchronised with the biasing pulse. This is connected through an Exfo FVA-3100 electrical variable optical attenuator (EVOA) (μ_{variable}), which is also fitted with a 25 dB fixed optical attenuator (μ_{fixed}) on the input port to protect it from damage. With a measured average pulse energy of $E_{\text{pulse}} = 248$ fJ for the pulsed laser, and photon energy $E_{\text{photon}} = 1.28 \times 10^{-19}\ J$, a total of 62.87 dB of attenuation is required to attenuate the pulses to a single photon level, $\bar{n} = 1$. Total system attenuation is given by

$$\mu_{\text{total}} = \mu_{\text{fixed}} + \mu_{\text{variable}} + \mu_{\text{coupling}}. \tag{3.6}$$

Using this, the average number of photons per pulse can be calcuated using

$$\bar{n} = \frac{E_{\text{pulse}}}{E_{\text{photon}} \cdot 10^{\frac{\mu_{\text{total}}}{10}}}. \tag{3.7}$$

μ_{variable} is then adjusted to set $\bar{n}$ to the desired level.

Synchronisation of the pulses with the biasing pulse is performed using the pulsed laser unattenuated. With a large diode (low coupling loss) reverse biased to a high gain region, it is possible to view the incoming laser pulse using an oscilloscope. The offset relative to the pulse generator trigger out is recorded, and then used to delay the biasing pulse sufficiently such that the peak of the pulse falls as the biasing pulse peaks.

Dark and pulsed laser measurements are performed back-to-back, ensuring that the pulse detection probability is more reliable. With time, the DCR can drift, so it's important to measure these close together for a given voltage. In general, several different attenuations are used to provide an understanding of how detection probability changes with photon level. The pulse detection probability is given by

$$P_{\text{pde}} = \frac{\text{Number of pulsed counts}}{f} - P_d. \tag{3.8}$$

Doping type	**Material**	**Doping density (cm^{-3})**	**Thickness (nm)**
N	$In_{0.53}Ga_{0.47}As$	5.0E19	100
N	$In_{0.52}Al_{0.48}As$	2.0E18	300
Undoped	$In_{0.52}Al_{0.48}As$	1.0E14	1000
P	$In_{0.52}Al_{0.48}As$	4.2E17	69
Undoped	$In_{0.52}Al_{0.48}As$	2.0E15	25
Undoped	$Al_{0.15}Ga_{0.32}In_{0.53}As$	1.0E14	25
Undoped	$Al_{0.29}Ga_{0.18}In_{0.53}As$	1.0E14	25
Undoped	$In_{0.53}Ga_{0.47}As$	2.0E15	1500
Undoped	$Al_{0.15}Ga_{0.32}In_{0.53}As$	1.0E14	25
Undoped	$Al_{0.29}Ga_{0.18}In_{0.53}As$	1.0E14	25
Undoped	$In_{0.52}Al_{0.48}As$	2.0E15	100
P	$In_{0.52}Al_{0.48}As$	2.0E18	100
P	$In_{0.53}Ga_{0.47}As$	5.0E18	1000
-	InP Substrate	-	-

TABLE 3.5: SF0940 wafer structure, used for simulation and fabrication of double mesa APDs.

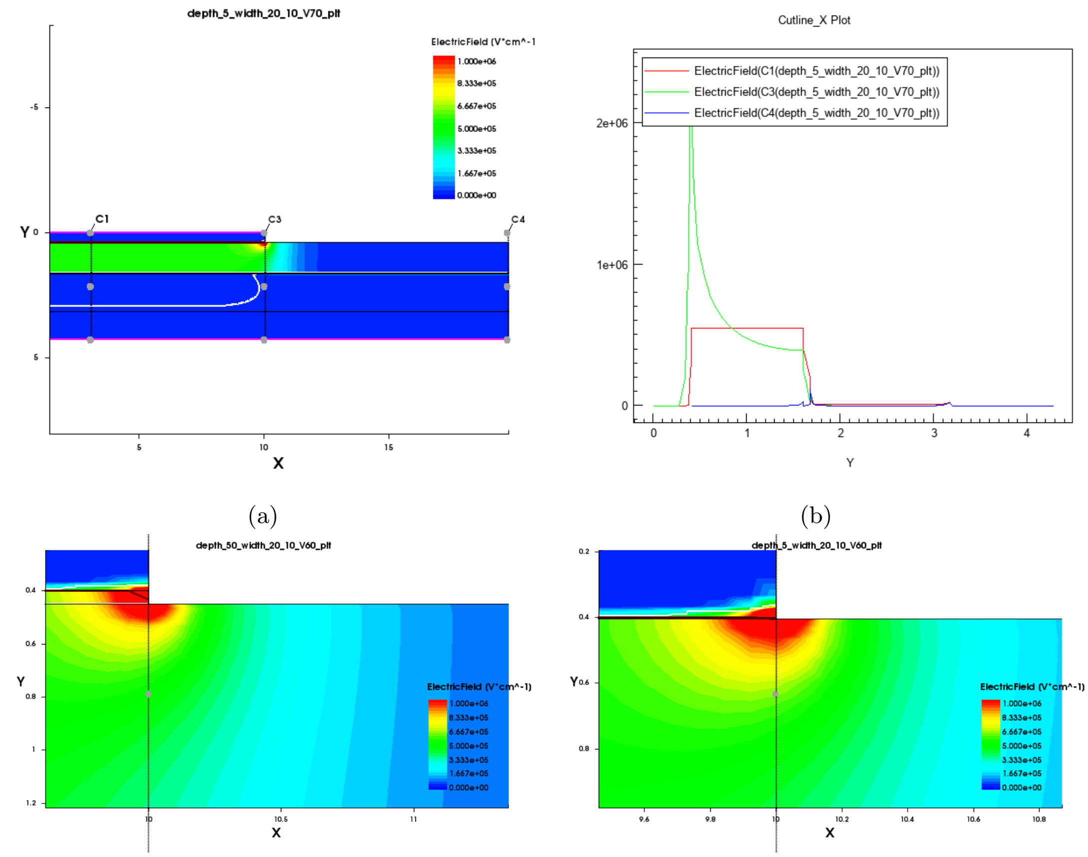
3.6 Double Mesa Design and Fabrication

Multi mesas are structures which are generally used to confine the electric field within a diode to a certain area, which is defined by the top mesa area. Specially designed wafers have in the past been used to realise these multi-mesa devices [27]. To test if there is any benefit to utilising existing non-specially designed wafers with a double-mesa fabrication process, simulation and fabrication of such a design was performed.

3.6.1 Design and Simulation

An APD wafer (SF0940) with relatively small punch-through voltage was selected. The existing wafer is given in table 3.5, and is used in the electric field profile simulations which are performed using semicondutor simulation software; TCAD Sentaurus. Several different mesa widths were tested, with the first (top) mesa width being varied (from 5 to 20 μm) to find a suitable distance that its edge had to be from the second mesa to provide adequate field confinement. Results of the simulations are presented in figure 3.16.

In figure 3.16a the full 2D modelled structure is shown, with electric field slices C1, C2 and C3 taken at important locations and plotted in figure 3.16b. C1 represents the regular field that one expects within a standard, single mesa diode with none of the



(a) (b)

(c) (d)

FIGURE 3.16: Double mesa simulations using the structure given in table 3.5 with the top mesa edge 10 μm away from the outer mesa. In all plots the diode has a reverse bias of 70 V applied to it. (a) shows the electric field across the diode in the region of interest, with slices C1, C2, and C3 taken at the indicated points and plotted in (b). (c) and (d) show the effects of etching further into the avalanche region, and how the electric field hot spot changes.

effects of electric field confinement visible. This is contrasted with the fields at C2 and C3, with C2 showing a large peak due to the sharp edge at the double mesa interface. C3 shows that the field is completely confined away from the second mesa edge. The electric field hotspot at C2 means that the voltage range the APD is able to operate at will be reduced due to edge breakdown effects, since the breakdown field will be reached in that area earlier than the junction as a whole.

Figures 3.16c and 3.16d explore the hotpsot further, showing a zoomed in view of the hotspot at two different etch depths into the InAlAs avalanche region of 50 nm and 5 nm respectively. Etching further into the region distributes the hotspot over a larger area, as well as causing a breakdown field across more of the first mesa surface. Therefore care must be taken during fabrication of the double mesa to minimise the depth etched into the avalanche region in order to maximise the APD performance.

With the field confined largely to within a 2-3 μm radius of the top mesa edge, 10 μm is chosen as the distance at which the second mesa will be defined and etched. This takes into account possible undercutting of the photoresist during etching, with a target total etch depth of around 4 μm possibly bringing parts of the sidewall to within 6 μm of the top mesa edge. This leaves around 3 μm for the alignment tollerance of the second mesa etch stage if field confinement is to be ensured.

3.6.2 Device Fabrication of Double Mesa Diodes

Fabrication was perfomed based on the standard procedure discussed in 3.1. One sample was cleaved from the SF0940 wafer and 3-stage cleaned. A spin coating of PMGI was applied, followed by a 5 minute bake on a hotplate at 180 °C. A layer of SPR350 was spun on top and baked for 1 minute at 100 °C. The IIG Mesa top contact mask was exposed in a UV mask aligner, and developed for 1 minute in MF26A photoresist developer. Ti/Au (20/200 nm) contacts were then deposited in an evaporator. After metal lift off in acetone, and 1 minute in MF26A developer to remove the PMGI, SPR350 was spun on and baked for 1 minute at 100 °C. The standard mesa mask was then exposed and developed for 1 minute.

At this stage the sample was cleaved in half before etching. The etchant used was H_2SO_4:H_2O_2:DIW (1:8:80). One half of the sample was etched to a depth of 4 μm, to be used as a reference sample to compare the double mesa against. The second half was etched 450 nm (to ensure full removal of the top n layers) after which the photoresist was stripped, and a spin coating of Hexamethyldisilazane (HMDS) adhesion promoter and SPR350 was applied and patterned with the deep mesa etch layer. The HMDS was crucial at this stage, as the etched mesa shapes negatively affect photoresist coating and adhesion, which can cause serious issues during etching. The deep mesa sample was then etched a further 3.6 μm down to the bottom contact layer for a total etch of 4 μm. The importance of HMDS is demonstrated in figure 3.17, which shows SEM images of a failed double mesa where HMDS was not used. The smaller mesa is visible above the second mesa, however the outer mesa is extremely rough due to the photoresist pealing off around the edges as the etch was performed.

The final stage of processing was the patterning of the grid contact layer onto a spin coating of SPR220. The SPR220 was baked on a hotplate for 1 minute 40 seconds at

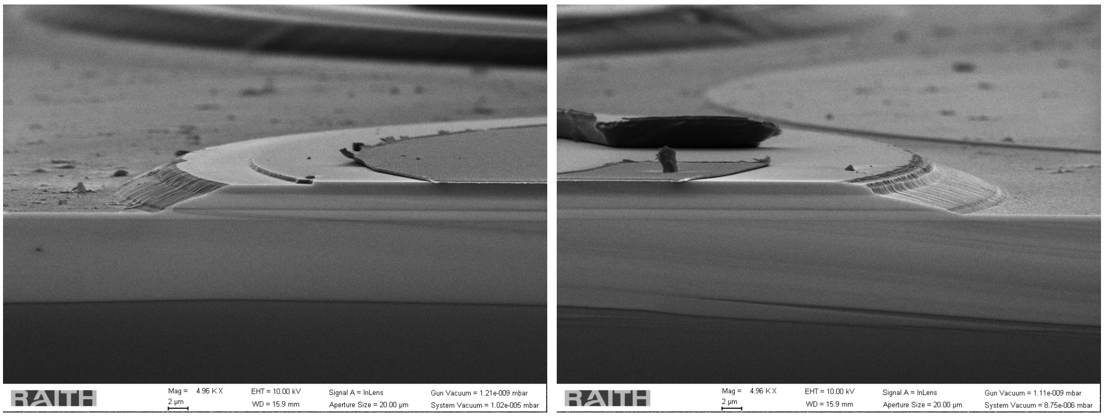


FIGURE 3.17: SEM images of a double mesa diode fabricated without using HMDS. Extreme undercutting of the photoresist has brought the wider mesa edge close to the top mesa, and led to a rough sidewall.

100 °C before exposure, and developed for 1 minute 10 seconds in MF26A after exposure. 20/200 nm Ti/Au contacts were then evaporated onto the sample, and lift-off performed. Completed devices are shown in 3.18.

3.6.3 Results of Double Mesa Fabrication

IV data of the fabricated single mesa and double mesa diodes are compared in figure 3.19. The double mesa diodes show a smaller breakdown voltage, confirming the electric field hotspots observed in the simulations cause premature breakdown. The breakdown was not catastrophic, and repeat measurements could be made on the same double mesa diode. The dark current is lower in the double mesas prior to punchthrough. This indicates that the double mesa structure is effective at reducing dark current density in the InAlAs avalanche region by up to an order of magnitude. After punchthrough occurs however, the dark currents are indistinguishable between the single mesa and the double mesa diodes.

These data indicate the dominant dark current mechanisms in the InAlAs avalanche and InGaAs absorption regions. The InAlAs dark current is likely dominated by surface effects from the etching, since confining the electric field away from surface reduces dark current significantly. The double mesa structure could therefore be useful to quantify the effects of various different etches and optimise for reduced surface dark current. Dark current in the InGaAs region is dominated by thermal generation - recombination current,

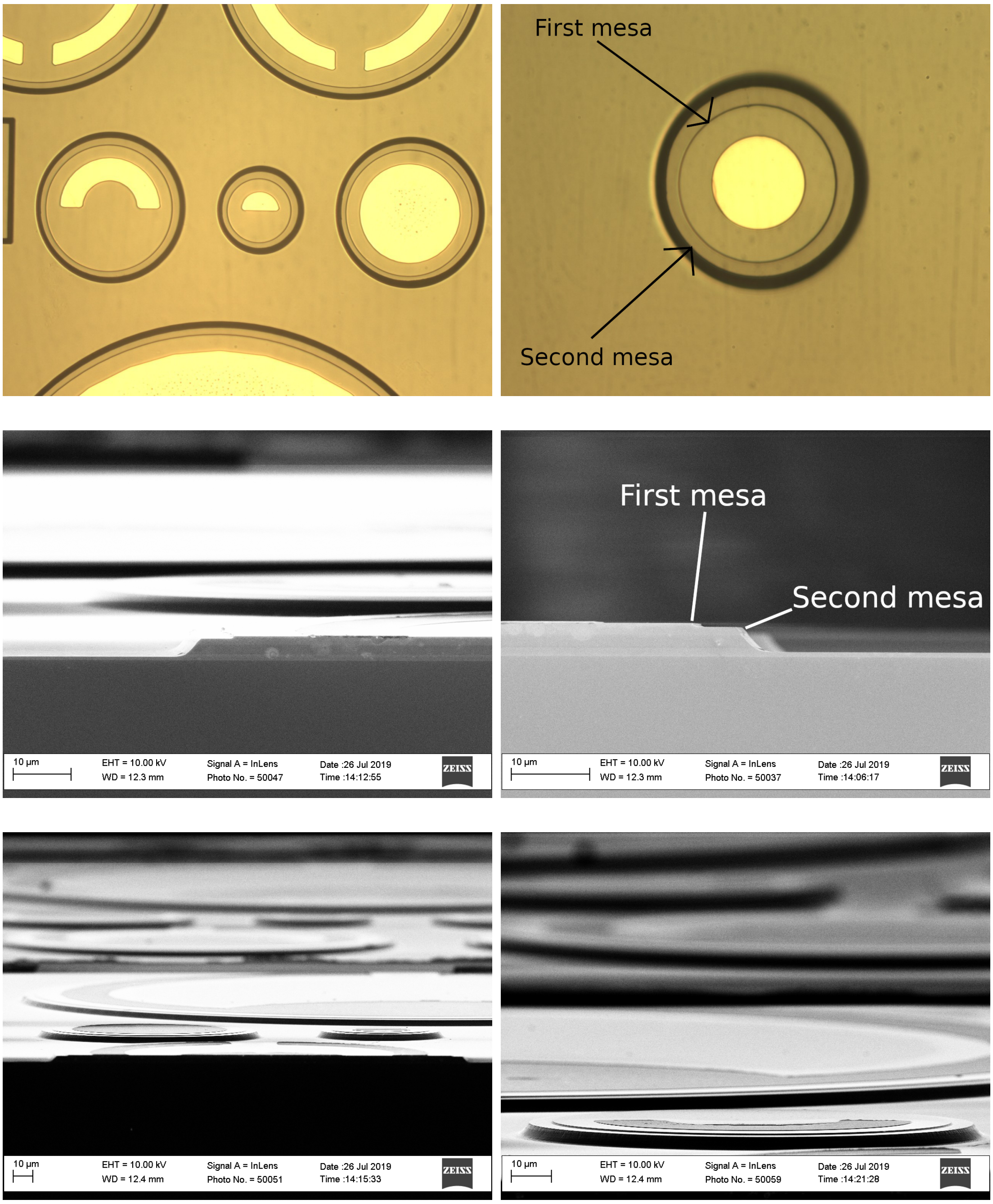


FIGURE 3.18: Microscope (top row) and SEM images of successfully fabricated double mesa diodes, coated with HMDS prior to the second etch step. For reference, the diodes shown in the top row have diameters of 70 and 120 μm, with the 70μm diode labelled.

with negligible surface effects. This can be deduced from the lack of difference between single mesa and double mesa dark currents.

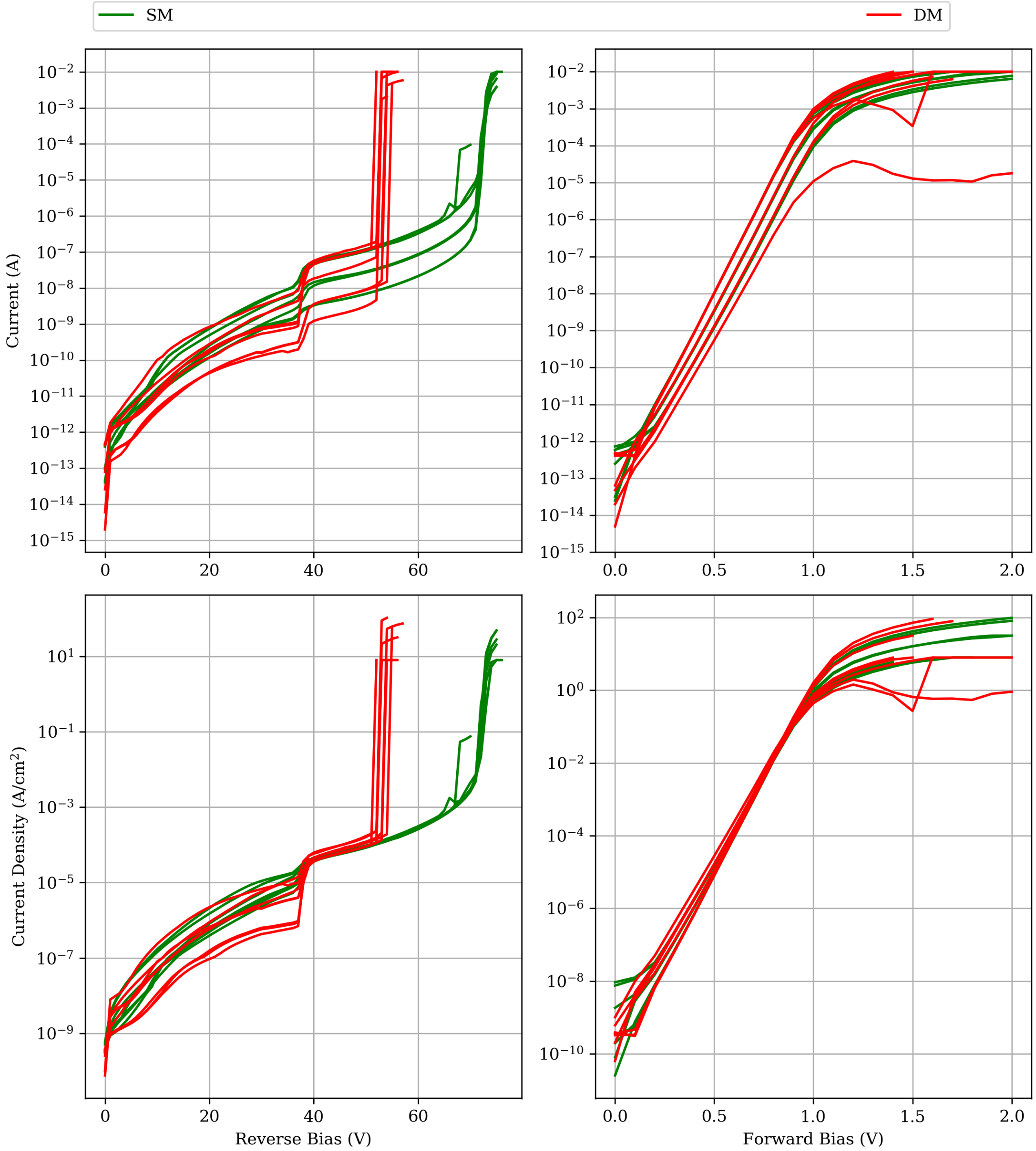


FIGURE 3.19: Comparison of IV data from fabricated single (green) and double (red) mesa diodes.

Chapter 4

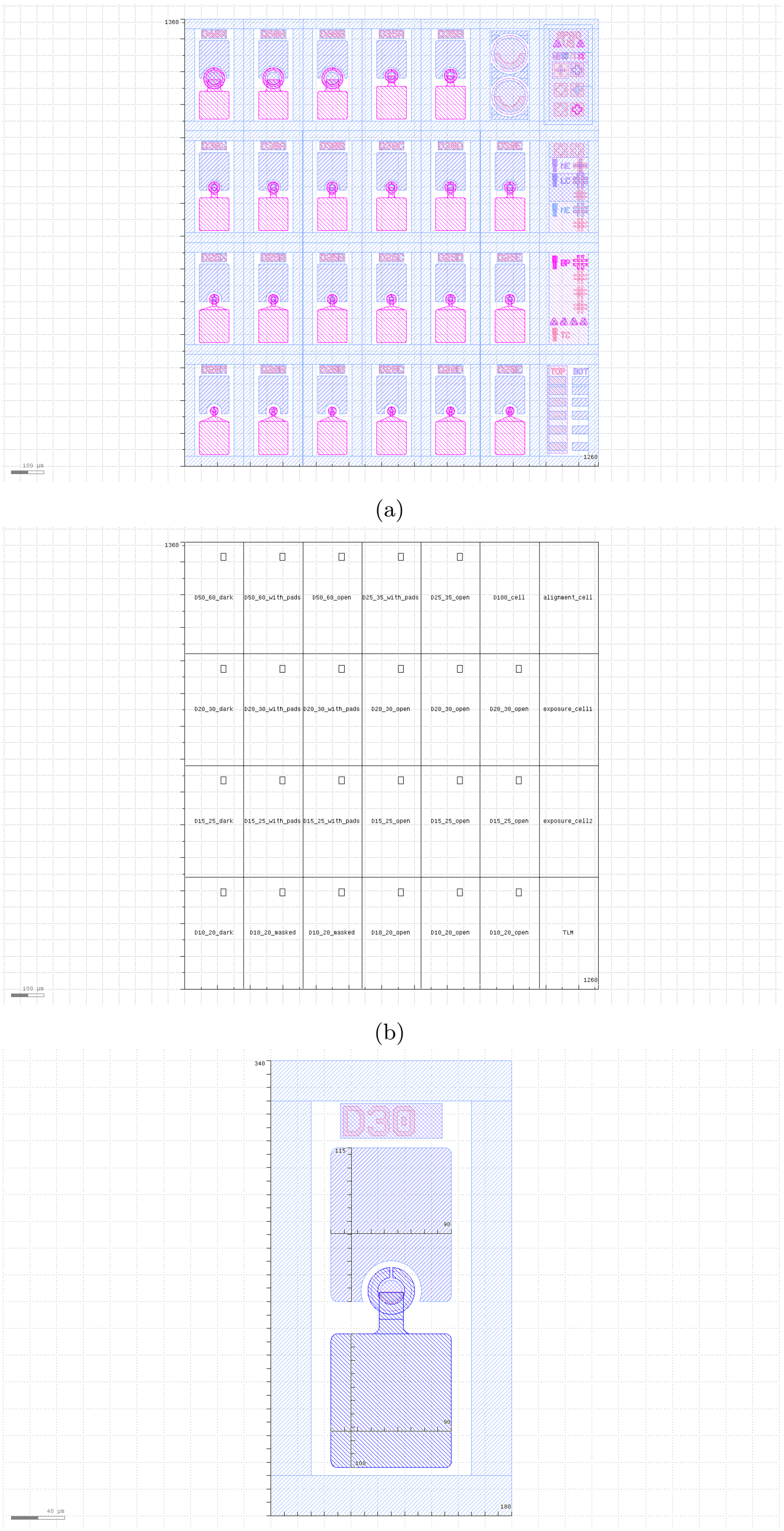
InGaAs/InAlAs Based Single Photon Avalanche Photodiodes

In this chapter, the design, fabrication, and characterisation processes are discussed in detail for two new APD wafers designed for SPAD operation. The wafers are then assessed for low and single photon detection.

4.1 Mask Design

Due to a relatively large dark current density arrising from the InGaAs absorber region, highly sensitive InGaAs/InAlAs SPAD diodes are generally less than 60 μm in diameter to minimise dark count rate (DCR) while maximising single photon detection efficiency (SPDE). A photolithography mask optimised for small diodes was designed for this purpose, as well as to utilise available wafer efficiently.

The designed mask is presented in figure 4.1. The naming convention used for the diodes is D<X> where X is the diameter of the diode in μm. The diodes with bondpads range from diameters of 60 μm (D60) down to 20 μm (D20). They are labelled on the mask, with effective sizes generally 10 μm smaller than indicated after wet etching 4 μm down. In figure 4.1c the lower contact can be seen above the diode, while the top contact bond pad is below. D100 diodes are the largest, but cannot be bonded, so are mainly used for sample diagnostics. There are a total of 23 bondable diodes per 1260×1360 μm



(a)

D50_60_dark	D50_60_with_pads	D50_60_open	D25_35_with_pads	D25_35_open	D100_cell	alignment_cell
D20_30_dark	D20_30_with_pads	D20_30_with_pads	D20_30_open	D20_30_open	D20_30_open	exposure_cell1
D15_25_dark	D15_25_with_pads	D15_25_with_pads	D15_25_open	D15_25_open	D15_25_open	exposure_cell2
D10_20_dark	D10_20_masked	D10_20_masked	D10_20_open	D10_20_open	D10_20_open	TLM

(b)

(c)

FIGURE 4.1: Overview of the designed mask set. Dimensions of the whole mask cell are 1260 × 1360 μm. Some layers are omitted for clarity. (a) shows the overall design, while (b) shows the cell IDs. A close up of a 180 × 340 μm D30 diode cell is shown in (c) with various dimensions indicated.

mask cell, with the cell design ensuring that individual diodes can be cleaved out without damaging neighbours.

Incorporated into the design are various features, such as diodes with sidewall shielding to prevent side photon injection. Due to the small diode size the possibility for failed window lift off means that some diodes are engineered without shielding. In addition, the diodes on the left hand column of the mask are completely covered in metal. These are used as reference diodes during initial testing of experimental set ups to ensure light tightness, since their DCRs can be compared against DCRs from diodes without the metal covering.

An additional mask set called IIG Mesa is used in order to assist with wafer characterisation, shown in figure 4.2. This is a modification of an existing mask set, and contains larger diodes than those in the SPAD mask set. The device diameters are 70, 120, 220, and 440 μm, with only the larger 220 and 440 μm diodes being bondable. More detailed overviews of the IIG SPAD and IIG Mesa mask sets are available in appendix A and appendix B, respectively.

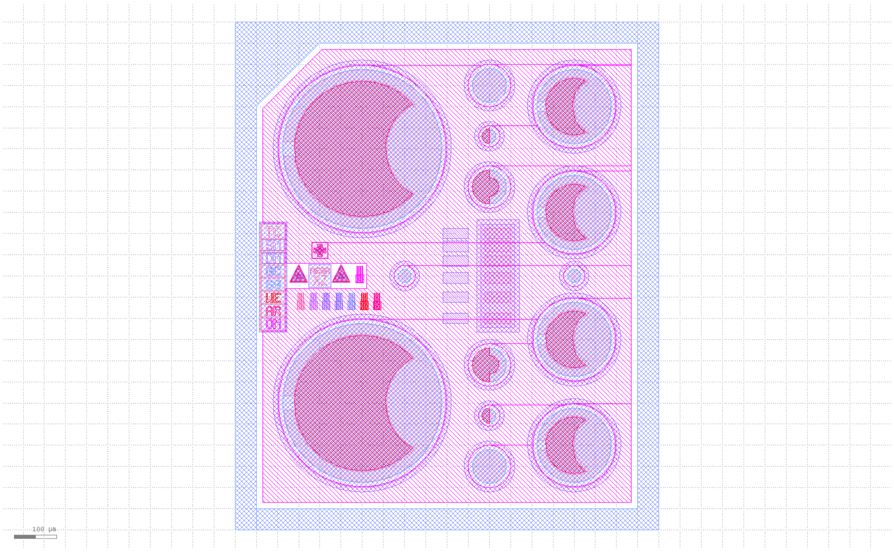


FIGURE 4.2: IIG Mesa mask set.

4.2 Wafer Design and Considerations

Using equations presented in chapter 2, the design of wafers optimised for SPAD operation can be performed. For InGaAs/InAlAs based SAMAPDs, the limiting factor for device operation is generally the high dark current from the InGaAs absorption region. In order to minimise this, it is benefitial to maximise the punch-through voltage as much as possible, so that the electric field is kept low in the absorber. Punch-through must, of course, be kept below the desired operating voltage. In the case of SPAD operation this is generally breakdown voltage and several percent over, where avalanche breakdown probability is high. The lower limit for punch-through voltage is determined by the breakdown field of the absorber. If this is exceded before the breakdown field is reached in the avalanche region, it is no longer a SAMAPD, and SPAD operation will be severely hampered.

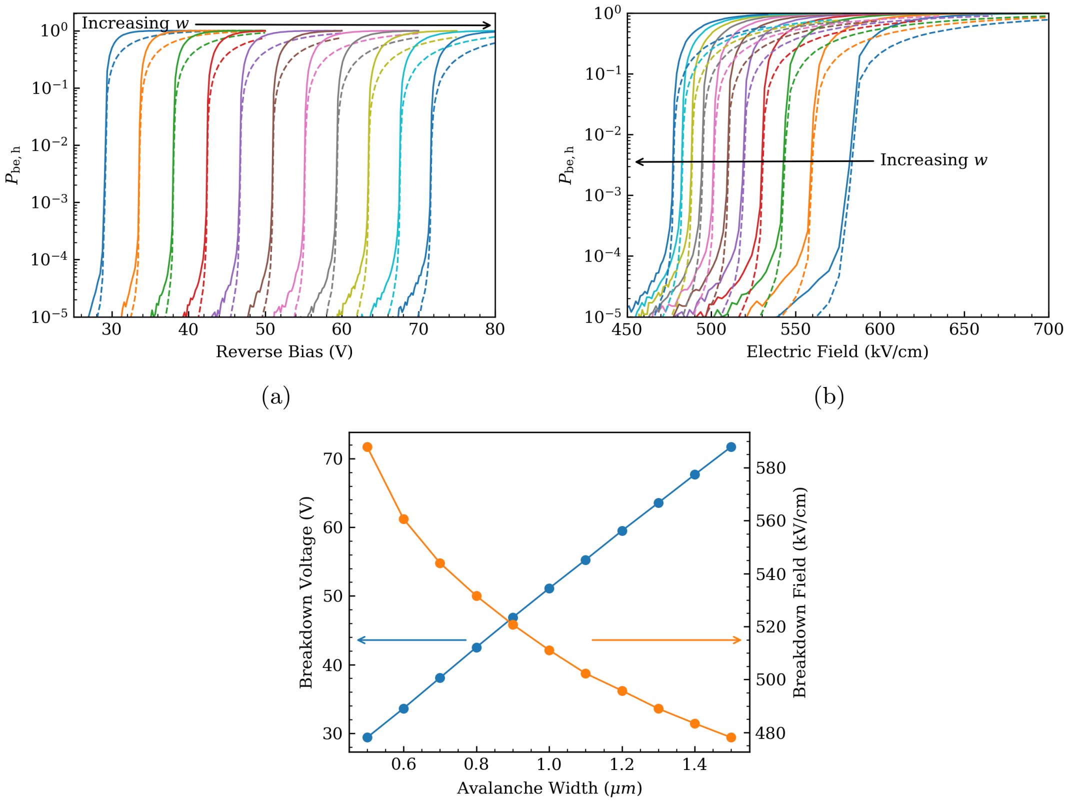
As well as punch-through, the absorption region needs to be wide enough to maximise quantum efficiency. Making this too wide however, leads to a higher dark current since it is a relatively narrow bandgap material, and degraded timing characteristics. The absorption region also needs to be wide enough such that at the desired operating voltage, the field across the region will not cause quantum tunneling.

The avalanche region is engineered to achieve high gain, while also having acceptable timing characteristics. Since SPAD operation is generally performed with circuit quenching, the usual excess noise considerations are not relevent.

The diodes in this work were engineered to punch through as close to breakdown voltage as possible. This minimises the electric field along the InGaAs absorber region with the aim of minimising dark current and any chances of breakdown events, while maintaining a high field in the InAlAs avalanche region. Two diodes with similar characteristics were designed using the recurrence equations (2.16) and (2.17) and electric field modelling using Poisson's equations (see section 2.1). In the following sections, the design considerations are quantitatively explored in detail.

4.2.1 Breakdown Probability Modelling

The breakdown probability for a perfect InAlAs *pin* diode was simulated, as to inform the breakdown behaviour of a SAMAPD structure. Using equations (2.16) and (2.17), P_{be} and P_{bh} were calculated for i region widths of 0.5 μm to 1.5 μm which are shown in figure 4.3. The InAlAs ionisation coefficients used are taken from [22].



(a) (b)

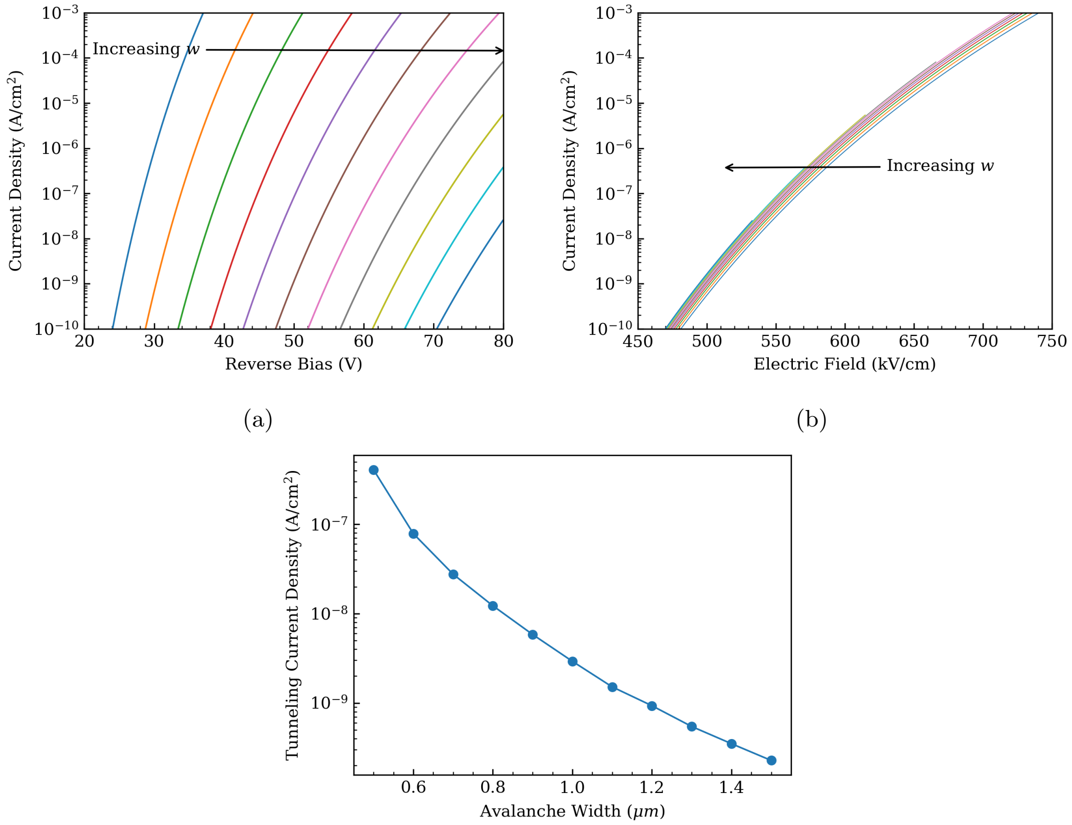
(c)

FIGURE 4.3: (a) and (b): Simulated avalanche breakdown probabilities for perfect InAlAs *pin* structures with widths range from 0.5 μm to 1.5 μm in 0.1 μm steps. Solid lines are P_{be} and dashed lines are P_{bh}. (c) Breakdown voltages and fields extracted from (a) and (b) at $P_{\text{be}} = 0.01$.

As the i region width increases, the reverse bias required to achieve breakdown increases. The electric field for the junction to enter breakdown is however reduced due to carriers acquiring higher kinetic energies before exiting the junction. Depending on the operating voltage requirement, the avalanche region width can be chosen accordingly. However, tunneling current (discussed below) often places upper limits on operating electric fields so there are lower limits for the avalanche region widths.

4.2.2 Tunneling Effects

As mentioned previously, it is essential to consider tunneling current effects when designing the avalanche region width. For this purpose, the InAlAs avalanche region is once more modelled as an InAlAs *pin* diode, for which the tunneling current density can be calculated using equation (2.23).



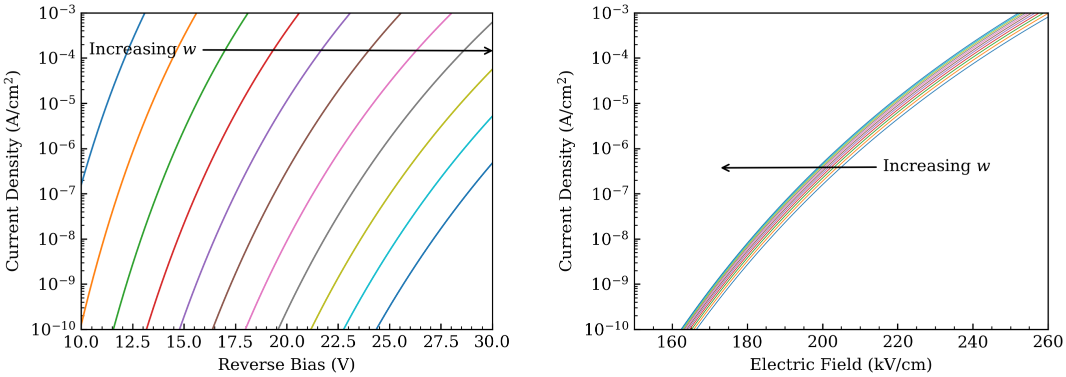
(a)

(b)

(c)

FIGURE 4.4: Modelled tunneling current densities versus (a) reverse bias and (b) reverse electric field for InAlAs *pin* diodes with *i* region widths from 0.5 to 1.5 μm in 0.1 μm steps. (c) Tunneling current density versus avalanche width at $V_{\text{bd}} \times 1.1$ - a typical operating voltage for a SPAD.

Figure 4.4 shows how tunneling current density increases with voltage and electric field as a function of avalanche region width. These were simulated using $E_g = 1.45$ eV, $m^* = 0.069m_0$, and $\sigma_T = 1.26$ [76]. The tunneling current as a function of electric field is not significantly affected by the *i* region width, though tunneling current depends sensitively on overbias. As avalanche width decreases, the tunneling current at avalanche breakdown increases, since the required avalanche breakdown field itself increases.



(a) (b)

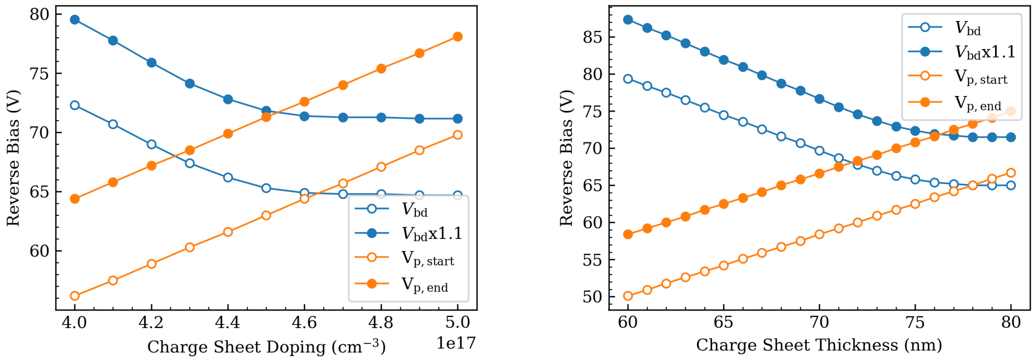
FIGURE 4.5: Modelled tunneling current densities versus (a) reverse bias and (b) reverse electric field for InGaAs *pin* diodes with *i* region widths ranging from 0.5 μm to 1.5 μm in 0.1 μm steps.

The absorption region's tunneling current is also considered. The tunneling current densities are shown in figure 4.5 for InGaAs pin diodes. Parameters used are $E_g = 0.738$ eV, $m^* = 0.068m_0$, and $\sigma_T = 1.16$ [77]. Since InGaAs has a narrower bandgap, tunneling current is more significant. Tunneling becomes significant at far lower fields than in InAlAs, so SPAD design must minimise this field.

4.2.3 Electric Field Modelling

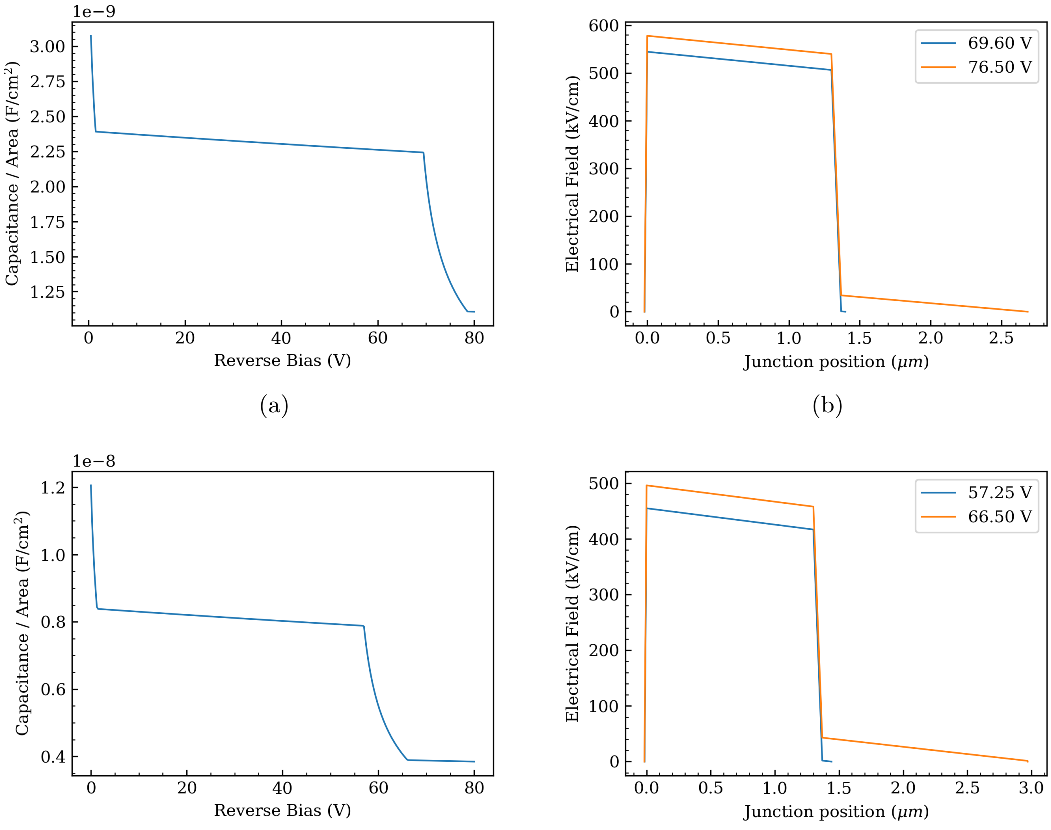
Since both breakdown probability and tunneling current depend on electric field, SPAD design requires electric field profiles as a function of reverse bias. For SPAD designs with a SAM structure, the charge sheet doping N_c and charge sheet thickness w_c heavily influence the electric field profiles. Utilising the modelled breakdown fields for *pin* diodes (shown in figure 4.3c), and mapping the electric field of the *pin* diodes to the SAM structure, it is possible predict the dependence of avalanche breakdown and punchthrough behaviour across the SAM region due to N_c and w_c. These are illustrated in figure 4.6. V_{bd} and V_p vary linearly with N_c and w_c, until V_{bd} drops below V_p. At this point the entire applied reverse voltage is dropped across the avalanche region, so V_{bd} remains constant with N_c or w_c, with the diode behaving more like a simple *pin* diode.

Details of the two wafer structures designed, SF1318 and SF1319, are shown in table 4.1. Their predicted CV and electrical field characteristics are shown in figure 4.7, with full punchthrough achieved at 76.50 and 66.50 V for SF1318 and SF1319 respectively. Graphs combining the electric field values extracted from CV model with the breakdown



(a) (b)

FIGURE 4.6: (a) demonstrates the effects of varying N_c with a fixed $w_c = 69$ nm. (b) demonstrates the effects of varying w_c with $N_c = 4.1e17$ cm^{-3}. V_p and V_{bd} are extracted from the CV model. $V_{\text{p,start}}$ and $V_{\text{p,end}}$ represent where the absorption region starts to deplete and fully depletes, respectively.



(a) (b)

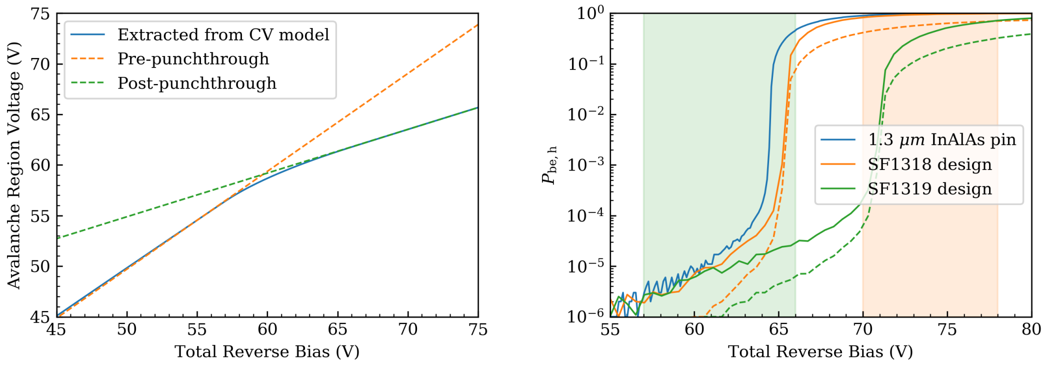
(c) (d)

FIGURE 4.7: Modelled CV of the (a) SF1318 and (b) SF1319 structures as designed. The two voltages indicated in (b) and (d) are the voltages at which the InGaAs absorber begins to deplete and fully depletes, respectively.

probability modelling are shown in figure 4.8. Figure 4.8a demonstrates how the effective voltage across the avalanche region changes with reverse bias as the structure begins to

Doping	Material	Doping density (cm^{-3})	Thickness (nm)
n+	$In_{0.53}Ga_{0.47}As$	5.0E19	100
n+	$In_{0.52}Al_{0.48}As$	2.0E18	300
p-	$In_{0.52}Al_{0.48}As$	2.0E15	1300
p+	$In_{0.52}Al_{0.48}As$	4.1E17 (SF1319) 5.0E17 (SF1318)	69
p-	$In_{0.52}Al_{0.48}As$	2.0E15	25
p-	$Al_{0.15}Ga_{0.32}In_{0.53}As$	2.0E15	25
p-	$Al_{0.29}Ga_{0.18}In_{0.53}As$	2.0E15	25
p-	$In_{0.53}Ga_{0.47}As$	2.0E15	1500
p-	$Al_{0.15}Ga_{0.32}In_{0.53}As$	2.0E15	25
p-	$Al_{0.29}Ga_{0.18}In_{0.53}As$	2.0E15	25
p-	$In_{0.52}Al_{0.48}As$	2.0E15	100
p+	$In_{0.52}Al_{0.48}As$	2.0E18	100
p+	$In_{0.53}Ga_{0.47}As$	5.0E18	300
-	InP Substrate	-	-

TABLE 4.1: SF1318 and SF1319 wafer structures as designed (charge sheet doping is later confirmed through CV measurements, and is different than the design).



(a) (b)

FIGURE 4.8: Extracting avalanche field from the predicted CV. Total reverse bias across the SAMAPD is shown on the horizontal axes. (a) shows avalanche region voltage extracted from CV modelling for SF1319, and how as the capacitance begins to drop, the electric field begins to increase more slowly in the avalanche region. Gradients pre and post-punchthrough are indicated to emphasise the voltage difference. In (b) the effective voltages are mapped to the breakdown probabilities calculated for a $w = 1.3\ \mu m$ pin diode. Orange and green highlighted areas show the voltage range across which the SF1318 and SF1319 absorption regions are depleting, respectively.

punch-through. The breakdown of SF1318 is engineered such that the absorption region starts to deplete as the avalanche region achieves 10% over breakdown field. SF1319 is designed to fully punch-through before breakdown occurs, with only a small seperation between achieving full depletion and high P_b.

The effects of temperature on electric field profiles is also considered. As the diode is cooled, the field required for breakdown to be achieved reduces. At the same time, punchthrough voltage will also move slightly higher due to an increasing bandgap. With this in mind, the SF1318 structure is designed for room temperature operation, while SF1319 is more optimised for cooling.

4.3 Wafer Growth and Device Fabrication

MBE wafer growth for the InGaAs/InAlAs SPADs (SF1318 and SF1319) was performed by Dr Shiyong Zhang at the EPSRC National Centre for III-V Technologies at the University of Sheffield. Table 4.1 shows the requested wafer structures. The structures are designed to produce *nipip* diodes, with additional AlGaInAs grading layers between the InGaAs/InAlAs layers added to avoid trapping carriers at the interface band-gap discontinuities. The i layers were lightly p doped to ensure consistency with the surrounding p doped layers, and prevent any undesirable pn junctions within the structure.

Several device fabrication rounds were carried out using the mask sets discussed in section 4.1, with one mask set containing D70, D120, D220, and D440 diodes, and the second containing the smaller D20, D25, D30, D60, and D100 diodes. Fabrication was carried out using the standard procedure discussed in detail in section 3.1, with a brief description given here. Samples were cleaved from the main wafer and a 3-stage clean performed. PMGI was spun onto the samples, followed by a 5 minute bake on a hotplate at 180°C. SPR350 was then spun on top of the PMGI layer, and baked for 1 minute at 100°C. All spin coating was performed at 4000 RPM. Top contact layers were then exposed using a UV mask aligner, and the photoresists developed for 1 minute in MF26A developer. Prior to metal evaporation, the samples were exposed to a O_2 plasma in an asher to remove any remaining dirt from the contact area. The Ti/Au (20/200 nm) top contacts were then deposited using a metal evaporation chamber. Lift off was performed first using acetone to remove the SPR350 and metal covering, followed by the sample being placed once more in MF26A developer to remove the PMGI layer, and then a 3-stage clean.

Following contact deposition the mesa etch layer was exposed onto a spin coating of HMDS + SPR220, with the HMDS added to ensure optimal photoresist adhesion during etching. Wet chemical etching was done using a solution of H_2SO_4:H_2O_2:DIW

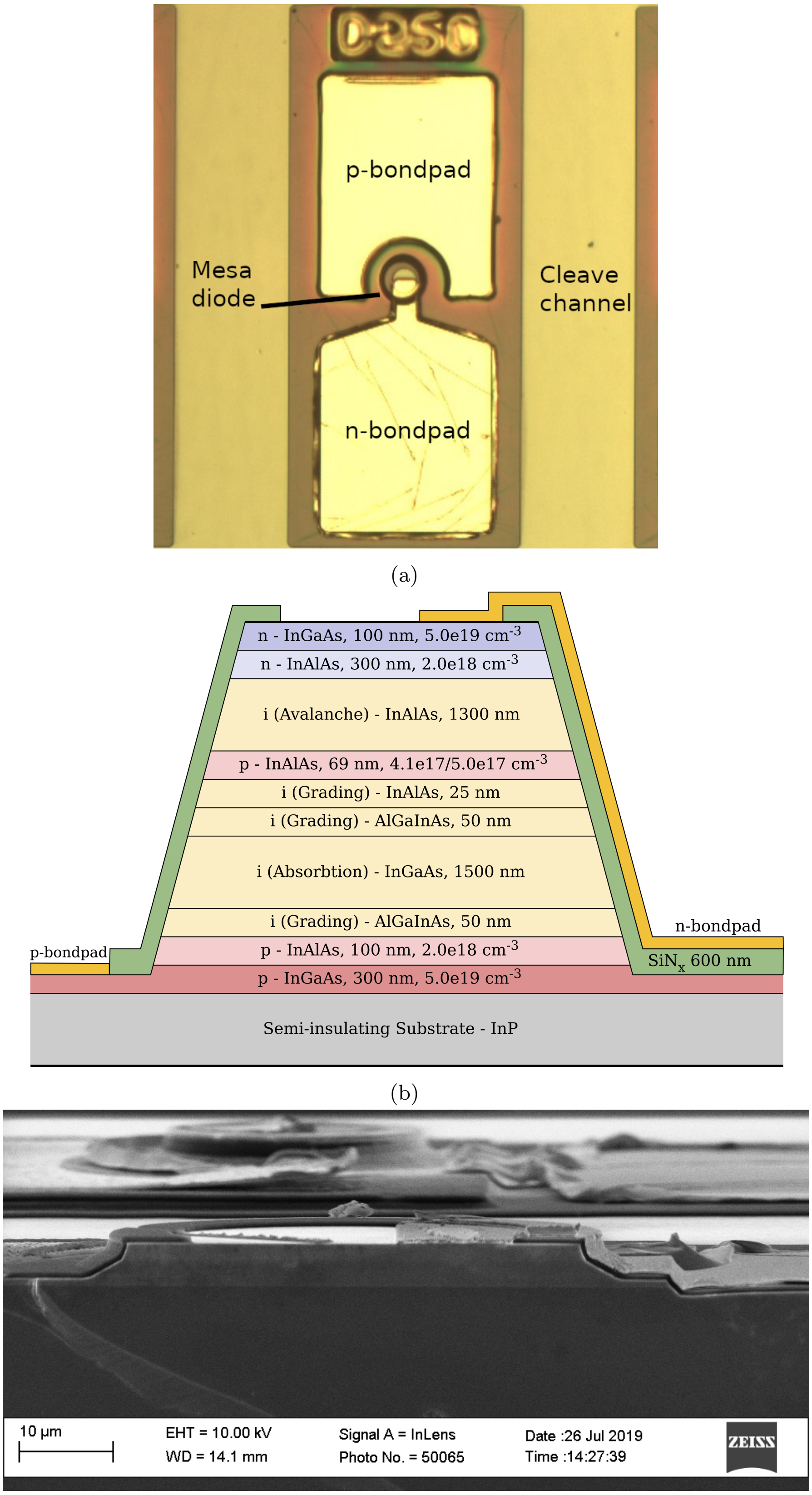


FIGURE 4.9: (a) shows an image of a D25 diode from the IIG SPAD mask after the final fabrication step, (b) is a cross section diagram of the diodes, and (c) is an SEM image of a D60 diode cleaved as in the cross section diagram.

(1:8:80), etching approximately 4 μm down to the lower InGaAs p-contact. Etch depth was verified using a Dektak surface profiler.

After stripping the etch photoresist in acetone, SPR220 was spin coated onto the sample and the lower contact/bondpad layer exposed and a Ti/Au (20/200 nm) contact deposited. After lift off the IIG Mesa fabrication ends, whereas the IIG SPAD fabrication procedure continues. IIG SPAD samples proceded to have a 600 nm $\mathrm{SiN_x}$ layer deposited using PECVD at 100°C. This layer passivates the sidewall of the mesa diodes and provides isolation for the bondpads from the mesa edge.

HMDS and SPR220 layers were spin-coated onto the nitride, followed by photolithography to expose part of the nitride layer. Using RIE with a $\mathrm{CHF_4/O_2}$ (35/5 sccm) gas mixture at 100 W of RF power, the nitride was etched and completely removed in contact areas and the mesa window. The final stage was the top bondpad deposition, with the mask layer exposed onto a HMDS+SPR220 spin coating. Prior to evaporation, the bondpad area was roughened using RIE with the previous recipe for 15 seconds to aid with bondpad adhesion. Ti/Au (40/400 nm) bond-pads were then deposited to connect to the top of the mesas. Examples of finished diodes are shown in figure 4.9.

4.4 Wafer Characterisation

4.4.1 IV Measurements

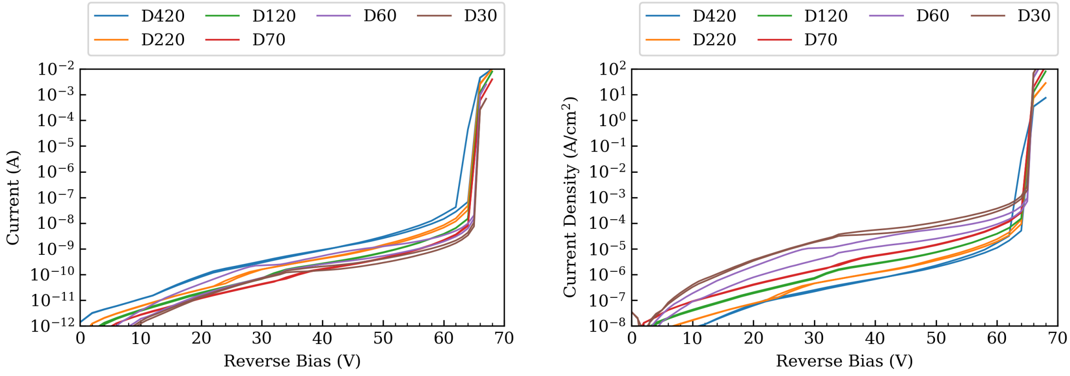
Room temperature IV measurements were carried out on diodes with 20 to 420 μm diameters. Diodes were characterised before nitride deposition and after the final bondpad deposition stage in order to determine the effects of the nitride and bondpad.

Room temperature IV data of SF1318 are shown in figure 4.10. These show a good consistency in breakdown voltage with most diodes breaking down sharply at the same voltage. There is no indication of punch-through, which would be indicated by a sudden rise in dark current before breakdown.

Room temperature IV data of SF1319 are shown in figure 4.11. The SF1319 diodes show less consistent data compared to SF1318. Nevertheless, breakdown voltage is largely consistent. There are more noticable increases in dark current with reverse bias than in

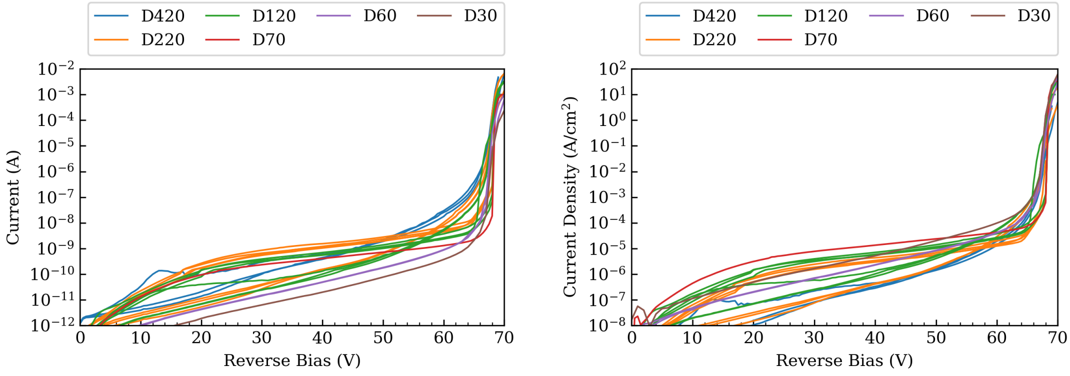
SF1318 which could be indicative of surface roughness due to wet etching issues. This effect is most pronounced in the larger diodes, with D70 diodes and smaller (those of interest for SPAD operation) largely avoiding this issue. As with SF1318, there is no clear indication of punch-through from the dark IV data.

Comparisons of diodes pre-nitride and post-bondpad deposition, are shown in figure 4.12. After bondpad deposition, the SF1318 IV data improved (decreased dark current) post-bondpad deposition, whereas SF1319 diodes were degraded. This indicates that in the SF1318 diodes the nitride suppressed negative surface effects, for example from chemical residue. SF1319 diodes were degraded from the nitride and bondpad deposition, though their IV characteristics largely similar to that of the SF1318 diodes post-bondpad deposition. It seems that the effects of the nitride largely normalise diode performance.



(a) (b)

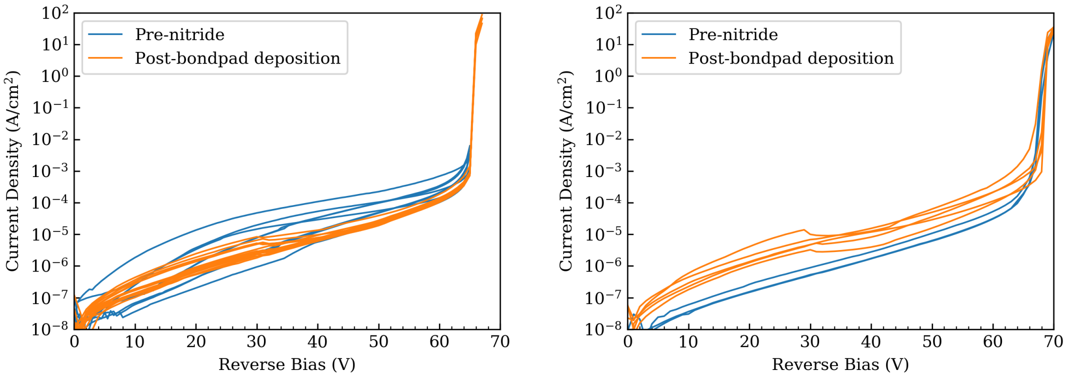
FIGURE 4.10: IV measurements of SF1318 diodes before nitride deposition.



(a) (b)

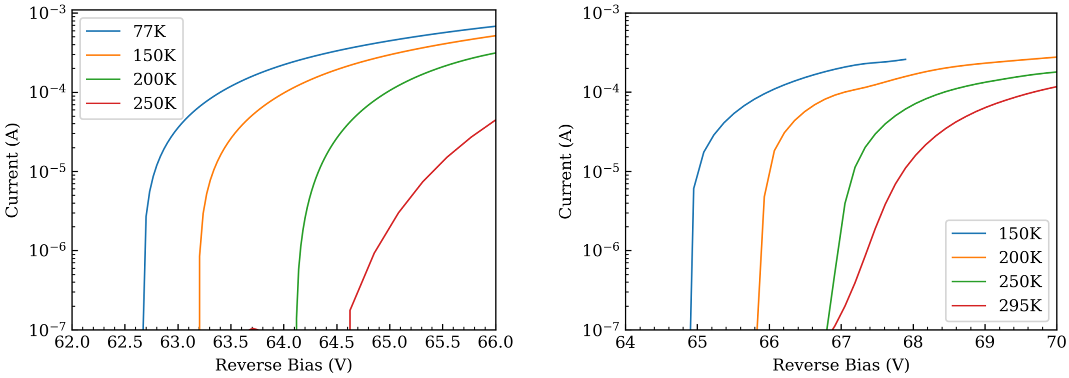
FIGURE 4.11: IV measurements of SF1319 diodes before nitride deposition.

The temperature dependence of the diodes' IV characteristics was obtained using a low-temperature probe station (Janis ST-500). Temperature dependent IV data of an



(a) (b)

FIGURE 4.12: Current density measurements of (a) SF1318 and (b) SF1319 diodes before (blue) and after (orange) nitride.



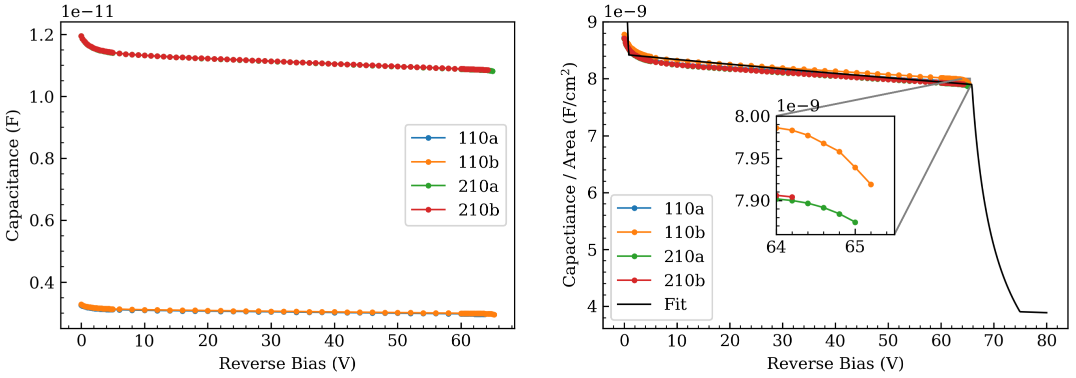
(a) (b)

FIGURE 4.13: (a) SF1318 and (b) SF1319 temperature dependent IV measurements to determine breakdown voltage temperature dependence.

SF1318 diode and SF1319 diode are shown in figures 4.13a and 4.13b, respectively. Values for the temperature dependences between 150 and 295 K are found to be 15.9 mV/K and 16.6 mV/K for SF1318 and SF1319, respectively.

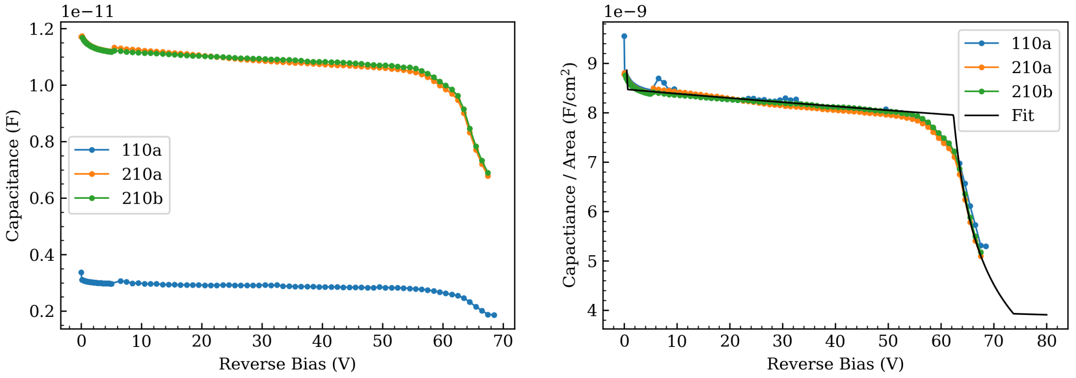
4.4.2 CV Measurements

Figures 4.14 and 4.15 show the capacitance data SF1318 and SF1319 diodes respectively. Large 420 μm and 220 μm diameter diodes are used for these measurements in order to maximise accuracy in determining capacitance per unit area. Measuring diodes with different sizes allows the actual device area to be extracted, since some area is lost to etching. For both SF1318 and SF1319 the effective reduction in diameter is found to be 8 μm, which is reasonable for the etch depth of 4 μm in device fabrication.



(a) (b)

Figure 4.14: (a) CV data of SF1318 diodes with nominal diameters of 420 and 220 μm. There is an indication of the start of punchthrough, shown in the inset. The capacitance density data in (b) has a correction of -8 μm applied to the diameter.



(a) (b)

Figure 4.15: (a) CV data of SF1319 diodes with diameters of 420 and 220 μm. A correction of -8 μm is applied to the diameter for the area calculation in (b).

From these data it is possible to deduce the effective dopings and layer thicknesses in the wafers. Figures 4.14b and 4.15b overlay fittings to the measured capacitance density versus reverse bias. The fittings yielded effective charge sheet dopings of 5.15×10^{17} and 4.45×10^{17} cm^{-3} for SF1318 and SF1319 respectively, using a 69 nm charge sheet width.

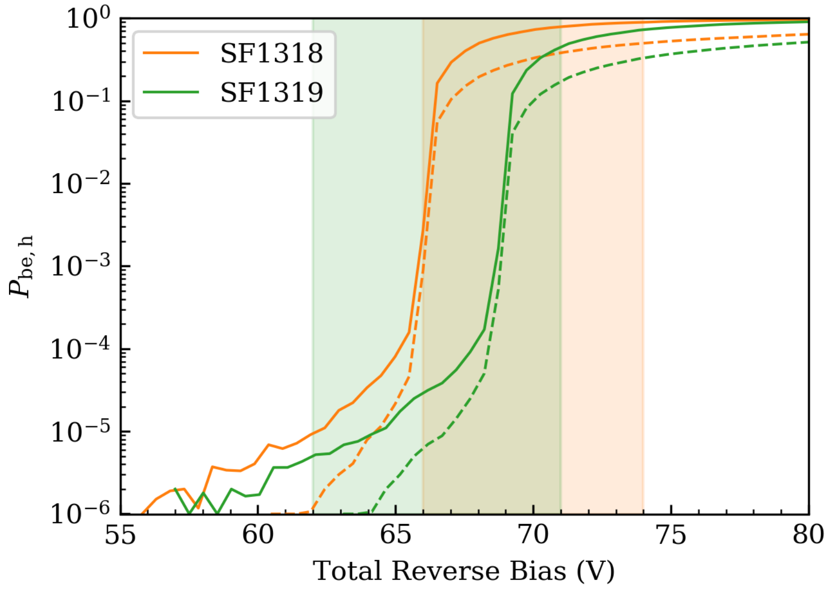


FIGURE 4.16: Punch-through and breakdown probabilities predicted for SF1318 and SF1319 based on IV and CV measurements. Orange and green highlighted areas indicate the voltages across which the absorption regions of SF1318 and SF1319 are depleting, with an overlapping area in the centre.

Doping	Material	Doping density (cm^{-3})	Thickness (nm)
n+	$In_{0.53}Ga_{0.47}As$	5.0E19	100
n+	$In_{0.52}Al_{0.48}As$	2.0E18	300
p-	$In_{0.52}Al_{0.48}As$	2.0E15	1300
p+	$In_{0.52}Al_{0.48}As$	4.45E17 (SF1319) 5.15E17 (SF1318)	69
p-	$In_{0.52}Al_{0.48}As$	2.0E15	25
p-	$Al_{0.15}Ga_{0.32}In_{0.53}As$	2.0E15	25
p-	$Al_{0.29}Ga_{0.18}In_{0.53}As$	2.0E15	25
p-	$In_{0.53}Ga_{0.47}As$	2.0E15	1500
p-	$Al_{0.15}Ga_{0.32}In_{0.53}As$	2.0E15	25
p-	$Al_{0.29}Ga_{0.18}In_{0.53}As$	2.0E15	25
p-	$In_{0.52}Al_{0.48}As$	2.0E15	100
p+	$In_{0.52}Al_{0.48}As$	2.0E18	100
p+	$In_{0.53}Ga_{0.47}As$	5.0E18	300
-	InP Substrate	-	-

TABLE 4.2: SF1318 and SF1319 wafer structures based on CV data fitting.

These are slightly higher than the charge sheet dopings in the designs in table 4.1. The updated structure based on CV data fitting is shown in table 4.2.

Based on the SF1318 data, the diodes have not punched-through by the point of avalanche breakdown. There is an indication at the highest voltage points that structure may be beginning to punch-through, however practical measurement range is limited by degrading phase angle when the diode is close to breakdown. The fit for SF1318 assumes that the start of punch-through is indeed around breakdown.

The absorption region of SF1319 does begin to deplete before breakdown voltage, however it does not fully deplete which indicates that the charge sheet doping is higher than expected, as shown by the fit. As with the SF1318 measurement the range is limited by the phase angle diverging significantly from -90^{o}.

It is important to note that the CV fittings are only an approximation. They assume a uniform doping across each region which is not necessarily the case. The actual distribution of dopants is likely to be more diffused the different regions.

Based on the IV and CV data of SF1318 and SF1319, it is possible to predict the behaviour of P_{be} and P_{bh} in each of the grown structures. Figure 4.16 shows how each structure is expected to deplete and the dependence of breakdown probability on reverse bias.

4.4.3 Responsivity

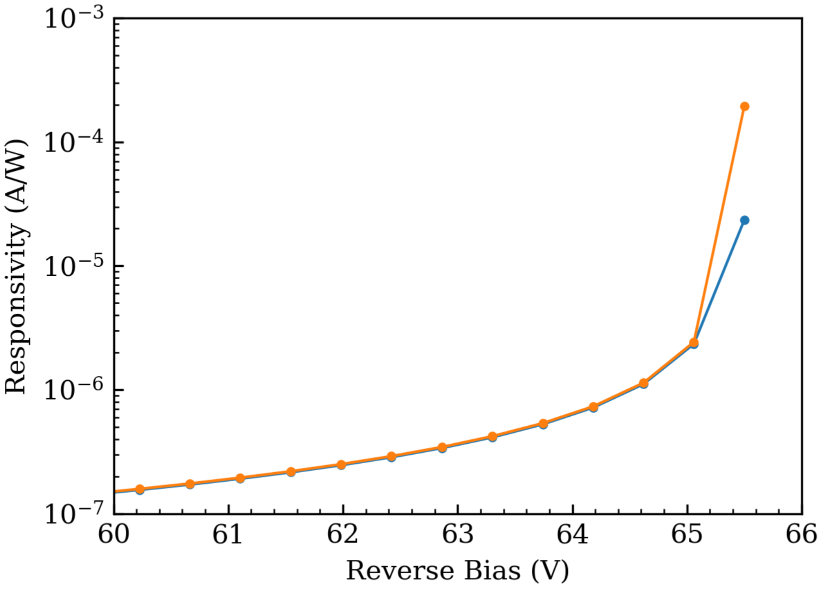
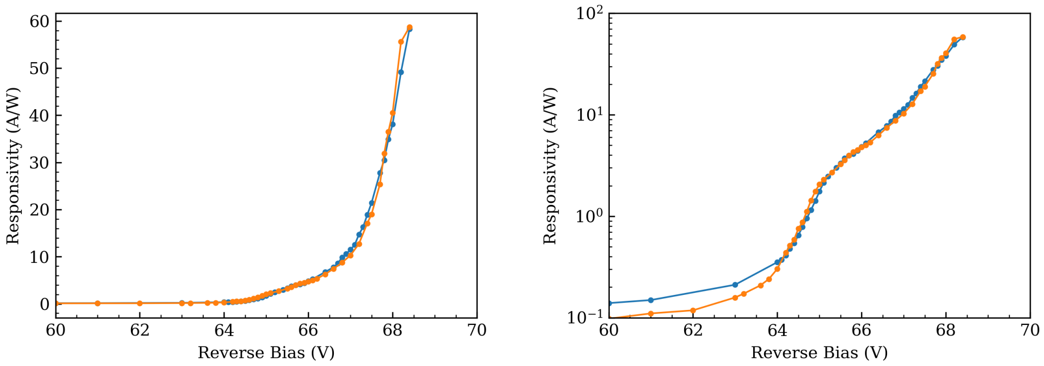


FIGURE 4.17: SF1318 responsivity data for two D220 diodes, including the effects of avalanche gain.

Responsivity was measured using large diodes to ensure all of the laser spot is incident on the mesa top. The responsivity data for two D220 SF1318 diodes is shown in figure



(a) (b)

FIGURE 4.18: Responsivity data for SF1319 obtained from two D220 diodes. (a) shows responsivity on a linear scale while (b) shows it on a log scale to provide more clarity in the low gain region. The responsivity shown includes gain, since it is difficult to define a region of gain 1.

4.17. As expected from CV data, there is very little photoresponse for reverse biases below 65 V, i.e. before the InGaAs absorber is depleted. Measurement of responsivity beyond 66 V is not possible as the diodes begin to breakdown.

Similar responsivity measurements were carried out for SF1319 and the data from D220 diodes are plotted in Figure 4.18. The responsivity shown does include avalanche gain, as it is difficult to extract gain data from a diode that is still in the process of punching-through. It is likely that there is already gain by the time that the diode photoresponses reach 1.5 A/W, which is approximately the expected photoresponse for when the InGaAs layer is fully depleted which further complicates determining gain. It is clear that the diodes are able to achieve relatively high gain before breakdown. However, the drop in capacitance in figure 4.15 is consistent with the sharp increase in photocurrent in the responsivity data at 65 V.

4.5 DCR and Pulsed Photon Measurements

Pulsed DCR measurements were carried out. Unless otherwise stated, the experiments were carried out at a repetition frequency of 100 kHz, with the pulse width set to 10 ns on the pulse generator. Pulse heights were varied in measurements (often adjusted based on the transients observed), so the data below are presented with total reverse bias or overbias, which combines the AC and DC biases. The range of pulse heights was between 5 and 10 V, with the DC bias making up the rest of the total applied reverse bias.

4.5.1 DCR Measurements

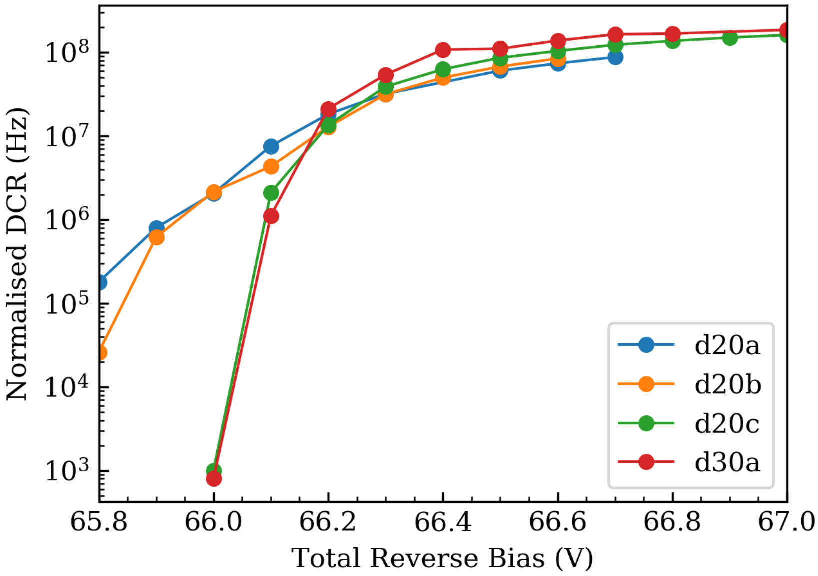


Figure 4.19: Pulsed DCR for SF1318 diodes.

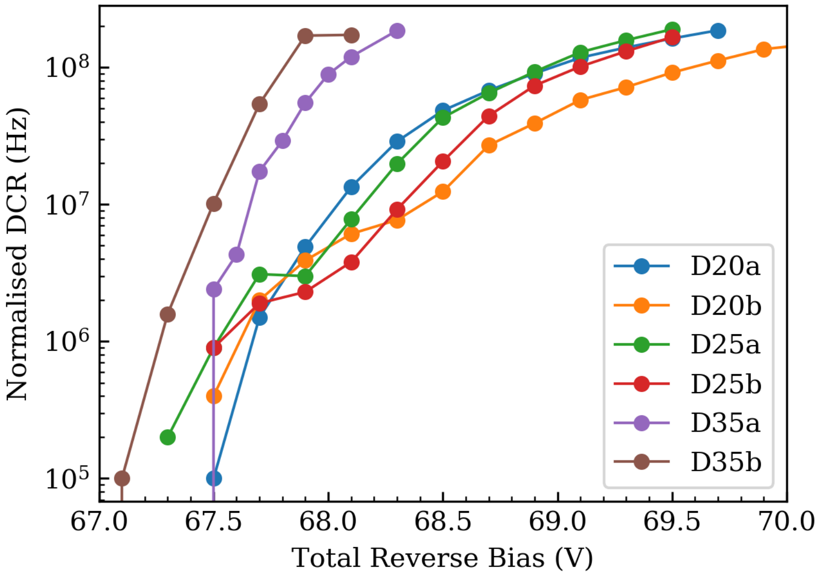


Figure 4.20: Pulsed DCR for SF1319 diodes.

Pulsed DCR measurements were performed on SF1318 using the CQC board described in section 3.5.4.1. The results of this are presented in figure 4.19. SF1318 diodes showed good, clear breakdown events well above the transients.

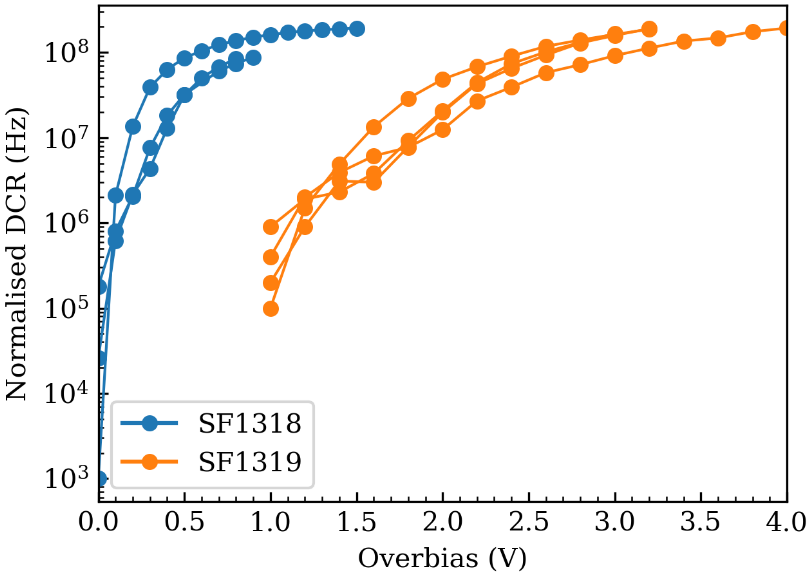


FIGURE 4.21: Comparison of DCR for SF1318 and SF1319, with D20 and D25 diodes plotted together.

Due to smaller observed breakdown events, SF1319 DCR data was gathered using the LeCroy oscilloscope method described in section 3.5.4.2. This limits the normalised DCR noise floor of the measurement to around 10^4 Hz, so the lowest overbiases are not measurable. Results are shown in figure 4.20. A comparison is shown in figure 4.21.

4.5.2 Afterpulsing

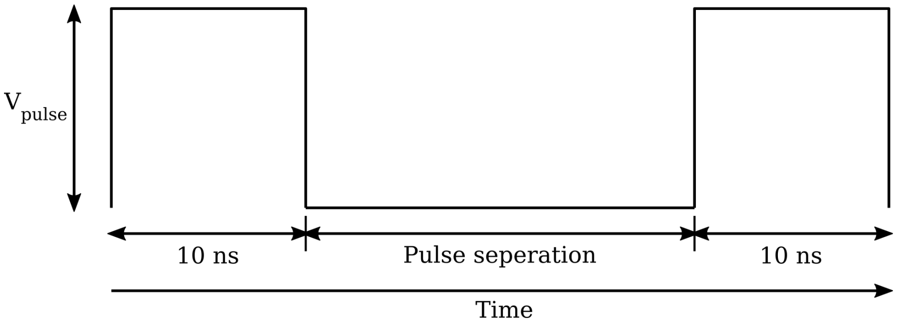


FIGURE 4.22: Double biasing pulse setup used for assessing afterpulsing.

To investigate afterpulsing, a double biasing pulse was used. The pulse is depicted in figure 4.22, and consists of two pulses with identical height and width, but varying pulse seperation. As seen in figure 4.23, the effects of afterpulsing are negligible after a seperation of 50 ns, which is equivalent to a repetition frequency of 20 MHz. This

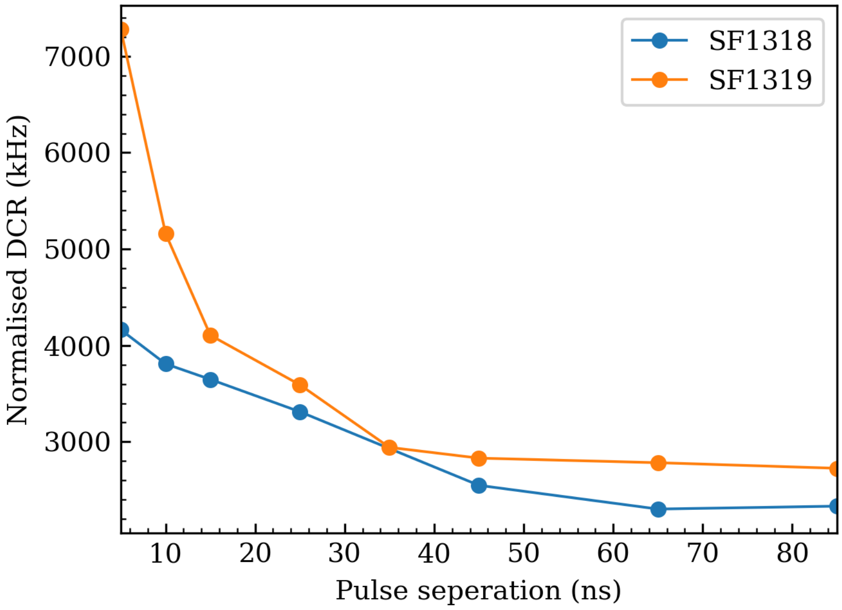


Figure 4.23: Double pulsing DCR measurements for SF1318 and SF1319. Two 10 ns, 4 V pulses were applied with the seperation between the end of the first pulse and the start of the second pulse indicated on the x-axis. The DC reverse bias was set to 62.2 V and 64.0 V for SF1318 and SF1319 respectively.

is essentially the upper limit on the repetition frequency for these diodes. With all subsequent measurements performed at repetition frequencies below 1 MHz, afterpulsing is not expected to affect the results significantly.

4.5.3 Temperature Dependence

The temperature dependence of DCR in SF1318 and SF1319 diodes was measured using a low-temperature probe station. The SF1318 diodes were placed in the low-temperature probe station and cooled down to 77 K from 295 K, with DCR measurements performed between. DCR results for a D20 diode are presented in figure 4.24. The same procedure was repeated for SF1319 down to 150 K, with results shown in figure 4.25.

In both cases, the temperature dependence of DCR is not as expected. SF1318 does show an improvement in DCR, though not as significant as in previous InGaAs/InAlAs SPADs [48]. SF1319 shows no improvement in DCR with cooling when considered against raw applied bias, however this must be considered in context of the breakdown voltage falling within the middle of the punch-through voltage range. This is discussed further in section 4.6.

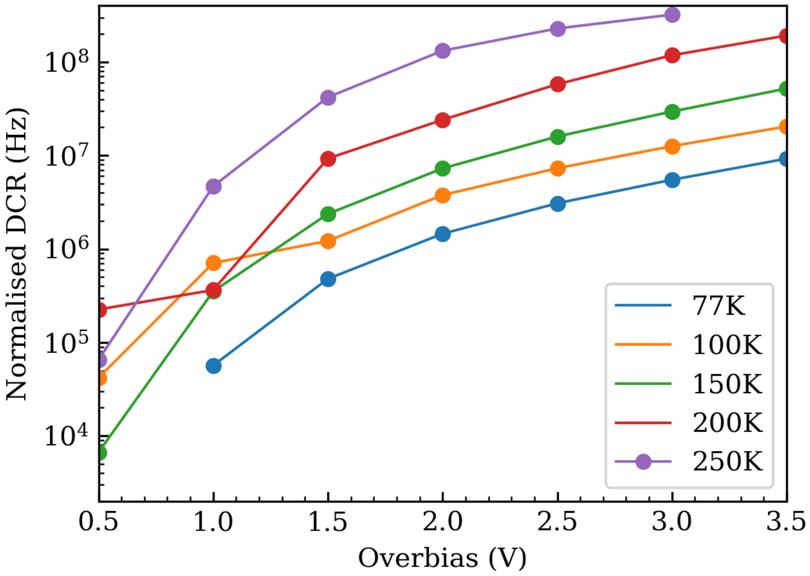


FIGURE 4.24: Temperature dependence of DCR in a D20 SF1318 diode. The measurements were performed with a 1.5 ns wide 5 V pulse at a repetition frequency of 100 kHz.

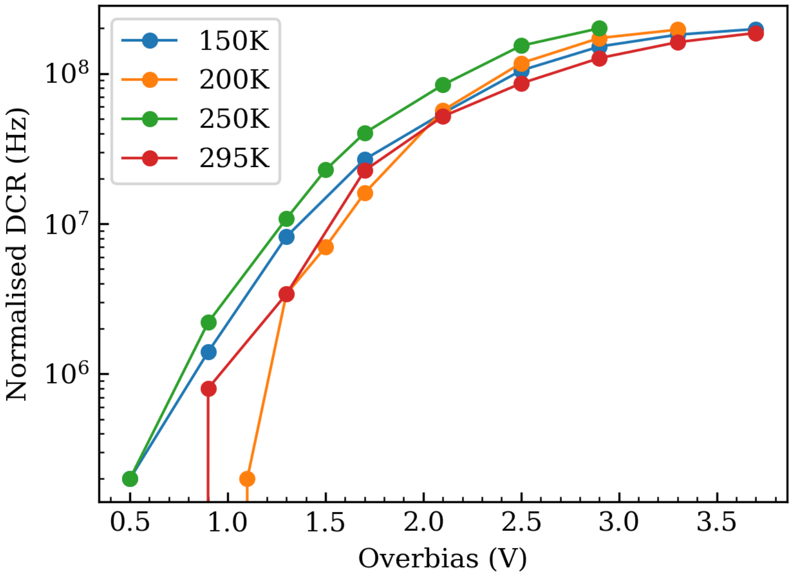


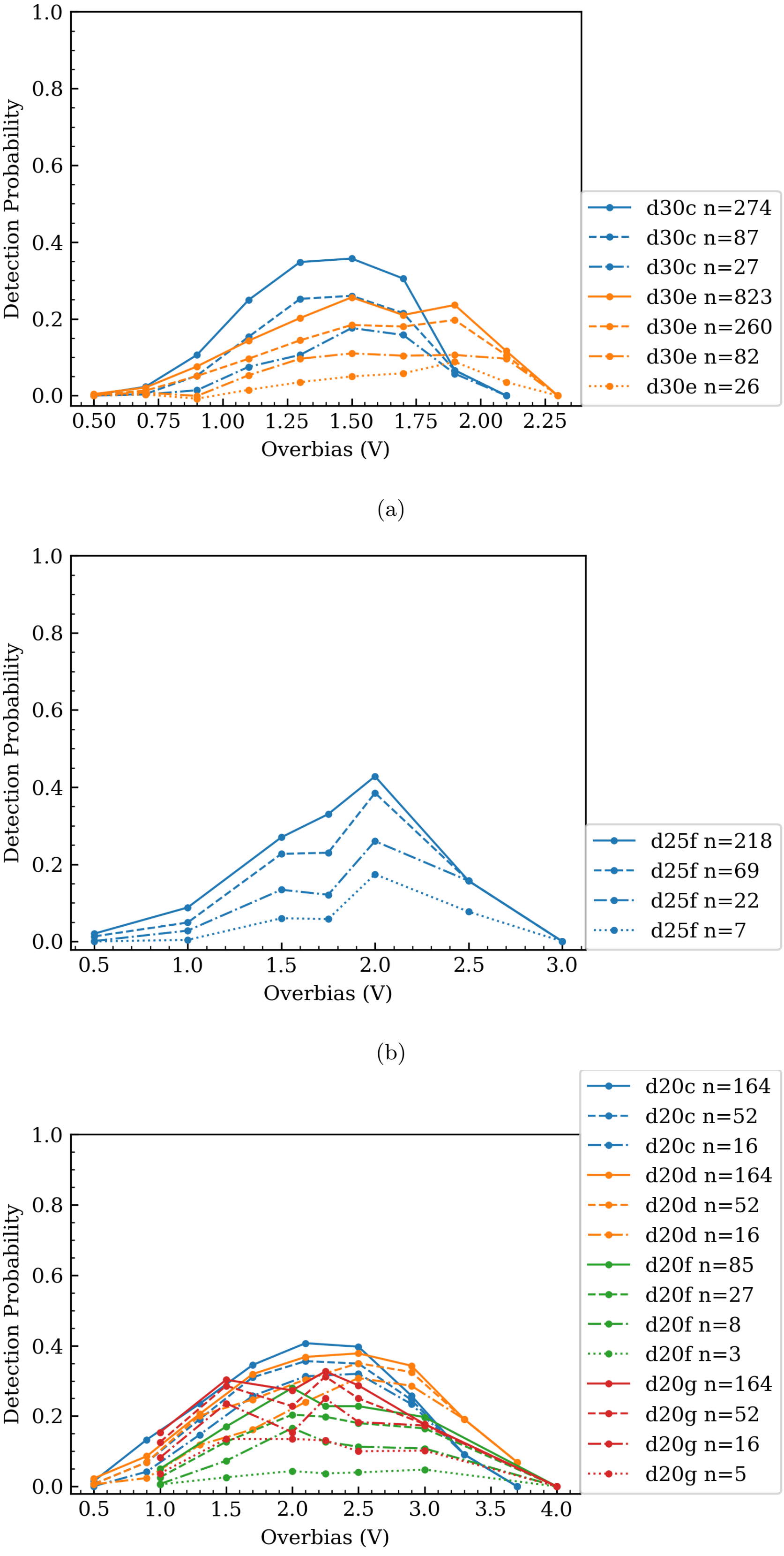
FIGURE 4.25: Temperature dependence of DCR in a D20 SF1319 diode. The measurements were performed with a 5 ns wide 10 V pulse at a repetition frequency of 100 kHz.

4.5.4 Pulsed Laser Measurements

Pulsed laser measurements were performed in the low-temperature probe station using the pulsed picosecond laser using the setup described in chapter 3. Due to the poor DCR characteristics of SF1318 and lack of adequate punchthrough before DCR saturation, pulsed measurements were performed successfully on SF1319 only. The low-temperature probe station was used due to packaging failures during bonding of fabricated SF1319 diodes. Results of the room temperature measurements are shown in figure 4.26.

Temperature dependent photon detection measurements were also carried out. The results for a D20 diode measured from 295 to 150 K are plotted in figure 4.27. Responsivity was found to change with temperature due to the proximity of punch-through and breakdown voltage, and the limited detection region. Therefore, all temperatures are assumed to have the same coupling losses as those determined at room-temperature.

For all pulsed laser measurements, detection probability represents the probability that a given pulse with n photons will be detected.

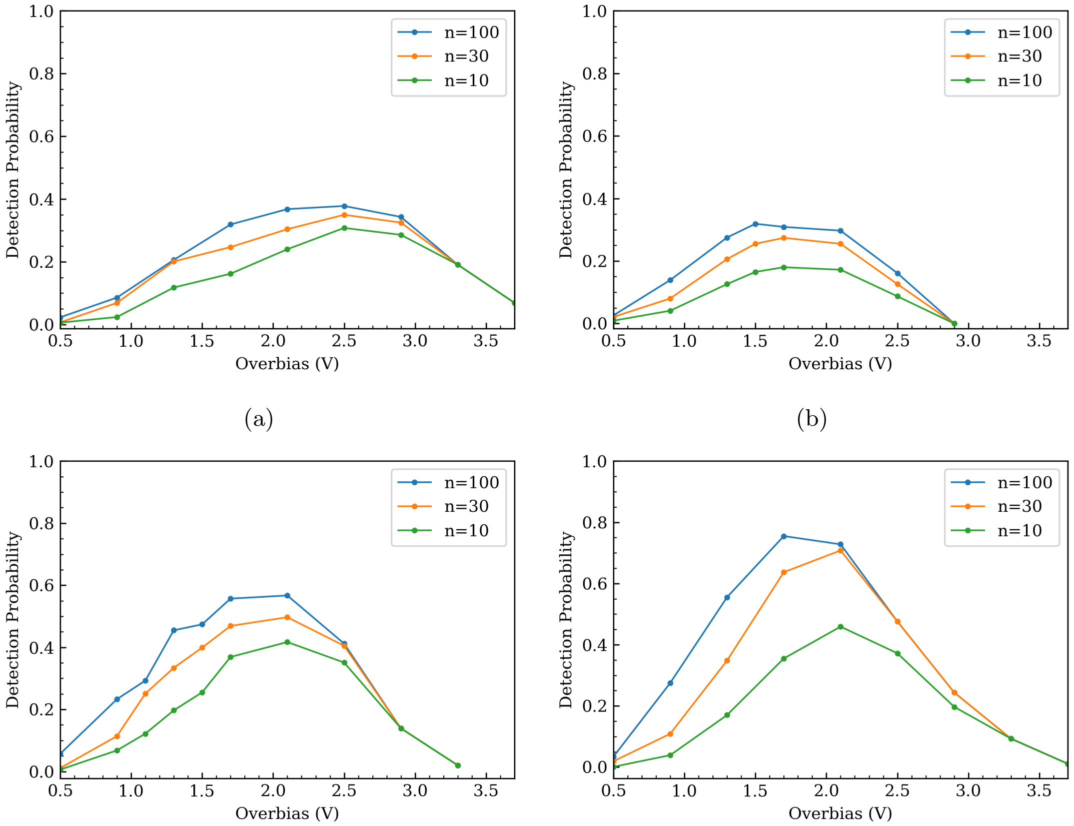


(a)

(b)

(c)

FIGURE 4.26: The detection probability of (a) D30, (b) D25, and (c) D20 diodes for 1550 nm pulses with n photons per pulses as indicated in the legend.



(a)

(b)

(c)

(d)

FIGURE 4.27: The detection probability of a D20 diode at (a) 295 K, (b) 250 K, (c) 200 K, (d) 150 K for different photon levels, with a 5 dB increase in attenuation between each level.

4.6 Discussion of Results

SF1318 diode performance is more limited than expected. SF1318 has a very limited operating range, with DCR saturating within 1 V overbias, corresponding to only 1.5 % overbias at room temperature. This limited voltage range is partly linked to the lack of punch-through which means that the majority of the voltage is dropped across the avalanche region, leading to a rapid rise in breakdown probability compared to if the voltage was more evenly distributed across the diode structure. With the lack of full punch-through it was expected that dark count should be low until the InGaAs absorber depletes, however this is clearly not the case. Cooling the diodes provides only a marginal improvement, and indeed is not desirable since breakdown moves even further from the expected punch-through voltage.

The reduction in DCR from cooling is less than expected. This could be indicative of tunneling, however the avalanche region was engineered to be wide enough that this is not expected to be a significant component of dark current. A possible mechanism to explain the rapid increase in DCR is that since the diodes only start punching-through as breakdown is reached, dark carrier density in the InGaAs absorber is higher than if the diode had punched-through since carriers in the region would only have diffusion and recombination as removal mechanisms. This is different in the case of a depleted absorber, where the carriers are free flowing and would be removed at a higher rate. In a partially depleted structure such as SF1318, when the biasing pulse is applied, the absorber could be partially depleting, causing a surge of dark carriers to be released into the avalanche region which dominates the DCR, and limits the operating range.

SF1319 achieves a wider operating rang with DCR saturation occurs by 3.5 V (5%) overbias, which is comparable to previous high-performance InGaAs/InAlAs SPADs at room-temperature [48]. Detection probabilities for weak laser pulses are in the range of 40 % for several hundred photons, which reduces to 15 to 20 % for several photons. Single photon measurements were not possible due to counts dropping below the backround DCR - partially a limitation of the detection method.

Low temperature measurements of DCR in SF1319 show that there is no increase in DCR saturation voltage. This has to be interperated carefully, since as the diode is cooled, the breakdown voltage is decreasing. When combined with the breakdown being between $V_{\text{p,start}}$ and $V_{\text{p,end}}$, an overbias of 1 V at a lower temperature will actually produce a higher breakdown probability. At the same time, the punch-through effect on DCR observed in SF1318 may be coming into effect, causing a rapid increase in the DCR due to a surge of dark carriers from the InGaAs absorber. This effect appears to become worse as breakdown voltage approaches $V_{\text{p,start}}$, cancelling out some of the beneficial effects of cooling.

Room-temperature photon counting measurements show that the SF1319 diodes are unable to operate in a single-photon regime. Due to the variability in EQE between diodes arising from a variable amount of top contact/bond-pad, and slightly differing fibre alignment, direct comparison between diodes is complicated. Three or four different variable attenuations were used for measurements - 10, 15, 20, and 25 dB. The effective photons per pulse (determined based on measured CW photocurrent) vary between diodes, and are noted in the legend. As the photon level approaches ñ=1, the detection probability begins to significantly drop off. With the combination of reflection losses at the air-semiconductor interface, the lack of full absporption region depletion, and expected P_b of around 60%, this is unsurprising.

Observations from SF1319 low temperature photon counting show mixed benefits to cooling. The diode performance was worse when cooled to 250 K compared to 295 K, with improvements seen at 200 and 150 K. At 250 K the degredation in detection probability was due to a more rapid increase in DCR with overbias. This prevented the diode from reaching higher P_b, and could be due to the effects from the diode depleting while at breakdown voltage. Detection probability increases at 200 and 150 K when compared to room temperature, with maximum detection probabilities of 76, 70, and 45 % for $n = 100$, 30, and 10, respectively, for 150 K. The large drop between $n = 30$ and $n = 10$ at 150 K could be due to only a small amount of the absorber being depleted, limiting QE.

The lack of full depletion has compromised the expected quantum efficiency of the diode, especially when cooled. With less of the absorber depleted, fewer photons will reach the avalanche region. Hence only optical pulses with at least several photons could be detected.

The measured DCR rates and PDE of SF1318 and SF1319 both compare poorly to previous InGaAs/InAlAs SPADs developed [48]. Cooling of diodes fails to improve voltage range significantly, with operation at overbiases above 10 % impossible. Previous InGaAs/InAlAs SPADs were able to operate up to operate up to 20 % overbias [48]. The main difference between those diodes and current diodes is where punch-through and breakdown fall in relation to each other, with previous diodes having approximately 30 V seperation between breakdown and punch-through. This further suggests that it is in fact an effect of punching-through during the biasing pulse that hampers the operational range, with the effect becoming more exagerated as SF1319 is cooled, reducing the benificial effects of cooling that were previously observed [48].

4.7 Conclusion

A comprehensive methodology for designing SPADs was presented. Two different wafer structures were designed and fabricated for low photon level detection. SF1318 diodes, due to lack of punch-through and ability to reach overbiases where the InGaAs abosrber would be depleted enough for reasonable collection efficiency, are found to be not suitable for this application. SF1319 diodes are shown to be able to detect low photon levels down to several photons. Peak laser pulse detection probabilities for n=100, 30, and 10 were found to be 73, 71, and 46 % at 150 K, and 39, 35, and 30 % at room temperature. Since the InGaAs absorber is close to, but not fully depleted, the collection efficiceny limits the diodes' ability to operate at lower photon levels, and impacts on low temperature performance.

Chapter 5

Optimising the Sensitivity of eAPD Based Optical Receivers

5.1 Introduction

APDs are important and highly prevelent optical detectors for use in high-speed optical receivers for long-haul optical communication systems. They offer high internal gain, which is the end result of the impact ionisation process taking place within them. When used in a high-speed optical receivers as part of an APD-TIA circuit, they aid in overcoming the post-amplifier noise, thereby increasing signal-to-noise ratio. In recent years with the large growth in demand for high-speed and throughput data lines, there has been great interest in developing highly sensitive 25 and 40 Gb/s receivers to reduce bit-error rates and allow for reductions in the required optical power needed for long distance transmission.

The sensitivity of an optical receiver is defined as the minimum optical power required to reach a target bit-error rate (BER)- the probability of an error in bit-identification by the receiver. In general, the performance of APD-based optical receivers is limited by several factors which include (i) the excess noise factor; (ii) the stochastic avalanche duration, which increases with gain and decreases the APD's speed, and (iii) the APD's dark current.

Previously accurate simulation models have been made which can simulate the behaviour of APD-based receivers to determine their sensitivity [78–80]. Within these models there is an assumption that both electrons and holes undergo impact ionisation, which is true for the majority of avalanche materials. The avalanche build-up time in such materials is non-deterministic. This means that the previous models cannot be applied, without modification, to an important subclass of APDs known as electron-APDs (e-APDs). In these diode, only a single carrier type - electrons - undergo impact ionisation, which leads to a unique and deterministic form for the impulse response function that must be treated carefully. Examples of materials with these properties include InAs [69] and HgCdTe [81]. Due to these unique properties, e-APDs have the potential to be useful within high-speed, low noise optical receivers.

In this chapter, a model is presented for application to any e-APD. The model captures the effects intersymbol interference (ISI), tunneling current, diffusion current, generation-recombination current, avalanche pulse duration, avalanche gain, and amplifier noise. Calculation of sensitivity is then performed using InAs e-APDs, which demonstrates the model and its funcionality. Previously InAs e-APDs have not been considered for such applications due to their high dark currents, so it is important to explore if their low excess noise and unique e-APD properties offset this limitation.

5.2 Bit-error Rate Model for APDs

This chapter is based on modifications applied to the bit-error-rate (BER) model that was developed in [78] and [79], with significant modifications which need to be explored. This is chiefly due to the unique e-APD impulse response which invalidates many of the assumptions that the previously developed models make, namely the assumption of an infinitely decaying mean impulse response.

In [78], to calculate the receiver sensitivity requires analytic expressions for the mean and variance of the photocurrent mean impulse repsonse function, which arises from an avalanche breakdown event (also known as the APD's stochastic impulse response function). To obtain these expressions, the mean impulse response is modelled using a simplification where the duration of the event is modelled in a non-deterministic way - explicitly, the duration is random and dependent on the impact ionisation coefficients. Such an assumption is not true for an e-APD, where the duration of an avalanche event is well defined since only electrons impact ionise, leading to a well defined, finite mean impulse duration. New expressions are required for the mean and auto correlation functions of the mean impulse response in order to use the comprehensive framework developed in [78] and [79].

The formulae for the mean and variance of the ISI contributions ($\mu_{\text{isi,n}}$ and $\sigma^2_{\text{isi,n}}$) to receiver output from the nth past bit with duration T are given by equations (A3) and (A7) from [79]:

$$\mu_{\text{isi,n}} = \phi_0 \int_0^T \int_{-nT}^{-nT+T} \langle I_p(t-\tau)\rangle \, d\tau dt, \tag{5.1}$$

$$\sigma^2_{\text{isi,n}} = \phi_0 \int_0^T \int_0^T \int_{-nT}^{-nT+T} \langle I_p(\mu-\xi) I_p(\nu-\xi)\rangle \, d\xi d\mu d\nu, \tag{5.2}$$

where ϕ_0 is a constant photon flux present between times $-nT$ and $-nT+T$, while $\langle I_p(t)\rangle$ is the mean impulse response function.

For an arbitrary past bit pattern I_j of length L, the total contributions from all ISI terms, dark current mean and variance (μ_{dark} and σ^2_{dark}), and Johnson noise σ_J, the mean and variance for a "0" bit (μ_0 and σ_0^2) are given by

$$\mu_0(I_j) = \sum_{n=1}^{L} (a_n(I_j)\mu_{\text{isi},n}) + \mu_{\text{dark}}, \tag{5.3}$$

$$\sigma_0^2(I_j) = \sum_{n=1}^{L} \left(a_n(I_j)\sigma^2_{\text{isi},n}\right) + \sigma^2_{\text{dark}} + \sigma_J^2, \tag{5.4}$$

which are equations (9) and (11) from [79]. Here $a_k(I_j) = 1$ when the kth bit in the pattern I_j is a "1", and 0 otherwise. The mean and variance when the current bit is a "1" (μ_1 and σ_1^2) are then obtained by adding these contributions to $\mu_{\text{isi,0}}$ and $\sigma^2_{\text{isi,0}}$, with $n = 0$:

$$\mu_1(I_j) = \mu_0(I_j) + \mu_{\text{isi,0}} \tag{5.5}$$

$$\sigma_1^2(I_j) = \sigma_0^2(I_j) + \sigma^2_{\text{isi,0}} \tag{5.6}$$

The pattern specific BER for $I_j = 1, \cdots, 2^L$ can then be calculated with

$$\text{BER}(I_j) = \frac{1}{4}\left[\text{erfc}\left(\frac{\theta - \mu_0(I_j)}{\sqrt{2}\sigma_0(I_j)}\right) + \text{erfc}\left(\frac{\mu_1(I_j) - \theta}{\sqrt{2}\sigma_1(I_j)}\right)\right]. \tag{5.7}$$

θ is the decision threshold, which is adjusted computationally to minimise the overall BER, given by

$$\mathrm{BER} = \frac{1}{2^L} \sum_{j=1}^{2^L} \mathrm{BER}(I_j). \tag{5.8}$$

which is (14) from [79].

A crucial limitation of models [78, 79] is that the analytical expressions for the stochastic photocurrent's mean and variance must be known. In the process of deriving these, approximations were made for an exponential form of the mean and autocorrelation functions (discussed further in chapter 6). This assumes that these functions are non-deterministic and unbounded in duration, which is not correct for an e-APD since only the electrons impact ionise. To properly model an e-APD, the model must be modified.

5.3 Modification of the Model for eAPD-based Optical Receivers

The basis for the modifications come from the work performed by Saleh *et al.*, who were able to derive dead space dependent analytical expressions to describe the behaviour of single carrier APDs [82]. Additionally, they demonstrated how e-APD properties could offer SNR advantages over regular APDs for optical receiver applications. The mean impulse response function derived in [82] is given by

$$\langle I_p(t)\rangle = \begin{cases} \frac{qv_e}{w} & \text{for } 0 \le t \le \tau_e \\ \frac{cq}{w}\left((v_e+v_h)\left(e^{\beta v_e t} - e^{\beta d}\right) + \frac{v_e}{c}\right) & \text{for } \tau_e < t \le \tau' \\ \frac{cq}{w}\left((v_e+v_h)\, e^{\beta v_e t} - v_h e^{\beta v' t} - v_e e^{\beta d} + \frac{v_e}{c}\right) & \text{for } \tau' < t \le T_e \\ \frac{cq}{w} v_h \left(e^{\beta w} - e^{\beta v' t}\right) & \text{for } T_e < t \le T_f \\ 0 & \text{otherwise.} \end{cases} \tag{5.9}$$

The terms here are given by

$$1/v' = (1/v_e + 1/v_h), \tag{5.10}$$

$$\tau_e = d/v_e, \tag{5.11}$$

$$\tau' = d/v', \tag{5.12}$$

$$T_e = w/v_e, \tag{5.13}$$

and

$$T_f = w/v'. \tag{5.14}$$

d is deadspace, w is avalanche width, v_e and v_h are electron and hole drift velocities respectively. c and c' are constants defined by (23) and (24) in [82], calculated using

$$c = \frac{\alpha + \beta}{2\beta(\alpha d + \beta d + 1)} \tag{5.15}$$

and

$$c' = \frac{c^2(\beta/\alpha + 1)^2}{1 + 2\beta/\alpha - (\beta/\alpha)^2}, \tag{5.16}$$

where α is the impact ionisation coefficient for electrons, and β is derived from the scaled Malthusian parameter, solution to the equation

$$2e^{-\alpha d} - \beta/\alpha = 1. \tag{5.17}$$

This impulse shape is unique to e-APDs, since each stage of the mean impulse behaviour is known precisely. An example is plotted in figure 5.1. In addition, the following

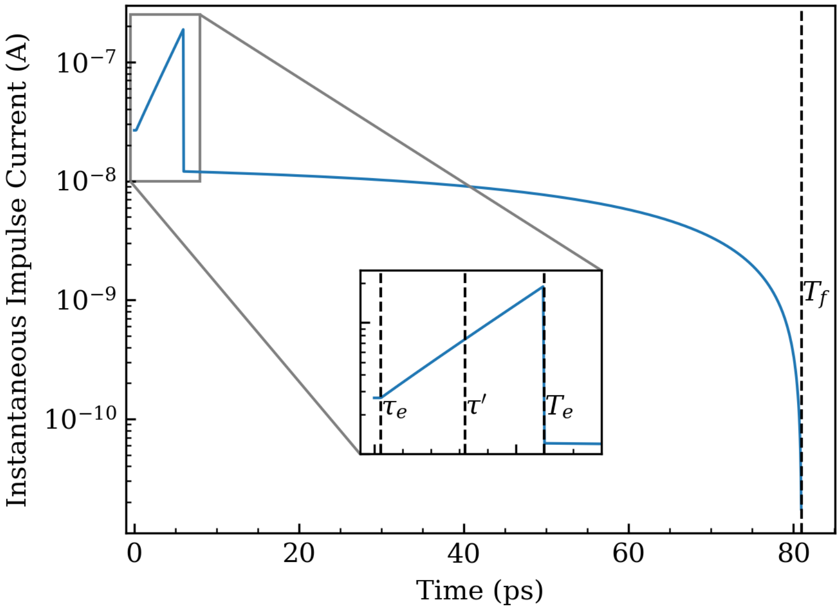


FIGURE 5.1: The calculated mean impulse current for an e-APD with $v_e = 5\times10^5\ ms^{-1}$, $v_h = 0.4 \times 10^5\ ms^{-1}$, and $w = 3\ \mu m$ reverse biased at 18 V. The important times are indicated, where the mean impulse changes shape.

asymptotic expression is used;

$$\langle I_p(t_1)I_p(t_2)\rangle = \frac{c'}{c^2}\langle I(t_1)\rangle\langle I(t_2)\rangle\,. \tag{5.18}$$

This asymptotic result is valid for the case $w \gg d$. These formulae have been validated by Jamil *et al.*, who were able to accurately predict the behaviour of gain and excess noise characteristics in InAs APDs over a range of avalanche widths from 500 nm [83].

An analytical expression reduces computation time, in contrast to a Random Path Length (RPL) model, when generating the mean impulse response function [84]. As in [79], the past bit pattern and the the present bit state is calculated. A key modification is the use of a precise solution for L defined using

$$L = \left\lceil \frac{T_f}{T} \right\rceil \tag{5.19}$$

where L is an integer and the brackets represent the ceiling function. Using this expression the full bit history, for determining the mean and variance, which affects the current detection bit can be encapsulated. Previously this was deduced through trial and error [79].

With the full bit pattern length known, the effects of bulk dark current and tunneling current on the current bit can be calculated precisely. Modifying equations (5.1) and (5.2);

$$\mu_{\text{dark}} = \phi_{\text{dark}} \int_0^T \int_{-T_f}^0 \langle I_p(t-\tau)\rangle\, d\tau dt \tag{5.20}$$

$$\sigma^2_{\text{dark}} = \phi_{\text{dark}} \int_0^T \int_0^T \int_{-T_f}^0 \langle I_p(\mu-\xi)I_p(\nu-\xi)\rangle\, d\xi d\mu d\nu \tag{5.21}$$

where ϕ_{dark} is given by

$$\phi_{\text{dark}} = \frac{I_{\text{tunn}} + I_{\text{bulk}}}{q}. \tag{5.22}$$

$$I_{\text{tunn}} = \frac{(2m^*)^{0.5}q^3EVA}{h^2E_g^{0.5}} \exp\left(-\frac{2\pi\sigma_T(m^*)^{0.5}E_g^{1.5}}{qhE}\right) \tag{5.23}$$

$$I_{\text{bulk}} = J_0A \tag{5.24}$$

I_{tunn} and I_{bulk} are the tunneling current [73], and combined thermal and generation-recombination currents [70], respectively, m^* is the effective electron mass, E is electric field, V is voltage, A is device area, h is planck's constant, k is Boltzmann's constant,

E_g is band-gap, σ_T is the tunnelling parameter, J_0 is bulk dark current density, and q is elementary charge.

After substitution of (5.20) and (5.21) into (5.3) and (5.4), the pattern specific and overall BER can be calculated. The integrals are evaluated using the QUADPACK numerical method [85].

5.4 Model Parameters

Presented here are the parameters used to run all the simulations unless otherwise noted. Most of the parameters used are summarised in table 5.1. The Johnson noise of the TIA in the simulations was given by

$$\sigma_J = \frac{\sqrt{B_{\mathrm{TIA}} i_n^2}}{q R_b}, \tag{5.25}$$

where i_n and B_{TIA} are the input noise current density and bandwidth of a typical TIA at a given bit rate R_b [79]. The electron impact ionisation coefficient was used from [69], while the holes were assumed to not impact ionise. As there is a lack of experimental measurements of InAs electron drift velocities above 10 kV/cm (above which lie the relevant fields required for impact ionisation), simulated values were used from Brennan and Hess [86]. Due to a lack of data on hole drift velocities for InAs, field dependent hole drift velocities were taken from GaAs data [87] due to similar hole mobility characteristics, 400 cm^2/Vs for GaAs, compared to 460 cm^2/Vs for InAs [12]. The fittings to these velocities are shown in figure 5.2. In addition, the total dark current calculation is based on experimental measurements of saturation current, tunneling current, and avalanche gain [69, 70, 88]. A key approximation made, $w \gg d$, is valid for all simulations, with dead-space to avalanche width ratios ranging from 0.016 to 0.037 for operational gain.

Using the modified equations from section 5.3 with these parameters, calculation of the overall sensitivity of the e-APD-TIA system were performed for a variety of avalanche widths and target BER ranging from 10^{-12} to 10^{-3}. A constant electric field profile across the avalanche region was assumed, with no depletion into the p and n regions. For each given avalanche width and data rate, the field-dependent parameters were calculated for each reverse voltage. In order to ensure that the assumptions made for the analytical

Parameter		**Value**	**Ref.**
Temperature T (K)		293	
Device diameter (μm)		20.0	
TIA bandwidth (GHz)	10 Gb/s	8.0	[89]
	25 Gb/s	16.0	[90]
	40 Gb/s	35.0	[91]
TIA noise current (pA/$\sqrt{\text{Hz}}$)	10 Gb/s	6.5	[89]
	25 Gb/s	15.0	[90]
	40 Gb/s	14.0	[91]
Electron effective mass m^*		0.023 m_0	[69]
Bandgap E_g (eV)		0.356	[69]
Tunneling parameter σ_T		1.16	[88]
Bulk dark current density J_0 (A/m^2)		68.8	[70]

TABLE 5.1: Model parameters for InAs e-APDs.

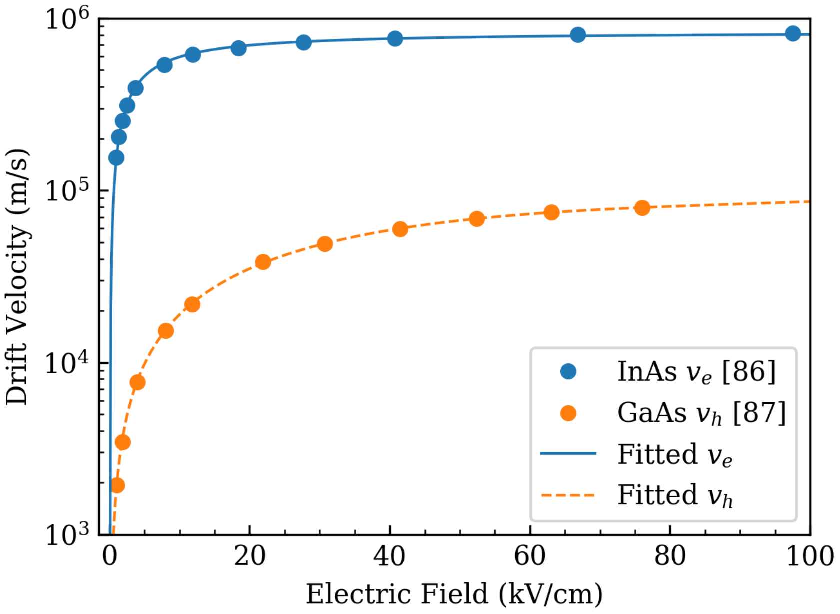


FIGURE 5.2: Field-dependent electron drift velocity simulations of InAs [86], and experimental field-dependent hole drift velocity of GaAs [87]. The lines shown are the fittings obtained and used in the simulation of InAs.

expression of mean impulse response are valid, minimum avalanche width was limited to 1 μm.

To verify that the numerical integrator used was working correctly, the mean impulse response was integrated over $[0, T_f]$ to obtain gain, and the result compared to the gain

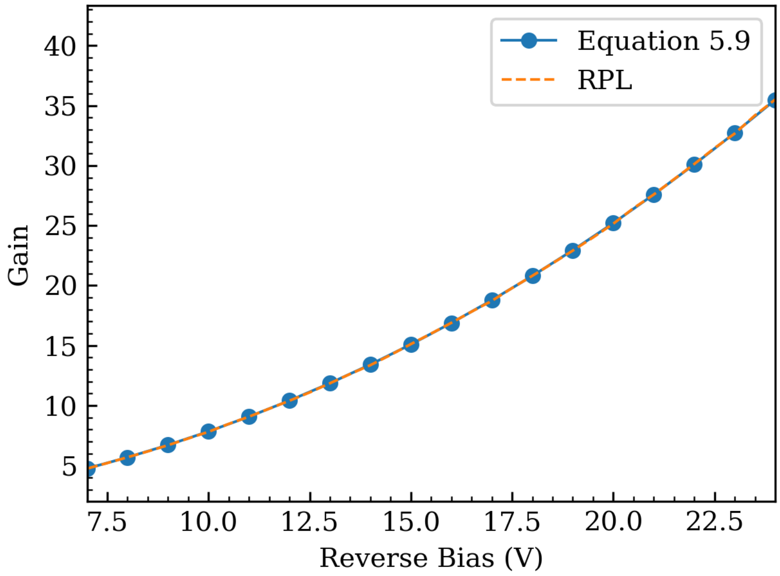


FIGURE 5.3: Verification of the integral of equation (5.9) from $t = 0$ to $t = T_f$ using QUADPACK integration, against RPL simulations for a 3.0 μm InAs APD.

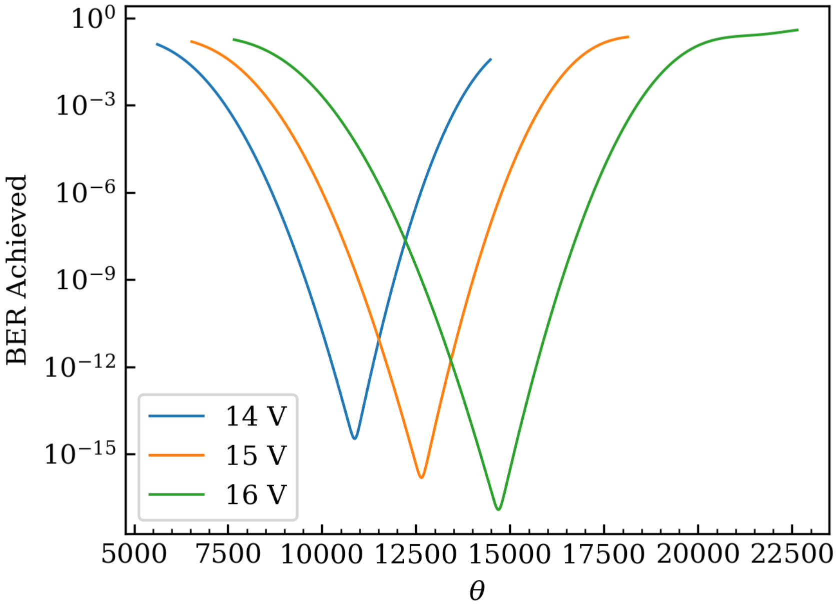


FIGURE 5.4: Varying θ for 14, 15, and 16 V at a fixed photon flux.

obtained from a well tested RPL model [84]. These results are plotted in figure 5.3 and show good agreement.

To accurately find BER, the overall BER needs to be minimised by finding the optimum value for the decision threshold θ. As mentioned previously, this is done computationally. A "sweep" is performed, and overall BER calculated at each value. The point before overall BER begins to increase is chosen as θ. This is process demonstrated in figure 5.4. As voltage is increased for the simulated device, the previous value of θ is chosen as the starting value for further sweeps, thereby speeding up the process after an initial θ is found.

A flowchart of the general process for calculating sensitivity is shown in figure 5.5. The process shown is the calculation of the required number of photons to achieve the requested target BER for a given avalanche width, device radius, reverse bias, and bit rate. The process is performed for multiple different voltages for each width. The device is implemented as an object which contains all of the information about the diode's state. The device object has a bias method which, when called, applies the desired reverse bias, after which the object recalculates all the field dependent parameters. Sensitivity simulations are implemented as seperate functions which act on the device. In principal the device object developed is highly portable and can easily be used with unrelated simulations.

5.5 Simulation Results

5.5.1 Effects of Gain, Bandwidth, and Dark Current

For each e-APD avalanche width, as the reverse bias is increased, the gain, bandwidth, and dark current characteristics will change. Figure 5.6 demonstrates this for a InAs APD with a 3 μm avalanche width operating with a 10 Gbit/s data rate and a target BER of 10^{-12}. As the reverse bias voltage increases, the gain, bandwidth, and dark current all increase. Sensitivity improves as the gain and bandwidth do, however the dark current works to limit the beneficial effects somewhat. The major limiting factor becomes the tunneling current which, when it becomes significant, degrades sensitivity. This is consistent with BER simulations for other avalanche materials [79, 80], although the effects in InAs are more drastic due to the relatively small, direct bandgap. This high tunneling current becomes even more pronounced at smaller avalanche widths, which means that it is not optimal to operate an InAs e-APD with avalanche widths smaller than 1 μm.

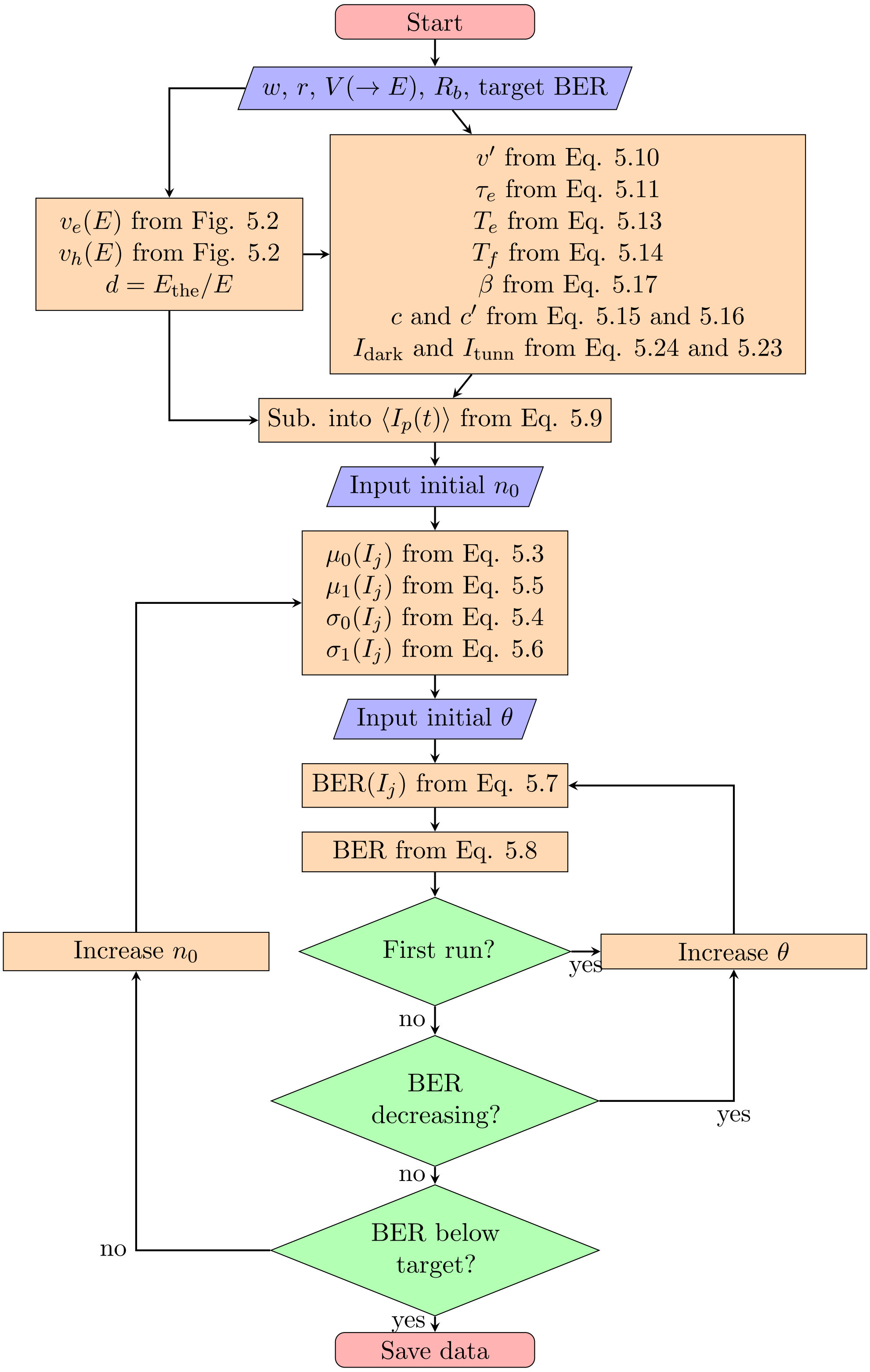


FIGURE 5.5: Flowchart for calculating sensitivity at a given avalanche region width w, voltage V, bit rate R_b, and target BER. In practice this is part of a larger process in which V, w, R_b, and target BER are also varied.

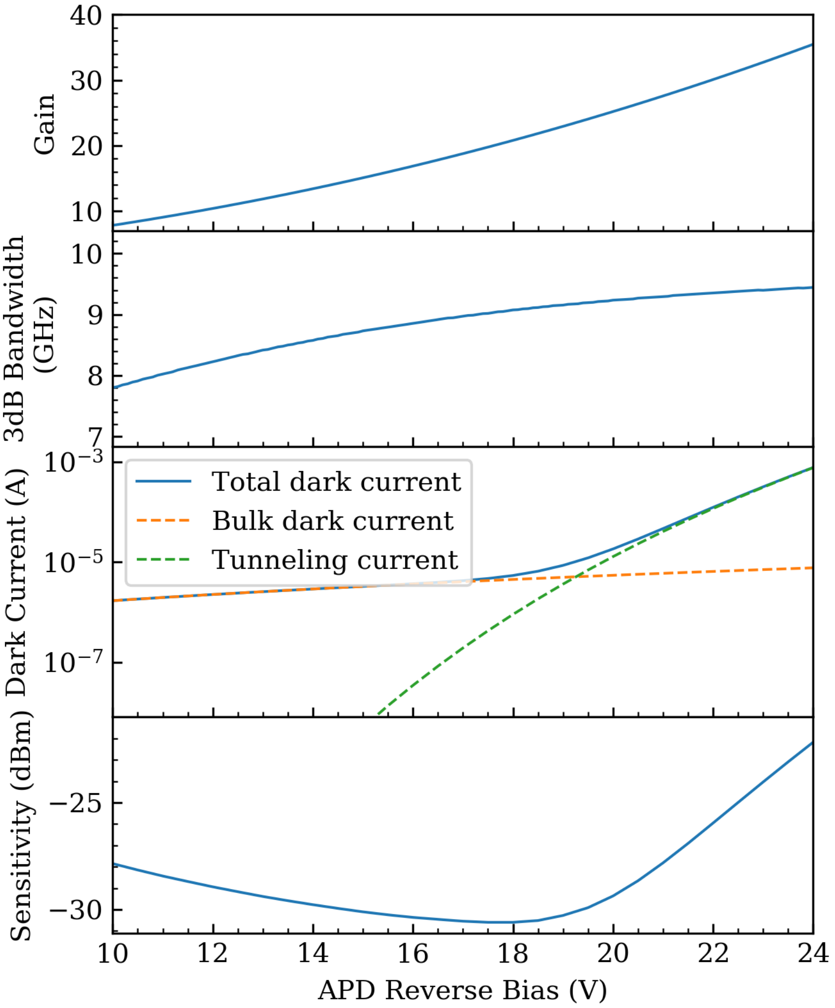


FIGURE 5.6: Gain, 3 dB bandwidth, dark current, and sensitivity simulations with $w = 3.0$ μm. The optimal operating voltage is at -18 V for a 10 Gb/s bit rate and 1×10^{-12} BER.

An interesting observation to note is the behaviour of the 3 dB bandwidth. It can be seen in figures 5.6 and 5.7 how the bandwidth does not decrease with gain, and indeed actually increases slightly due to the field dependent drift velocities used. This demonstrates a major property of e-APDs. Due to a lack of hole impact ionisation, there is no longer a gain-bandwidth limitation as seen in two-carrier APDs. This behaviour is as predicted by Emmons [75].

5.5.2 Sensitivity at Different Avalanche Widths and Bit Rates

As well as allowing for the optimal operating voltage and characteristics to be predicted for a certain eAPD, it is possible to vary the avalanche region width to determine the

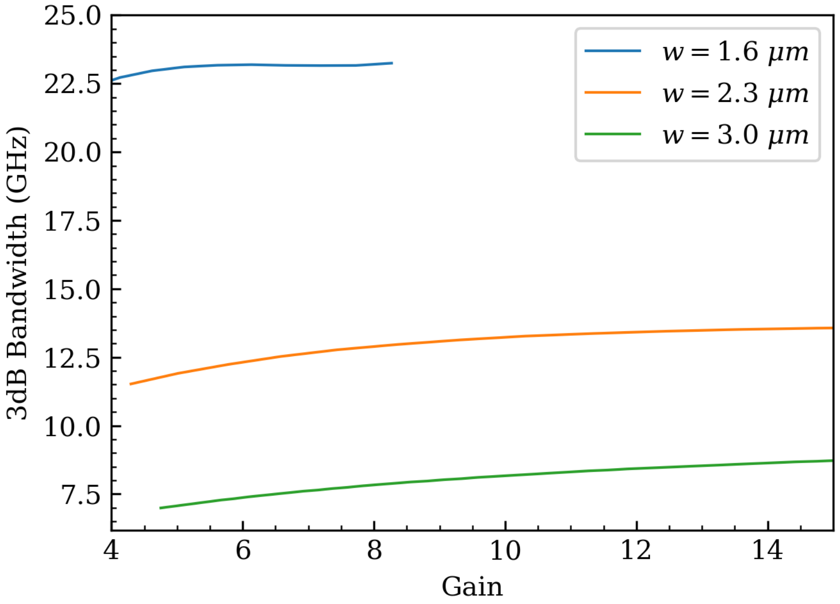


Figure 5.7: Gain - bandwidth properties of three simulated InAs e-APDs. The bandwidth is largely independent of gain, other than an initial increase due to the field dependent drift velocities.

optimal width for operation at a given bit rate. Simulations were performed for 10, 25, and 40 Gb/s data rates with a target BER of 10^{-12} for avalanche widths ranging from 1 μm to 5.5 μm. The simulations yielded minimum sensitivities of -30.60, -22.74, and -19.17 dBm for data rates of 10, 25, and 40 Gb/s, respectively. The optimal avalanche widths corresponding to these were 3.0, 2.3, and 1.5 μm. The sensitivity as a function of reverse voltage is plotted in figure 5.8 for a selection of avalanche widths. It can be seen how the optimal sensitivity moves to a higher voltage as width is increased, since the tunneling current decreases, which allows the diode to reach a higher gain before tunneling current becomes significant. The overall sensitivity worsens for larger avalanche widths, since the bandwidth begins to decrease which leads to higher intersymbol interference.

The optimal sensitivities from each of these simulations are compared together in figure 5.9 for the different bit rates. The optimal width is largely determined through an interplay between tunneling current and bandwidth. With a reducing bandwidth requirement, InAs is able to exploit the lack of gain-bandwidth product limiation further. Based on these observations, e-APD avalanche width needs to be maximised (subject to bandwidth) to maximise gain and reduce tunneling current.

The optimal widths found are larger compared to InP and InAlAs APDs [79, 80].

Unlike these materials, the gain in an e-APD is no longer limited by the bandwidth, and so there is no single optimum operating gain across widths, allowing for much larger avalanche widths [79]. This demonstrates how design considerations for high speed, highly sensitive e-APDs differ from other diodes, and is a key contribution of this work.

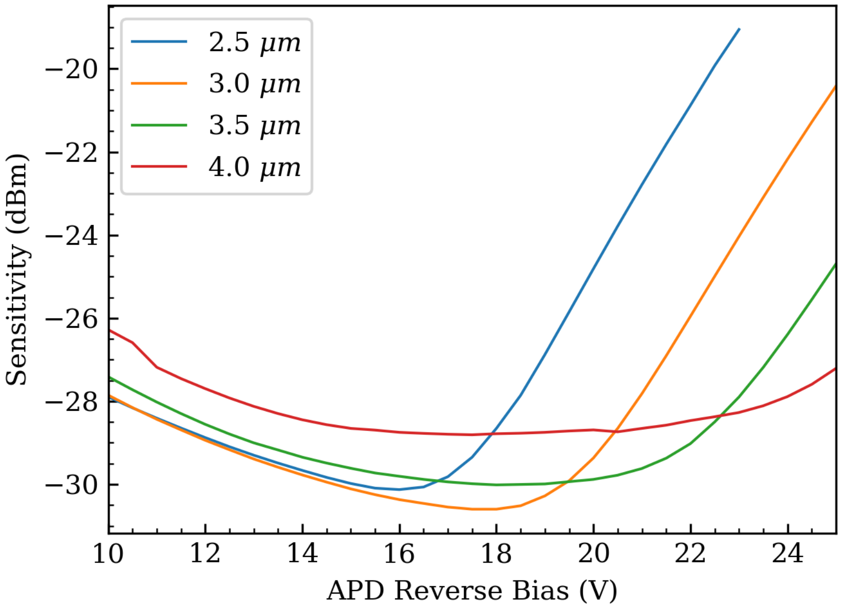


FIGURE 5.8: Sensitivity versus APD reverse bias for several avalanche widths. The data rate is 10 Gb/s, and the target BER is 1×10^{-12}.

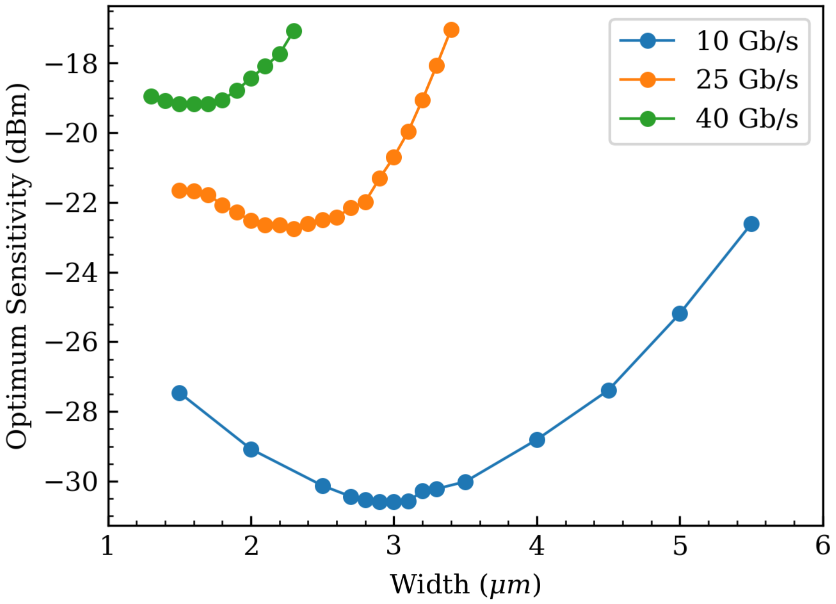


FIGURE 5.9: Gain-optimised sensitivity simulations for a range of different avalanche widths, performed for 10, 25, and 40 Gb/s bit rates with a target BER of 10^{-12}.

5.5.3 Sensitivity with Different Target BERs

For each of the optimal widths determined for a given bit rate in the previous section, simulations were performed with different target BER. These are of interest as depending on the exact requirements of a detection scheme, a higher BER is acceptable. In particular, PON schemes generally operate with a target BER of 10^{-3}.

Results of simulations for different target BER for InAs e-APDs detecting at 1550 and 1310 nm wavelengths, with a comparison to existing high speed, high sensitivity detectors, are plotted in figure 5.10. The previous optimum widths of 3.0, 2.3, and 1.5 μm are used for their respective BERs. Comparison receivers include InAlAs based APDs [59, 92], Si/Ge APDs [93], and SOA-PINs [63, 94, 95].

For 1550 nm operation, InAs e-APDs perform comparably well at 10 and 25 Gb/s to InAlAs APD based receivers [59, 92]. Extending to 40 Gb/s however, the peformance of InAs e-APDs is worse than comparable SOA-PIN from [63, 94, 95], other than at lower BER. Performance is better than Ge/Si APDs for 1550 nm 25 Gb/s detection (not shown due to sensitivity being too low), which have a sensitvity of -16 dBm for 10^{-12} target BER.

As seen in figure 5.10(b), for 1310 nm operation Ge/Si APD based receivers offer better sensitivity than InAs e-APDs, while closely matching InAlAs APDs for 25 Gb/s detection. These comparisons will change over time, as the performances of each of the respective technologies improve. The key to improving the sensitivity of InAs e-APD based receivers will be reducing their tunneling dark current density, which could be possible through engineering a modifications to tunneling barrier shape, and improvements in growth quality and fabrication to reduce bulk dark current density.

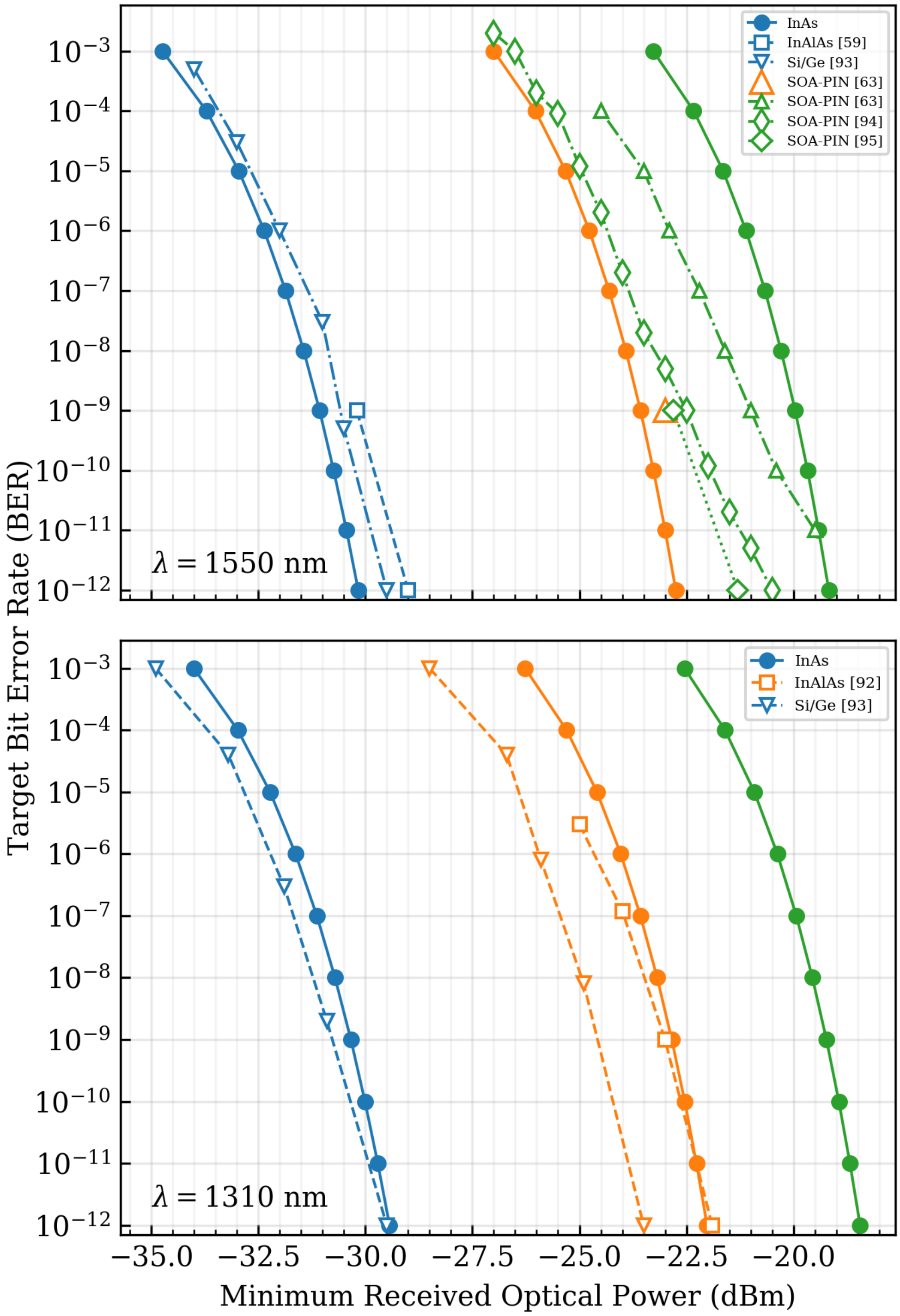


FIGURE 5.10: Variable target BER versus optimum sensitivity for InAs APD-TIA combinations for (a) 1550 nm and (b) 1310 nm wavelength operation. The data rates are 10 (blue circles, $w = 3\ \mu m$), 25 (orange circles, $w = 2.3\ \mu m$), and 40 Gb/s (green circles, $w = 1.5\ \mu m$). Other high speed detector systems (hollow symbols, using identical color scheme for the different bit-rates), such as InAlAs APDs [59, 92], Ge/Si APD [93], and SOA-PIN combinations [63, 94, 95] are included for comparison.

5.6 Conclusion

A sensitivity model taking into account the unique avalanche characteristics of e-APDs has been developed and presented for integrated eAPD-TIA systems. Demonstration of the model was done through simulation of InAs e-APDs at room temperature, using their current performance parameters. Simulations have shown optimum sensitivity values of -30.6, -22.7, and -19.2 dBm at 10, 25, and 40 Gb/s bit-rates, respectively. Results indicate that although InAs e-APDs can offer an improvement in sensitivity over InAlAs APDs at 10 and 25 Gb/s, they are not able to compete with SOA-PIN-TIA combinations at 40 Gb/s at a wavelength of 1550 nm. For 1310 nm operation, Ge/Si offer better sensitivity at 10 and 25 Gb/s, though these systems have not been demonstrated for a 40 Gb/s bit-rate. InAs e-APDs therefore need to be developed with lower dark currents to be competetive with existing technologies.

Chapter 6

A Comparison of InAs, InAlAs, and InP APDs for High Speed Optical Receiever Applications

6.1 Introduction

With the e-APD model presented in the previous chapter, and the simulations demonstrating its use for InAs based optical receivers, it is valuable to compare to other avalanche materials and to better understand the different design considerations and potential of these systems. This chapter investigates these differences, and compares optical receivers based on three different material systems; InAs, InAlAs, and InP. Previous studies carried out on these materials had differing simulation parameters, making a direct comparison difficult [79, 80]. The same is true when comparing experimental measurements to each other, since there are many variables differing between them, such as TIA noise and TIA bandwidth. In addition, since previous simulations were performed, TIAs with improved performance have been demonstrated [90, 91]. As well as this, the further miniaturisation of commerical APDs means that modern high-speed diodes can be as small as 12 μm in diameter [96]. It is therefore important to compare the different systems under the exact same conditions, with a modernised common parameter set to establish theoretical limits to their performances.

InP based APDs are a long established method for high-speed detection, utilised since the 1980s for 1 Gbps detection [57]. High-speed 40 Gbit/s detection with -19.0 dBm sensitivity has been demonstrated in a InP/InGaAs based APD [97]. Due to inferior impact ionisation coefficients, InP based APDs are being superceded by InAlAs based APDs.

With a smaller k ratio in comparison to InP, InAlAs has lower noise and higher bandwidth performance than InP. High-speed detection up to 50 Gbit/s was achieved using an InAlAs APD with a sensitivity of -10.8 dBm, which is the fastest detection with resonable optical power demonstrated in an APD [60]. In the following sections, InP and InAlAs based APDs are simulated together with the same common parameters as InAs in chapter 5 in order to evaluate the differences in performance. Parameters are then varied for InAs to determine how best to improve the overall performance of InAs based APDs.

6.2 Gaussian Approximation Model

To model the behaviour of the InP and InAlAs based optical receivers, the Guassian approximation model developed for APDs with non-zero k ratio is used [79]. The basis for the model was initially developed by Sun *et al.* and included the effects of ISI and dead space [78]. They used a recursive method for finding the joint probability distribution function for the APD's stochastic gain, M , and the stochastic avalanche duration, t. In addition, the APD's stochastic impulse-response function was approximated in terms of M and t. To account for the stochastic behaviour of M and t, and their influence on the buildup-time-limited bandwidth, a new variable was introduced called shot-noise equivalent bandwidth, defined as

$$B_{\text{sneq}} = \frac{\langle M^2/t\rangle}{2\langle M\rangle^2 F}. \tag{6.1}$$

B_{sneq} is used for calculation of the APD-amplified shot noise current, acting as a correction factor to the usual 3dB bandwidth, B_{3dB}. In general, due to the stochastic coupling between M and t the shot-noise equivalent bandwidth is higher than 3dB bandwidth. Additional parameters defined were analytical expressions for mean and variance of a "0" bit μ_0 and σ_0^2, and the mean and variance of a "1" bit, μ_1, and σ_1^2:

$$\mu_0 = \frac{n_0 M}{2\kappa\lambda}(1 - e^{-\kappa\lambda}), \tag{6.2}$$

$$\sigma_0^2 = \frac{n_0 M^2}{4\kappa\lambda^2}\frac{(1 - e^{-\kappa\lambda})^4}{(1 - e^{-2\kappa\lambda})} + \frac{n_0 M^2 F}{2\kappa\lambda}(1 - e^{-\kappa\lambda} - \kappa\lambda e^{-\kappa\lambda}) + \sigma_J^2, \tag{6.3}$$

$$\mu_1 = \mu_0 + \frac{n_0 M}{\kappa\lambda}(\kappa\lambda - 1 + e^{-\kappa\lambda}), \tag{6.4}$$

and

$$\sigma_1^2 = \sigma_0^2 + \frac{n_0 M^2 F}{\kappa\lambda}(\kappa\lambda - 2 + 2e^{-\kappa\lambda} + \kappa\lambda e^{-\kappa\lambda}) + \sigma_J^2, \tag{6.5}$$

where κ is the bandwidth correction factor, correcting for the discrepency between B_{sneq} and B_{3dB}, and λ is the detector-speed factor, with the terms defined by

$$\kappa = \frac{4B_{\text{sneq}}}{2\pi B_{\text{3dB}}} \tag{6.6}$$

and

$$\lambda = \frac{2\pi B_{\text{3dB}}}{R_b}. \tag{6.7}$$

The parameters in equations (6.2) - (6.5) represent the average values over all possible past bit patterns. Sun *et al.* modelled the output of the integrate-and-dump receiver as a Gaussian random variable, which included effects of ISI [78]. Ong *et al.* extended this model to account for the bit-pattern specific behaviour of the receiver mean and variance, more accurately modelling the effects of ISI [79]. This was the basis for the model presented in chapter 5. The mean impulse response in [79] was approximated as exponentially decaying and infinite, given by

$$\langle I_p(t) \rangle = a e^{-b_{\text{se}} t} \tag{6.8}$$

for $t > 0$, where $b_{\text{se}} = 4B_{\text{sneq}}$ and $a = b_{\text{se}} \langle M \rangle$. Applying equations 5.1 and 5.2 from section 5.2, the ISI contribution to the receiver mean and variance is given by

$$\mu_{\text{isi,n}} = \frac{2n_0 \langle M \rangle e^{-\kappa\lambda n}}{\kappa\lambda}(\cosh(\kappa\lambda) - 1) \tag{6.9}$$

and

$$\sigma_{\text{isi,n}}^2 = \frac{n_0 \langle M \rangle^2 F}{\kappa\lambda} e^{-\kappa\lambda n}(e^{-\kappa\lambda} - 1)(1 - \kappa\lambda e^{-\kappa\lambda} - e^{-\kappa\lambda}). \tag{6.10}$$

The mean and variance receiver of the receiver output caused by the ISI terms is given by

$$\mu_{0,\mathrm{isi}}(I_j) = \sum_{k=1}^{L} a_k(I_j)\mu_{\mathrm{isi,k}} \tag{6.11}$$

and

$$\sigma_{0,\mathrm{isi}}(I_j) = \sum_{k=1}^{L} a_k(I_j)\sigma^2_{\mathrm{isi,k}} \tag{6.12}$$

where $a_k(I_j) = 0$ unless the kth bit is a “1”. These equations are then combined with the contribution from dark carriers that have been generated during the lead up to the current detection bit. This is accounted for using equation 6.2 as the basis, with n_0 replaced with $2n_d$, where n_d is the number of dark carriers present within the bit detection window. The extra factor of 2 compensates for the lack of random modulation of the dark-carrier generation. The contribution from dark carriers within the present bit is obtained by using equation 6.4, with n_0 replaced by n_d. Combining these,

$$\begin{aligned}\mu_0(I_j) &= \mu_{0,\mathrm{isi}}(I_j) + \frac{n_d \langle M \rangle}{\kappa\lambda}(1 - e^{-\kappa\lambda}) + \frac{n_d \langle M \rangle}{\kappa\lambda}(\kappa\lambda - 1 + e^{-\kappa\lambda}) \\ &= \mu_{0,\mathrm{isi}}(I_j) + n_d \langle M \rangle .\end{aligned} \tag{6.13}$$

A similar principal follows for the receiver variance for a “0” bit, which is given by

$$\sigma_0^2(I_j) = \sigma_{0,\mathrm{isi}}(I_j) + \frac{n_d \langle M \rangle^2 F}{2\kappa\lambda}(\kappa\lambda - 1 + e^{-\kappa\lambda}) + \sigma_J^2 \tag{6.14}$$

with the Johnson noise contribution (σ_J) from the TIA also added.

When the current detection bit is a “1” bit, there is an extra contribution from the photon flux for the duration of that bit to equations 6.11 and 6.12. This contribution is the same as the second terms in equations 6.4 and 6.5, resulting in the expressions

$$\mu_1(I_j) = \mu_0(I_j) + \frac{n_0 \langle M \rangle}{\kappa\lambda}(\kappa\lambda - 1 + e^{-\kappa\lambda}) \tag{6.15}$$

and

$$\sigma_1(I_j) = \sigma_0(I_j) + \frac{n_0 \langle M \rangle^2 F}{\kappa\lambda}(\kappa\lambda - 2 + 2e^{-\kappa\lambda}). \tag{6.16}$$

The pattern specific and overall BER values are then calculated using equations 5.7 and 5.8.

6.3 Model Parameters

A table of the parameters used for the simulation of InP and InAlAs APDs is given in table 6.1. InAs simulations are the same as in chapter 5. TIA parameters, temperature, and device diameter are the same as in table 5.1. Due to a lack of reported electron and hole drift velocities for InAlAs, the drift velocities of InP are used in the simulations. The bulk dark current densities in InAlAs and InP are insignificant for the purpose of sensitivity simulations, so are excluded. The simulations consist of two parts, the RPL model [84], and the BER calculator, modified from [79]. The RPL model provides the gain, bandwidth, and noise data, with the simulation components visualised in figure 6.1. Utilising equations (6.13), (6.14), (6.15), and (6.16), the BER is calculated for a given n_0 with equation (5.8) for a bit pattern length of $L = 10$ (the optimum determined in [79]). As in chapter 5, n_0 is increased until the desired target BER is achieved.

Param.	**InP Value**	**InAlAs Value**
α (m^{-1})	$1.41 \times 10^8 \exp\left(-\left(\frac{1.69\times10^8}{E}\right)^{1.23}\right)$ [98]	$2.10 \times 10^8 \exp\left(-\left(\frac{1.62\times10^8}{E}\right)^{1.29}\right)$ [99]
β (m^{-1})	$2.11 \times 10^8 \exp\left(-\left(\frac{1.77\times10^8}{E}\right)^{1.15}\right)$ [98]	$2.40 \times 10^8 \exp\left(-\left(\frac{1.86\times10^8}{E}\right)^{1.36}\right)$ [99]
v_e (m/s)	0.68×10^5 [100]	0.68×10^5 *
v_h (m/s)	0.70×10^5 [100]	0.70×10^5 *
$E_{\text{th,e}}$ (eV)	2.8 [98]	3.2 [99]
$E_{\text{th,h}}$ (eV)	3.0 [98]	3.5 [99]
E_g (eV)	1.344 [76]	1.450 [101]
m_e^*	$0.08m_e$ [76]	$0.023m_e$ [101]
σ_{tunn}	1.16 [98]	1.26 [22]

TABLE 6.1: Parameters used to simulate InP and InAlAs APD sensitivity performance.

The modifications to the model in [79] consist of an improved method of determining θ, which is no longer calculated using an approximation. Instead, the same method is used as in the e-APD BER model, with θ found dynamically (see figure 5.4). This further minimises sensitivity, and is more realistic as to how θ would be set. An additional modification is the use of a single θ for all bit-patterns to determine total BER; ie. the use of θ instead of $\theta(I_j)$. In [79], θ is optimised on a per pattern basis, giving a more

---

*Assumed to be the same value as InP

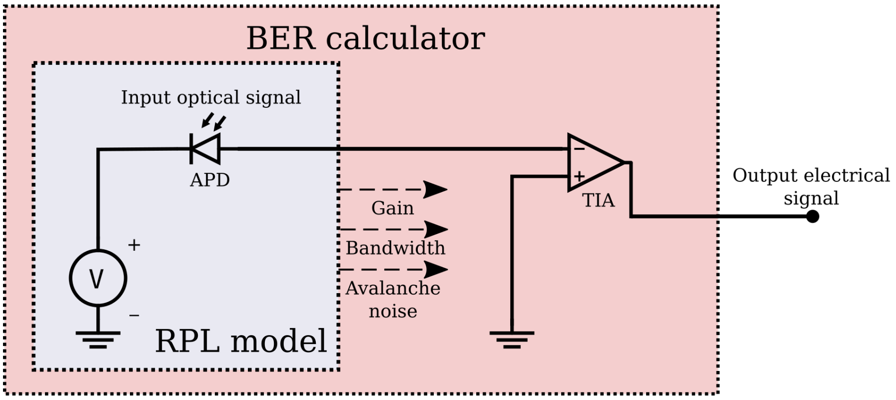


FIGURE 6.1: Block diagram of the simulation components and how they interact to calculate sensitivity.

idealised sensitivity. With the modification given, a single θ is used across all bit patterns, which is a more realistic method.

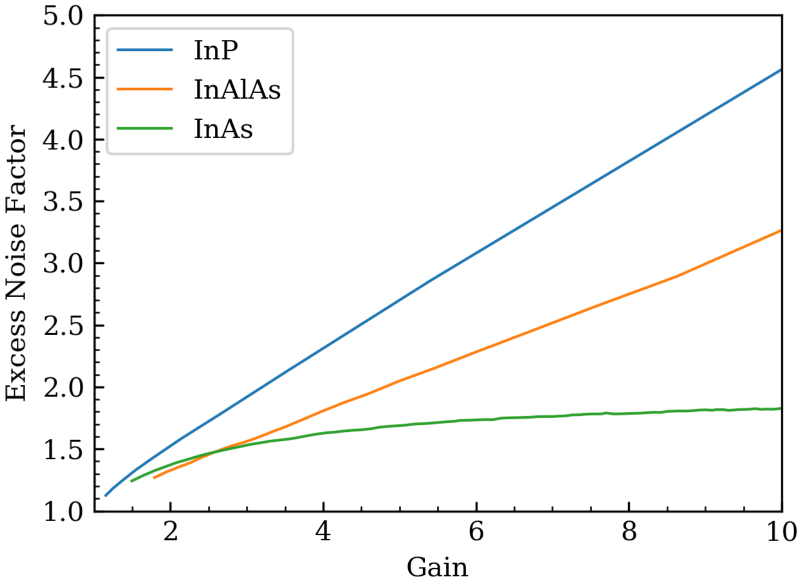
6.4 Results and Comparisons

This section explores the different characteristics between the three avalanche materials.

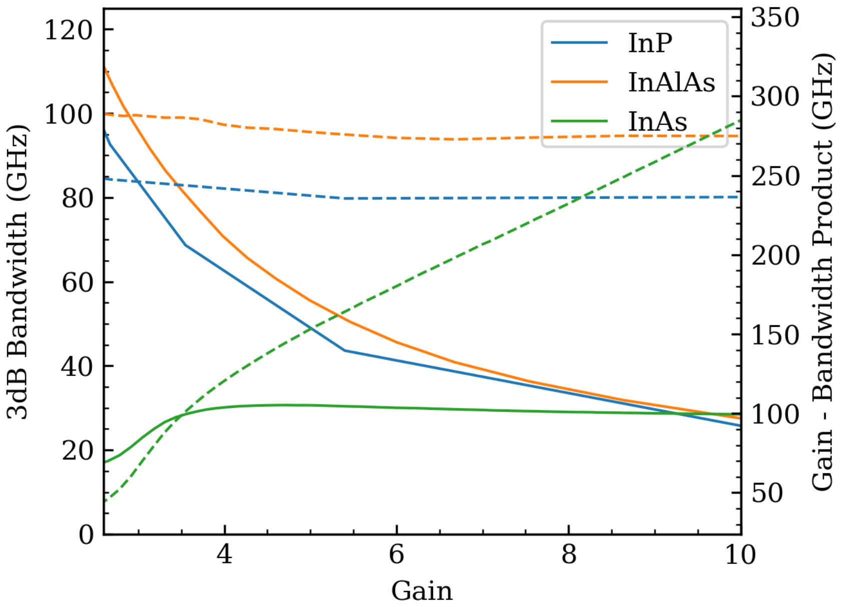
6.4.1 Bandwidth

The gain-bandwidth characteristics of the three materials are compared in figure 6.2. The excess noise increases with gain in all three materials, which acts to limit the 3dB bandwidth. In the case of InP and InAlAs, the excess noise is not constrained due to both the holes and the electrons both undergoing impact ionisation. As the gain increases, the mean impulse duration also becomes longer and so the diodes take longer to recover from a detection event. The implication of this is that the gain bandwidth product remains largely constant as gain increases, as seen in figure 6.2b. InAs, on the other hand, has no such constraint since excess noise never exceeds 2. It is therefore beneficial to operate InAs at as high a gain as possible in order to maximise the gain-bandwidth product.

These comparisons also demonstrate the impact of the ionisation coefficient ratio different between InAlAs and InP on the gain-bandwidth product. For a given avalanche gain and avalanche width, InAlAs will always have a higher gain-bandwidth product, since the average avalanche duration will be shorter. However, the disadvantage of InAlAs



(a)



(b)

Figure 6.2: A comparison of the gain-bandwidth properties of the three materials. An avalanche width of 0.11 μm is used for InP and InAlAs, while 1.5 μm is used for InAs. Due to the excess noise characteristics of InAs, it is able to overcome the gain-bandwidth limitation of regular APDs. In Figure (b) the solid lines represent 3dB bandwidth, and dashed lines represent the gain-bandwidth product.

is that InP is able to reach a higher gain at a lower voltage, due to InP's higher hole ionisation coefficient. This, of course, is not the most important factor for sensitivity.

6.4.2 Optimising Width

Sensitivity simulations were performed where the width of the InP and InAlAs devices increased, for a target BER of 10^{-12}, and are shown in figure 6.3. The overall optimum sensitivities from InAs simulations are also indicated for comparison. Due to the gain-bandwidth limitation imposed by the excess noise in InP and InAlAs, the optimal widths are far smaller than in InAs, which makes comparing on the same width scale difficult. Optimum widths for InP are determined to be 0.14, 0.11, and 0.10 μm for 10, 25, and 40 Gbps bit rates, respectively. For InAlAs they are found to be 0.12, 0.10, and 0.09 μm. These widths are then used to simulate the target BER dependent sensitivity.

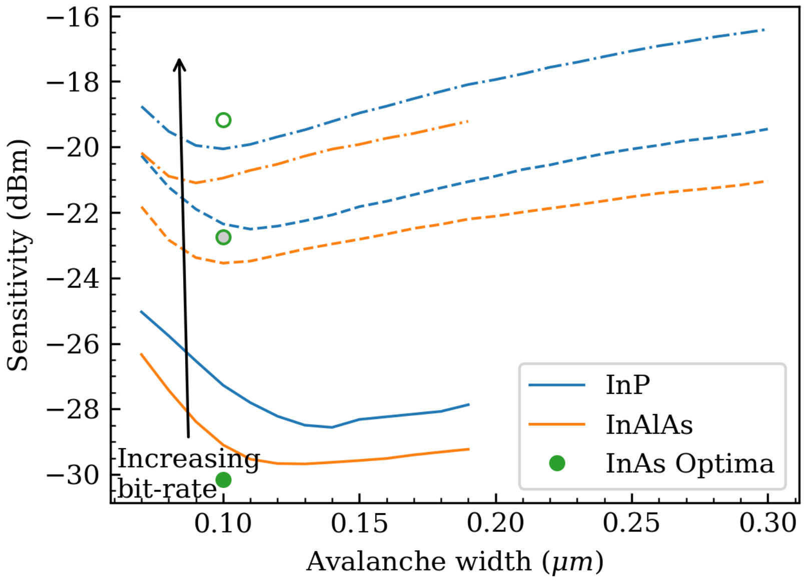


FIGURE 6.3: BER simulations for InAlAs and InP at different widths for a target BER of 1×10^{-12}. InAs results are also indicated at as green dots, though note that these are actually at their respective optimum widths of 3.0, 2.3, and 1.5 μm for 10, 25, and 40 Gbps respectively.

6.4.3 Variable BER

The optimum widths from the previous simulations are taken, and simulations performed for a changing target BER. These simulations are plotted in figure 6.4, along with the

InAs results from section 5.5.3. Results show a similar trend as the comparison to existing diodes in section 5.5.3. InAs is able to outperform InAlAs bit rates below 25 Gb/s, however by 25 Gb/s InAlAs begins to match InAs performance.

Good agreement is seen with realised InAlAs based receivers at 10 Gb/s, and a relatively close agreement at 25 Gb/s [59, 92]. The simulations suggest that InAlAs should be able to perform effectively compared to SOA-PINs at 40 Gb/s, though this has not been shown yet experimentally. A limitation of the current model is that the receiver APD is simulated as a *pin*, while practical APD receivers would use a SAMAPD structure with InGaAs as the absorber. This suggests that for higher bit-rates, this simplified model may be too idealised. This is not the case for InAs simulations, since it is able to act as both the absorber and the avalanche medium.

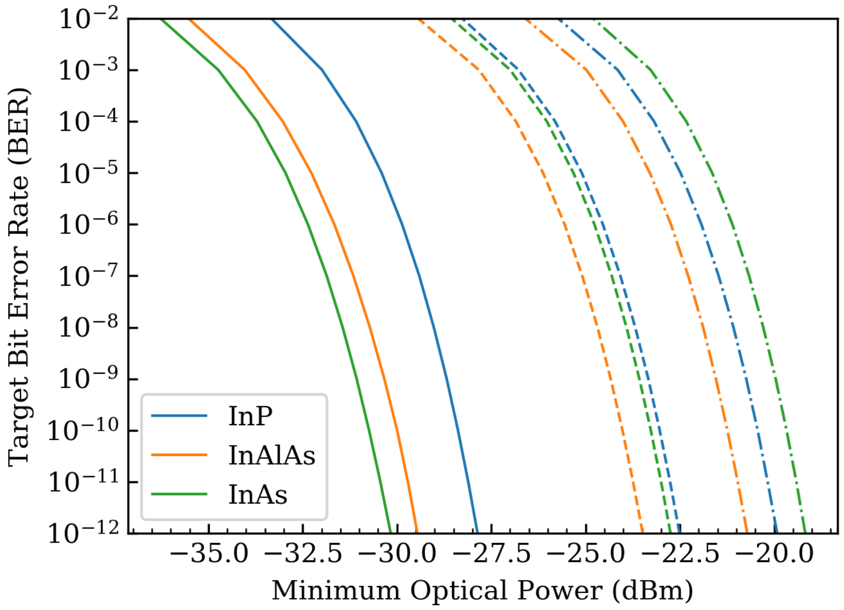


FIGURE 6.4: Comparison of the optimal sensitvities of InP, InAlAs, and InAs diodes for different target BERs. InP $w = 0.14$, 0.11, and 0.10 μm. InAlAs $w = 0.12$, 0.10, and 0.09 μm. 10, 25, and 40 Gb/s are represented by solid, dashed, and dot-dashed lines, respectively.

6.5 Further Optimization of InAs

It is clear InAs needs optimising to be more competetive with other APD materials. This section explores the effects of different parameters on the sensitivity of InAs and how sensitivity might be improved for 40 Gbps operation.

6.5.1 Varying Diameter

In recent years, practical diode sizes have been reducing with advances in fabrication and packaging technology. High-speed commercial APDs have been demonstrated, with effective light coupling, with diameters as low as 12 μm [96]. Waveguides also provide a way to reduce the area of a given diode, while maintaining the ability to couple the light into the diode.

Simulations comparing the effects of varying diameter (and therefore active area) were performed. This is a highly effective way of reducing dark current for InAs APDs. Shown in figure 6.5 is how diode diameter affects the performance of an InAs receiver, for diameters between 1 μm and 40 μm. There is a 0.3 dBm improvement in sensitivity between 20 μm and 10 μm, with the improvement arising from a higher possible operating gain due to a reduced tunneling current. Diameters below 10 μm would require special optics or waveguide structures to be pratctical, the implementation of which would introduce other noise and loss mechanisms, so these are provided as "ideal" values for such a configuration.

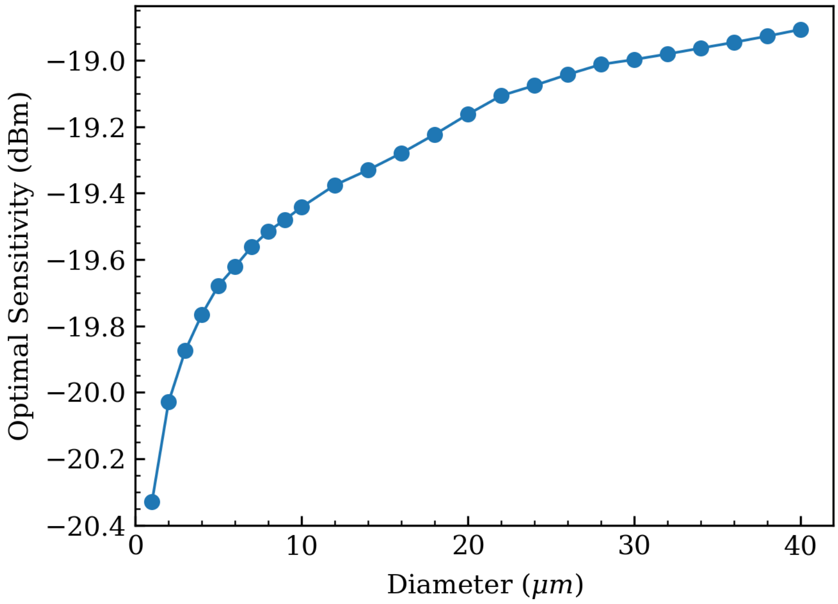
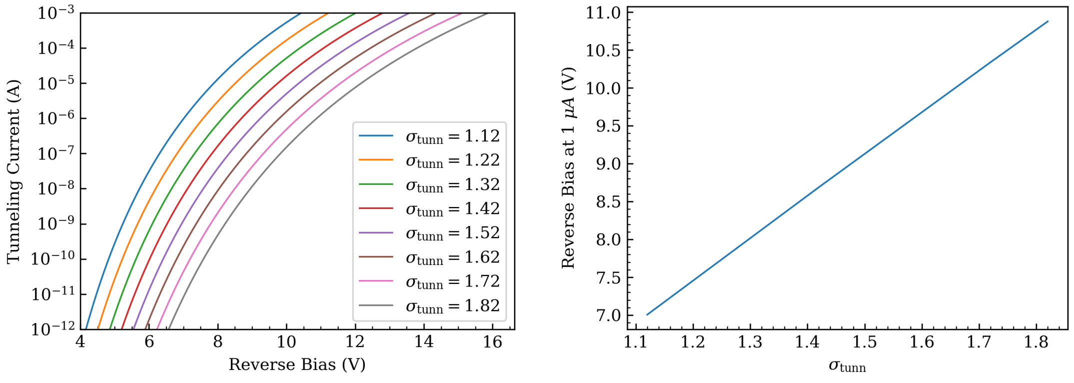


FIGURE 6.5: Sensitivity simulations with a fixed avalanche width $w = 1.5\mu m$, while varying diode diameter.

6.5.2 Tunneling Parameter

The tunneling parameter is dependent upon the shape of the tunneling barrier. These has been shown to change depending on the doping in the region, with a more parabolic barrier shape being described by a larger tunneling parameter [73]. With an increased, graded doping distribution for the avalanche region to increase the barrier parabolicity, the amount of tunneling could be reduced. Using a range of σ_T values from 1.12 to 1.72, the overall optimal sensitivity was calculated for a $w = 1.5$ μm InAs APD, with results plotted in figures 6.6 and 6.7. The impact of the parameter on tunneling current is shown in figure 6.6, with the onset of tunneling current being delayed with increased σ_{tunn}.



(a) (b)

FIGURE 6.6: Effects of varying σ_{tunn} for $w = 1$ μm on tunneling current. (a) shows raw tunneling current, (b) shows the reverse bias at which 1 μA of tunneling current is reached, where a significant impact on sensitivity begins to occur.

Significant improvements in sensitivity can be seen in figure 6.7, even with only slight changes in σ_{tunn}. It is important to note that an increase in doping would have a negative impact on other parameters, such as drift velocity, the effects of which have not been simulated. Crucially, a band structure engineered in such a way would have a differing electric field distribution across the junction. This would lead to a varying drift velocity across the junction, as well as ionisation coefficients, which would affect optimum sensitivity. It is possible that there is a middle-ground that can be found where such a barrier shape would be benificial to sensitivity.

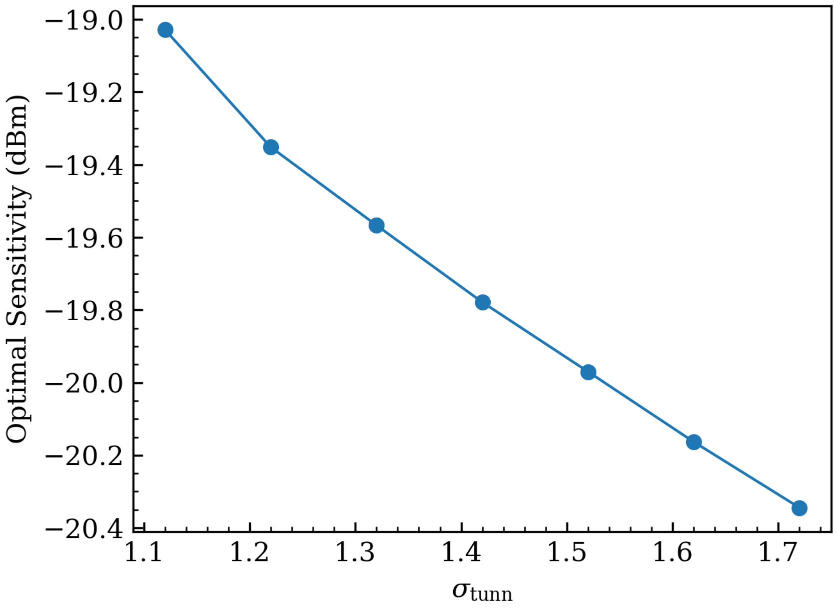
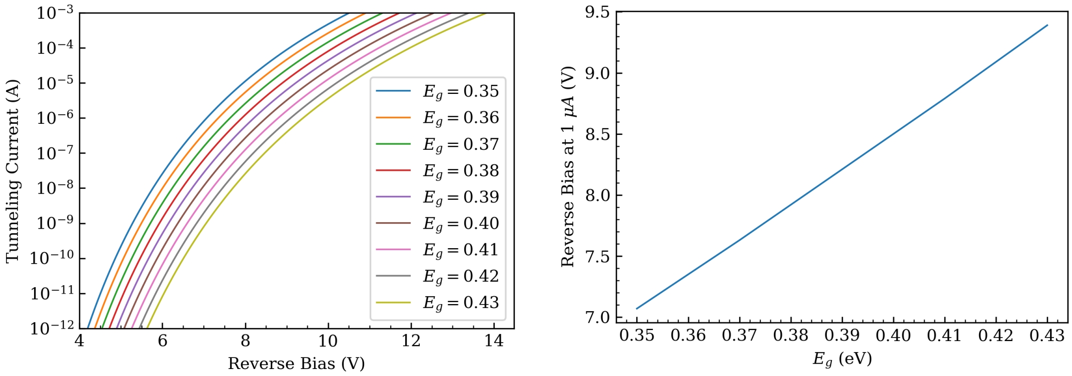


FIGURE 6.7: Sensitivity simulations with a fixed avalanche width $w = 1.5\mu m$, while varying σ_{tunn}.

6.5.3 Bandgap and Temperature

Another effective way of decreasing the effects of tunneling is through an increasing bandgap. This is achievable through cooling, with the bandgap increasing to 0.42 eV when InAs is cooled to 77 K [76]. Figure 6.8 shows how the operational voltage range increases with increasing E_g. By the point the diode is cooled to 77 K, range increases by around 35 % which allows it to reach higher gain before tunneling current becomes significant. At the same time, the thermal dark current would be greatly reduced, effectively to a level that can be excluded from sensitivity simulations.



(a) (b)

FIGURE 6.8: Effects of varying E_g for $w = 1.5\ \mu m$ on tunneling current.

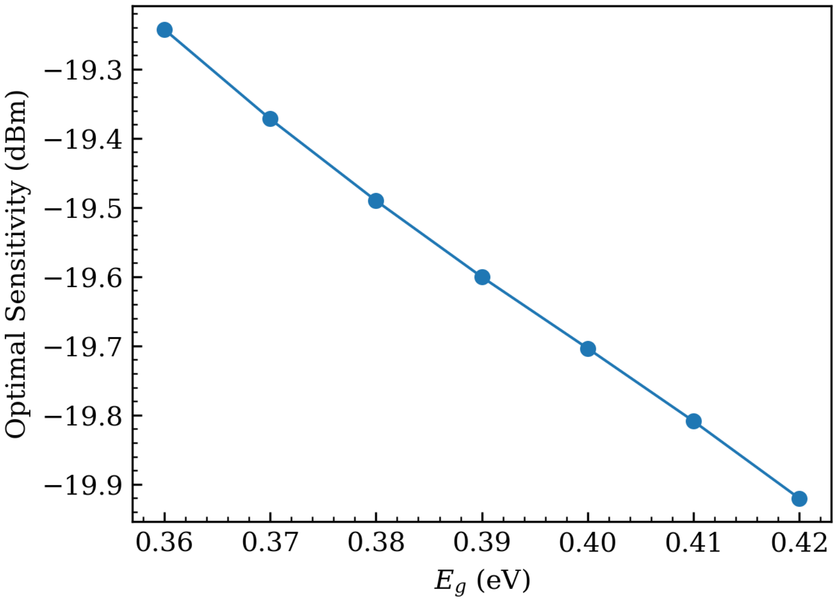


FIGURE 6.9: Sensitivity simulations with a fixed avalanche width $w = 1.5\mu m$.

Sensitivity calculations were performed with varying E_g and are presented in figure 6.9. All other parameters were fixed to their room temperature values. This provides an approximation for how the diodes would benefit from cooling.

Since there are impact ionisation coefficients avaiable at 77 K [102], it is of interest to more accurately model the sensitivity at this temperature. The e-APD properties of InAs are confirmed to persist at low temperatures, so the model is still valid [102]. Unusually, the ionisation coefficients are actually lower for the equivalent field when compared to those at room temperature. This means that with cooling there are two factors at play, the beneficial effects of a increasing band-gap and a reduced I_{dark}, and the negative effects of reduced gain for the same voltage. Other parameters such as the drift velocities are maintained the same. Results of these simulations are shown in figure 6.10 where a comparison is presented between the room temperature and 77 K result.

Overall the beneficial effects prevail over the reduced α, however this degrades performance by around 0.4 dBm when compared to only adjusting bandgap as seen in Figure 6.9. Therefore, cooling to 77 K is of minor benefit, especially when taking into account the added complexity of maintaining the receiver at cryogenic temperatures.

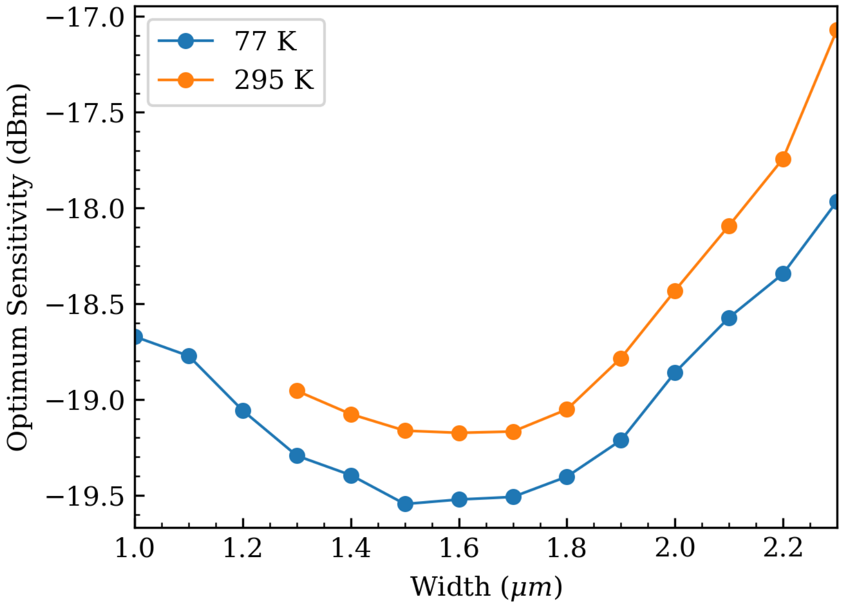


FIGURE 6.10: Comparison of InAs 40 Gbps sensitivity simulations at 77 K and 293 K.

6.6 Conclusion

Comparisons have been presented between InAs, InAlAs, and InP APDs. The same parameters were used where applicable, allowing for an accurate and direct comparison by ignoring any differences that are usually present when sensitivity is reported. InAlAs optimum sensitivities of -29.4, -23.5, and -21.0 dBm were found for 10, 25, and 40 Gb/s operation. For InP these were found to be -27.9, -22.5, and -19.9 dBm for the respective bit rates. InAs is found to offer the best performance at 10 Gbps, while InAlAs is superior at 25 and 40 Gbps. The benefits for switching to InAlAs from InP for the avalanche region for high-speed optical receiver applications are clearly demonstrated, with a minimum of 1 dBm sensitivity improvement shown.

The e-APD properties of InAs require a larger avalanche width to show their benefits due to a large tunneling current caused by the small band-gap. Further calculations were presented where different parameters were varied in order to determine how InAs' parameters would have to change in order for it to be more competetive at 40 Gb/s. The most realistic method for improving the sensitivity is cooling, however this is shown to not have a significant effect due to the decreased ionisation coefficients with cooling. The low temperatures required for improved sensitivity are not practical for data centre applications. Increasing σ_{tunn} is shown to be a possible method, though the negative effects of this were not modelled. A more advanced model of the device properties and electric field distribution within is required to accurately predict the effects of a changing σ_{tunn}. Smaller diode sizes provide a practical way of improving sensitivity, with a 0.3 dBm improvement by switching from 20 μm diode diameter to a 10 μm diameter. Overall it may be possible for an InAs based receiver to achieve performance levels seen in state of the art SOA-PIN systems, however this would involve reducing device radius as well as attempting to modify the tunneling barrier shape to reduce overall dark current.

Chapter 7

Conclusion and Future Work

7.1 Conclusion

Two wafer structures were designed, optimised for SWIR single-photon counting. Wafer SF1318 was designed to punchthrough after breakdown voltage, and SF1319 was designed to punchthrough just before. Due to a rapidly increasing DCR (saturating within 1 % overbias) limiting the operating range to below the punch-through voltage, SF1318 was found to be unsuitable for SWIR single-photon counting. The cause of this rapid rise in DCR is thought to be due to carriers trapped in the absorption region. As the reverse bias pulse raises total reverse bias over the breakdown voltage, the diode absorption region partially depletes, releasing the trapped carriers into the avalanche region, saturating count rate.

The second wafer, SF1319, performed better. However, due to an unexpectedly high doping in the charge sheet, punch-through voltage was higher than designed. With the absorption region not fully depleted, QE was limited and prevented single photon counting, however pulses with several photons were detectable. For larger optical pulses with 10s of photons, detection probability was as great as 75% at 150K, even with only a small part of the absorption region depleted. This work shows promising results for wafers designed to punchthrough just before target operating voltage, and also shows the issues encountered when breakdown voltage falls below the punchthrough voltage.

A new model was developed for the calculation of sensitivity of high-speed optical receivers based on e-APDs, and demonstrated using InAs as the e-APD component.

This model is the first demonstration of the potential of e-APDs to be used in such an application. Calculations showed sensitivities of -30.6, -22.7, and -19.2 dBm for 10, 25, and 40 Gb/s operation, respectively. These values compare favourably to existing high-speed receivers, especially at 10 and 25 Gbps when compared to InAlAs APDs, however for 40 Gbps operation SOA-PIN based receivers perform better [63, 92]. For 1310 nm optical systems, Ge/Si APDs were found to offer a better sensitivity than InAs e-APDs [93]. To improve, InAs e-APDs with lower dark currents need to be developed.

Simulations of InAlAs and InP APD based optical receivers were peformed, and comparisons made to the simulated InAs APD based optical receivers. The simulations used a common parameter set where appropriate, allowing for a direct comparison of the receivers. InAlAs optimum sensitivities were -29.4, -23.5, and -21.0 dBm for 10, 25, and 40 Gb/s operation. InP optimum sensitivities were -27.9, -22.5, and -19.9 dBm for the respective bit rates. InAs offers the best sensitivity for 10 Gb/s bit rates, while InAlAs is best for 25 and 40 Gb/s. InAlAs was shown to outperform InP at all bit-rates, demonstrating its superior gain-bandwidth properties. Further simulations of InAs APDs showed the sensitivity improvements from reducing diode diameter, modifying bandgap parabolicity, and increasing band-gap through cooling.

7.2 Suggestions for Future Work

Chapter 4 demonstrated the negative effects of engineering punch-through above breakdown voltage. Future wafer designs should take this into account, and ensure full punch-through before breakdown. Engineering full punch-through to be as close to breakdown voltage proved inconclusive, since the SF1319 diodes were not fully depleted at breakdown. There could be a potential benefit to such a design over previous designs where punch-through is significantly earlier than breakdown, especially if InGaAs quality is poor [48]. Future wafer designs should explore this further. Structures could be optimised to fully punch-through at the breakdown voltages expected at different temperatures. The DCR and SPDE of these SPADs could then be compared, and the effects of punch-through voltage in relation to breakdown voltage better understood.

To better understand the breakdown mechanisms in the InGaAs/InAlAs SAM structure, InAlAs *pin* SPADs could be designed and fabricated with widths matching those of the SAM avalanche layer. The DCR characteristics of these diodes could then be compared together with the matching SAM SPADs. The InAlAs *pin* SPADs could also then be tested for visible light detection, for which InAlAs has not yet been explored.

Simulations in chapter 5 of InAs based optical receivers at 10 Gb/s showed benefits over existing technologies at 1550 nm [92, 93]. By selecting a TIA with known noise and bandwidth, an InAs APD could be designed and fabricated with a width optimised around these parameters by using the new model. With the given 10 Gb/s TIA in this work, this width would be 3.0 μm. This would allow for the verification of the new sensitivity model, and further the understanding of the potential of high-speed e-APD based optical receivers.

The model for APD based optical receivers in [79] and chapter 6 only models the sensitivity properties of the avalanche region of the APD. To provide a more accurate model, the full SAM structure must be simulated. This would involve creating an improved RPL component for the model which takes into account how carriers absorbed at different positions along the absorption region would affect the bandwidth properties of the APD. This model could also take into account losses due to insufficient absorption thickness. With these improvements, APD based optical receivers could be more accurately simulated at higher bit-rates.

Appendix A

Mask Set: IIG Mesa

The IIG mesa mask set was a modification of the existing mask set known internally as *newpin*. Additions included an identifier for each layer, modification to etch two different radii, AR coating, and an optical mask layer.

A.1 Mask 1

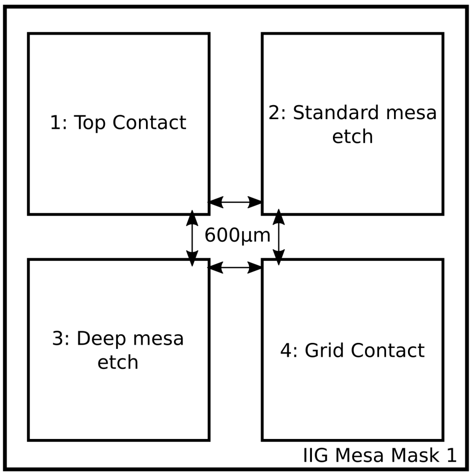


Figure A.1: IIG mesa mask plate 1

Layer	**Title on mask**	**Description**	**Data Polarity**
1	Top contact	Top metal on the mesa	Clear
2	Standard mesa etch	Define mesa area	Dark
3	Deep mesa etch	Define mesa area (10 μm extra radius)	Dark
4	Grid contact	Grid contact on etched area	Clear

Table A.1: Overview of IIG mesa mask 1 layers.

A.1.1 Frame 1: Top contact

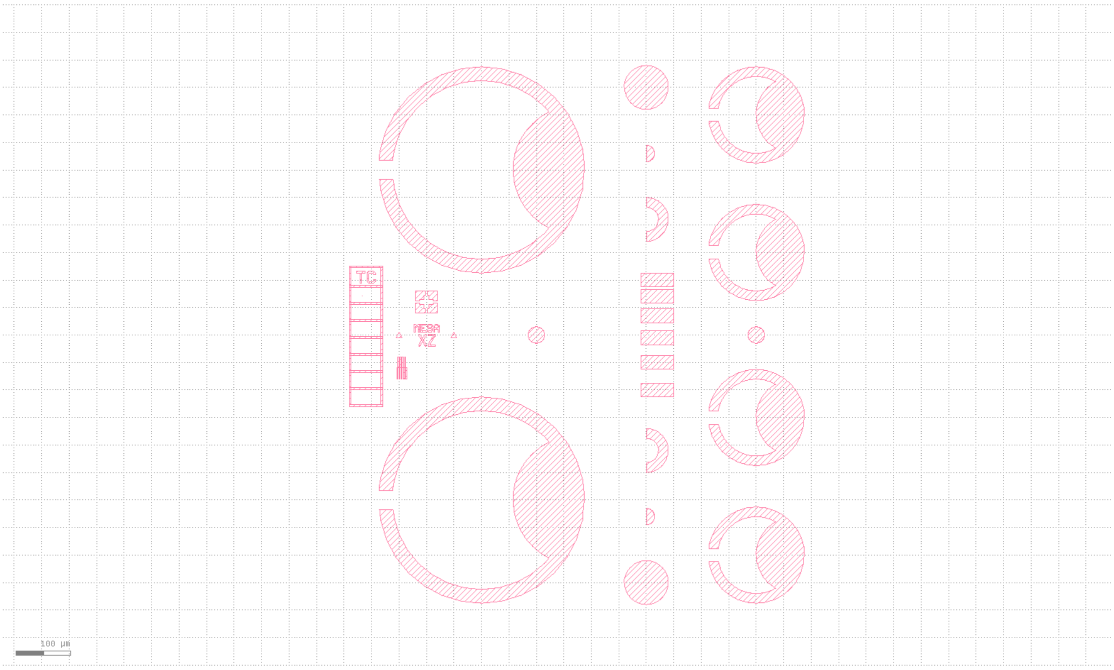


FIGURE A.2: Top contact deposition layer.

A.1.2 Frame 2: Standard Mesa Etch

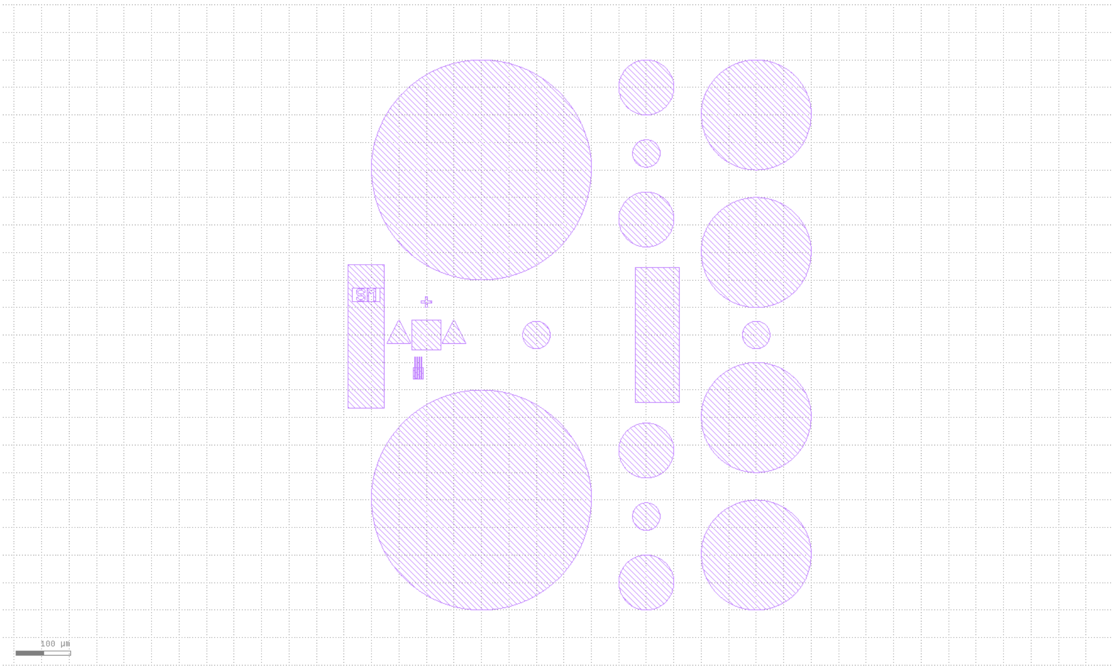


FIGURE A.3: Etching layer to define mesas.

A.1.3 Frame 3: Deep Mesa Etch

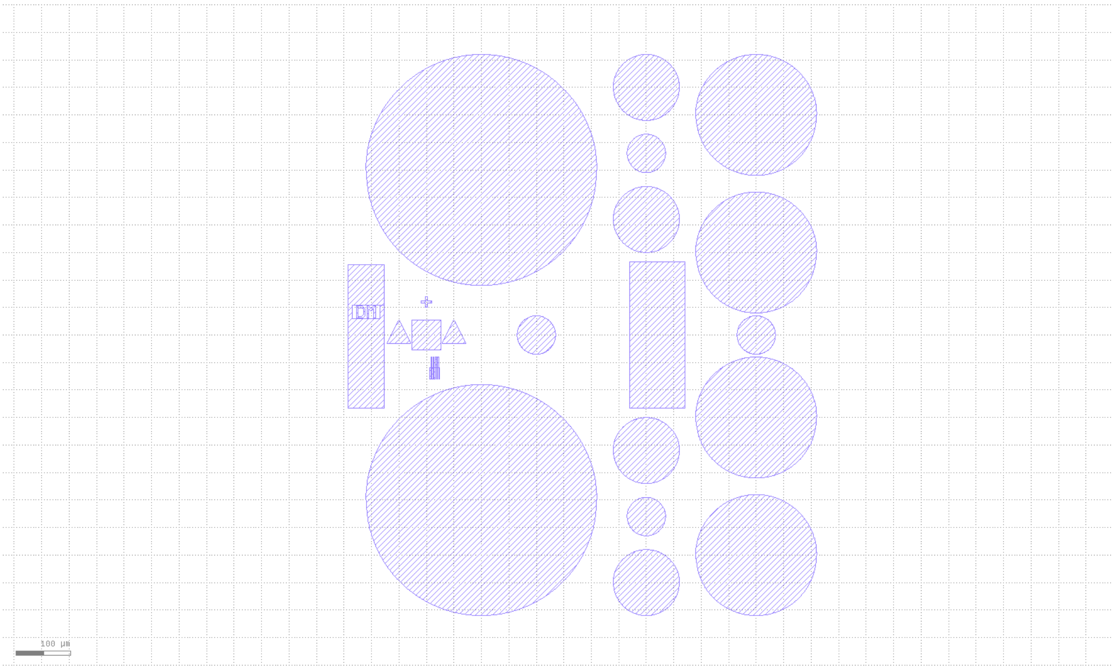


FIGURE A.4: Etching layer to define mesas, designed for depths where undercutting would reach the top contacts if using the standard etch.

A.1.4 Frame 4: Grid Contact

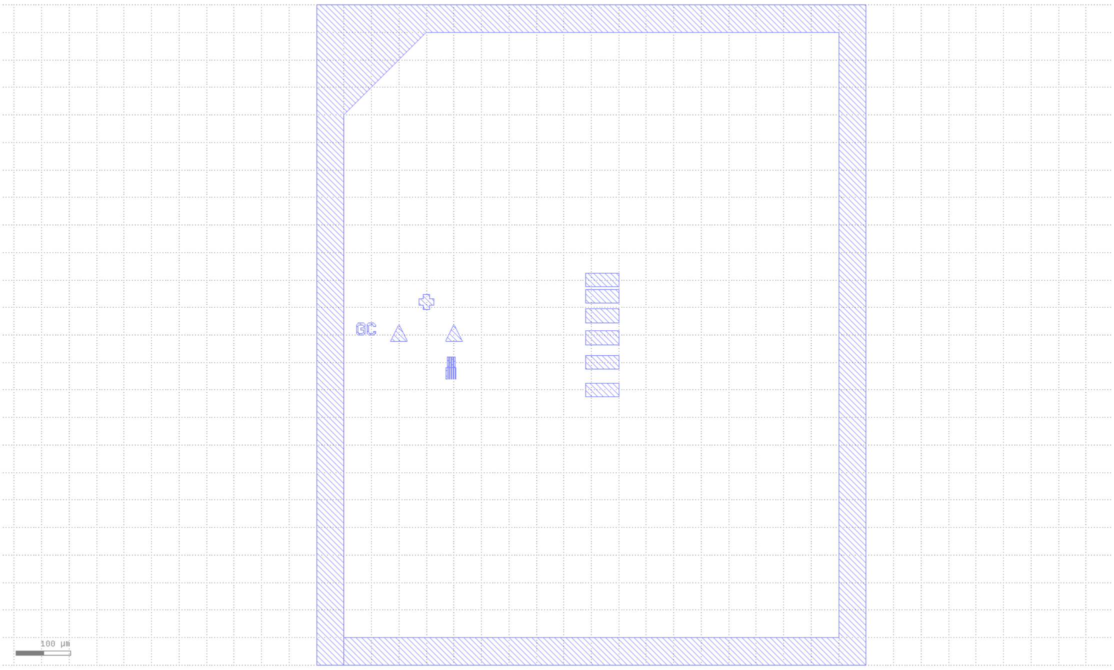


FIGURE A.5: Lower contact deposition layer.

A.2 Mask 2

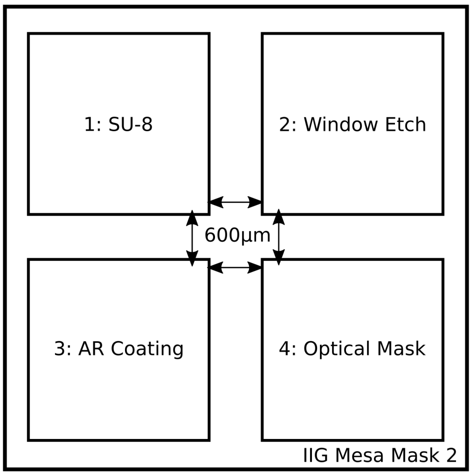


Figure A.6: IIG mesa mask plate 2

Layer	**Title on mask**	**Description**	**Data Polarity**
5	SU-8	SU-8 passivation	Dark
6	Window etch	Etch mesa windows	Clear
7	AR coating	Remove nitride from contact areas	Dark
8	Optical Mask	Deposit metal mask over SU-8	Clear

Table A.2: Overview of IIG mesa mask 2 layers.

A.2.1 Frame 1: SU-8

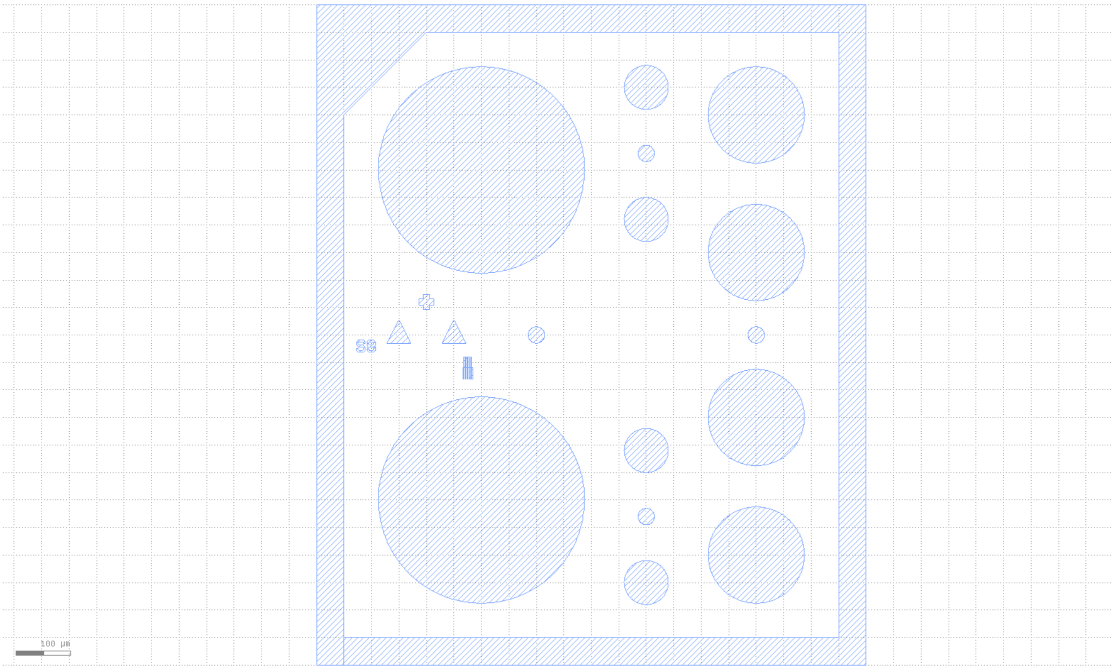


Figure A.7: Exposes SU-8, used to passivate the diodes.

A.2.2 Frame 2: Window Etch

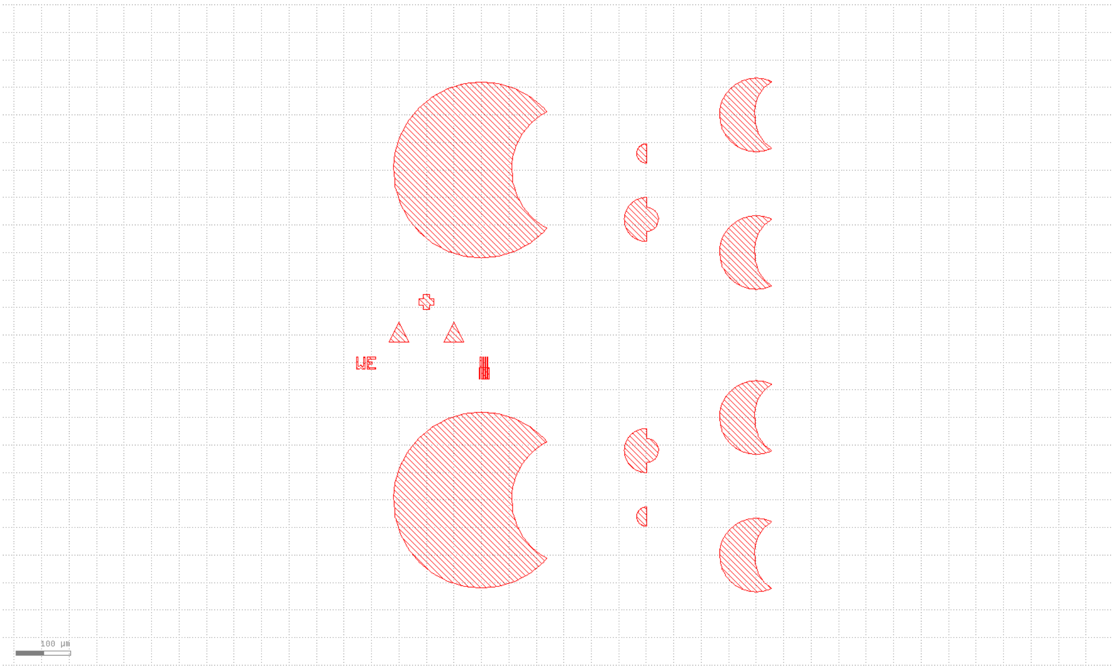


Figure A.8: Etch the windows of the mesa diodes. For use with diodes where the top layer absorbs too much light.

A.2.3 Frame 3: AR Coating

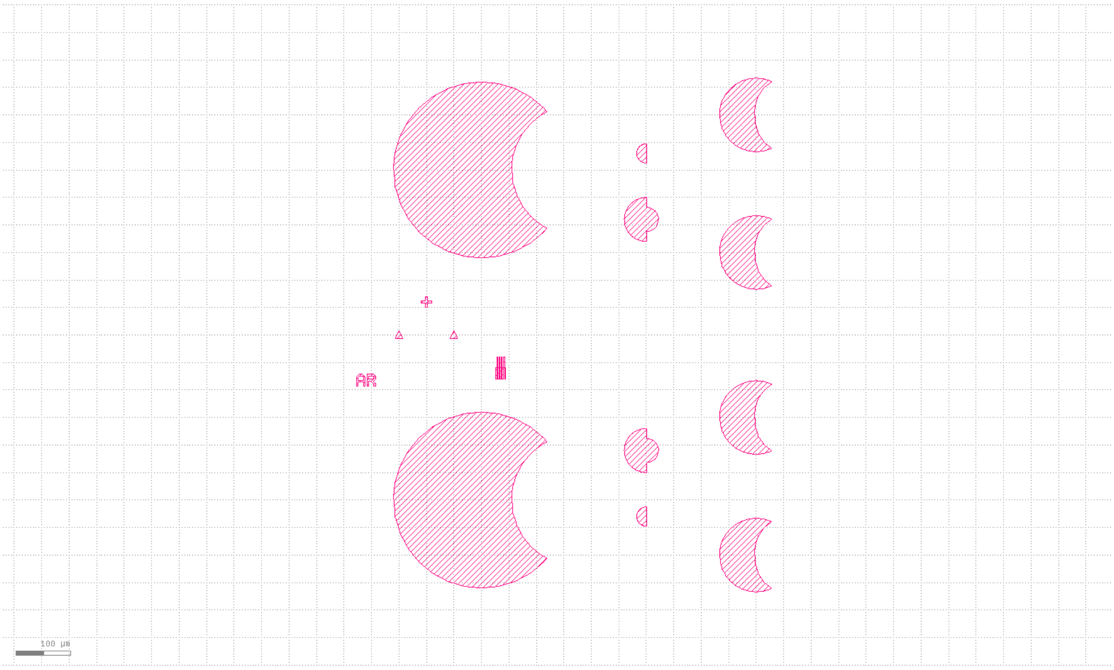


FIGURE A.9: Remove AR coating nitride from all areas other than the mesa windows.

A.2.4 Frame 4: Optical Mask

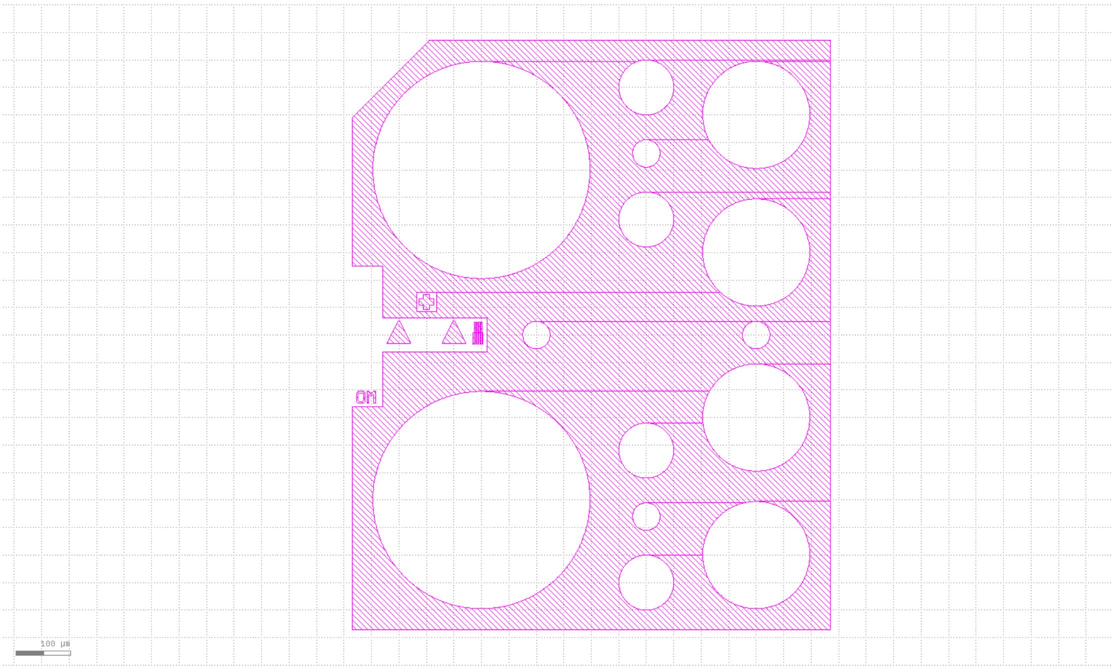


FIGURE A.10: Layer to deposit a metal optical mask, used to prevent side injection into the mesa sidewalls.

Appendix B

Mask Set: IIG SPAD

The IIG SPAD mask set was developed from scratch, optimised for small diodes and designed to be used for single photon detection.

B.1 Mask 1

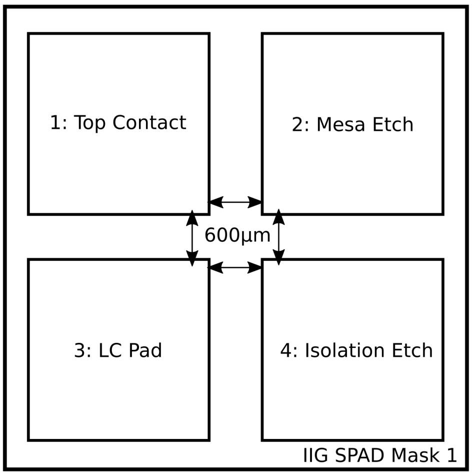


Figure B.1: IIG SPAD mask plate 1

Layer	**Title on mask**	**Description**	**Data Polarity**
1	Top contact	Top metal on the mesa	Clear
2	Mesa etch	Define mesa area	Dark
3	LC pad	Lower contact bond pad	Clear
4	Isolation etch	Etch into the SI substrate around the cells	Dark

Table B.1: Overview of IIG SPAD mask 1 layers.

B.1.1 Frame 1: Top contact

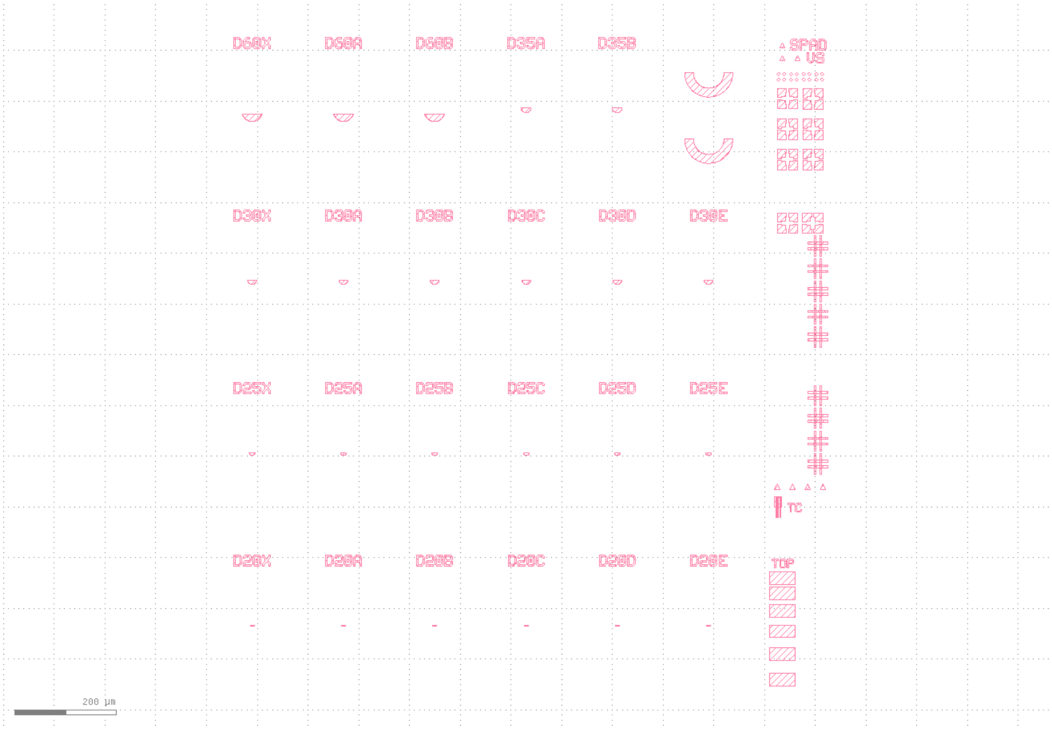


FIGURE B.2: Top contact deposition layer.

B.1.2 Frame 2: Standard Mesa Etch

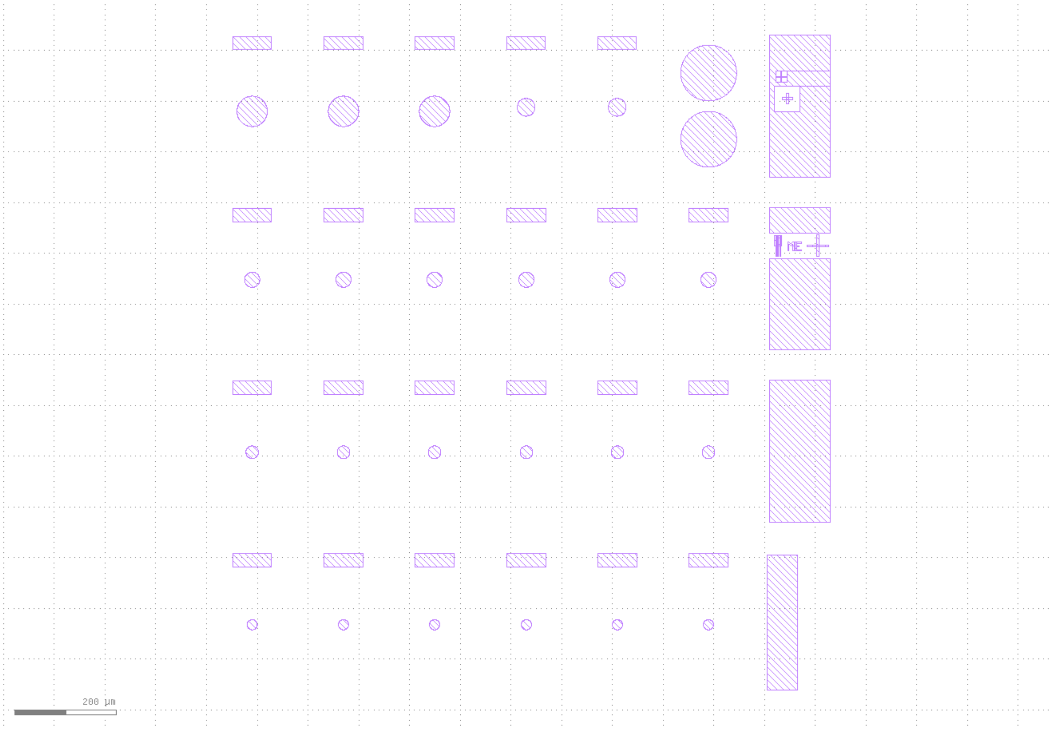


FIGURE B.3: Etching layer to define mesas.

B.1.3 Frame 3: LC Pad

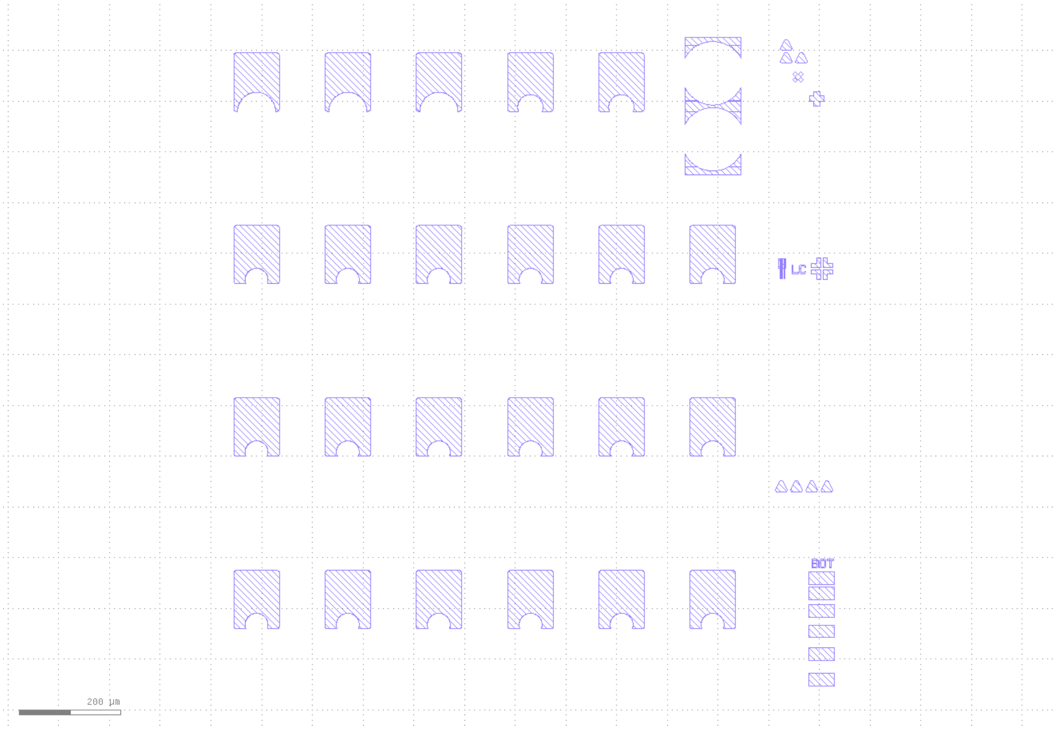


FIGURE B.4: Lower contact bond pad deposition layer.

B.1.4 Frame 4: Isolation Etch

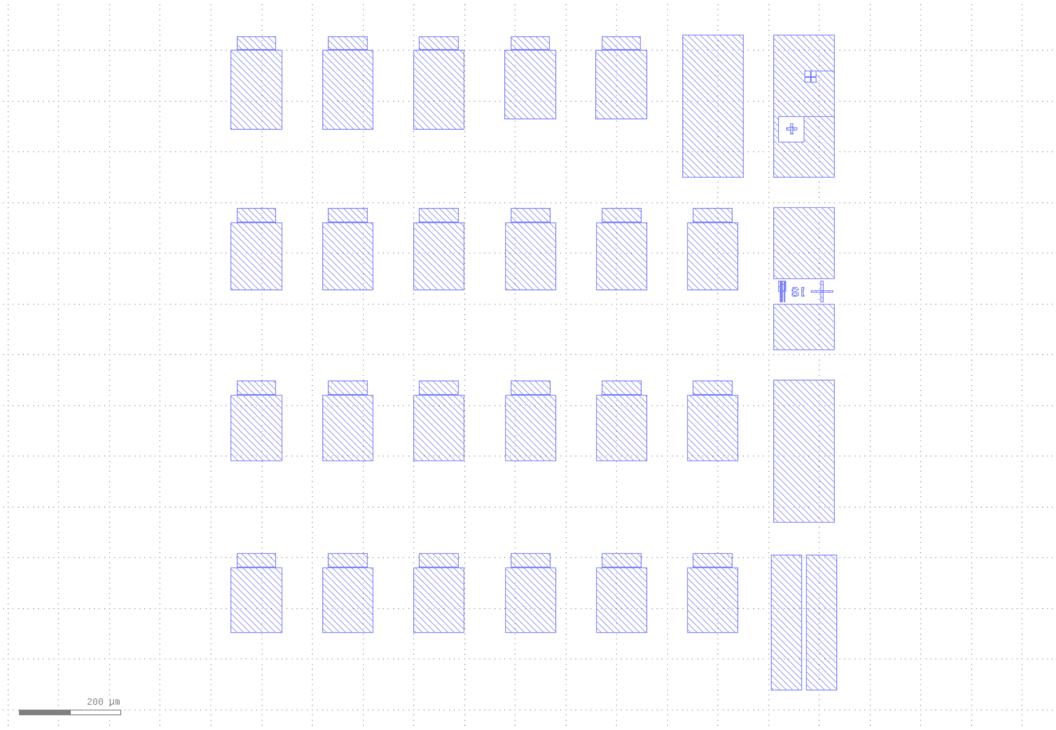


FIGURE B.5: Etching layer to isolate the diodes from each other, and to remove the contact layer from the top bond pad area.

B.2 Mask 2

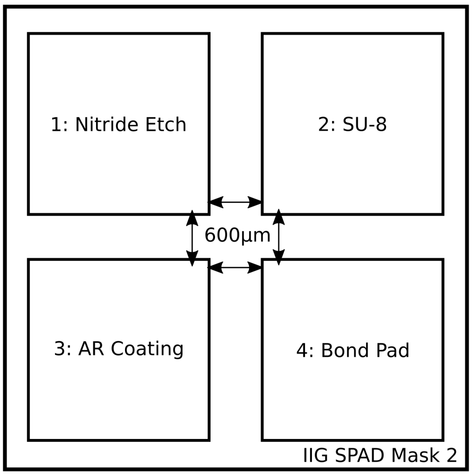


FIGURE B.6: IIG SPAD mask plate 2.

Layer	**Title on mask**	**Description**	**Data Polarity**
5	Nitride Etch	Remove nitride from pads and device window	Clear
6	SU-8	SU-8 passivation	Dark
7	AR coating	Remove nitride from contact areas	Clear
8	Bond pad	Deposit top contact bondpad and optical mask	Clear

TABLE B.2: Overview of IIG mesa mask 2 layers.

B.2.1 Frame 1: Nitride Etch

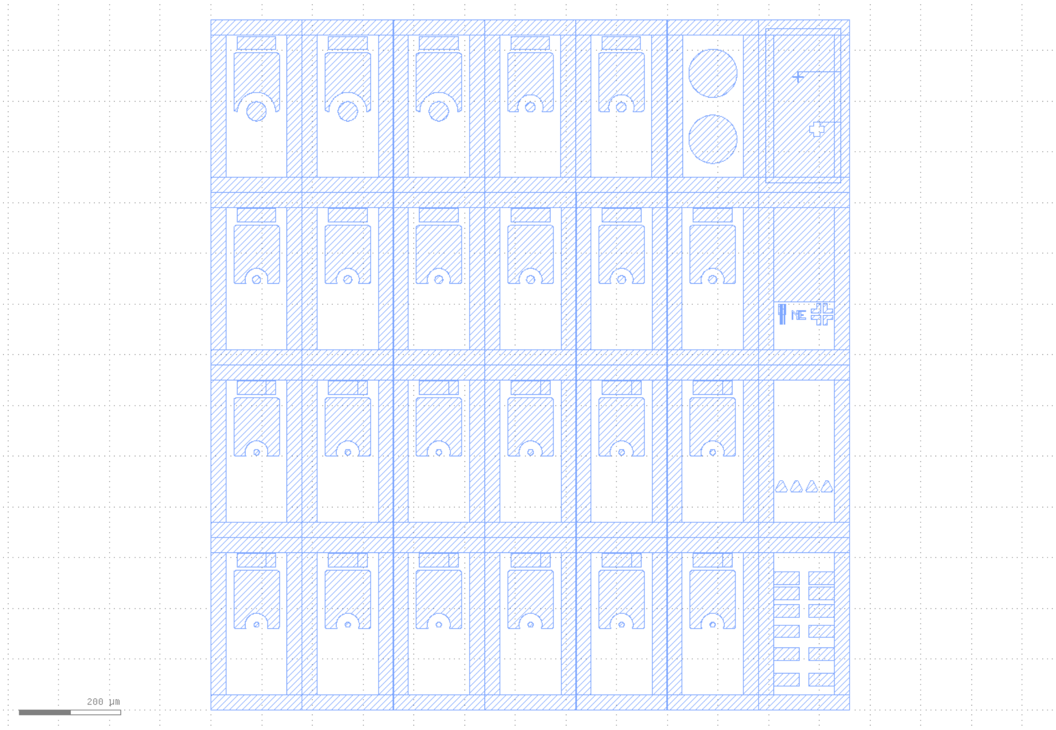


FIGURE B.7: Remove dielectric from device windows, contacts, and between diode cells.

B.2.2 Frame 2: SU-8

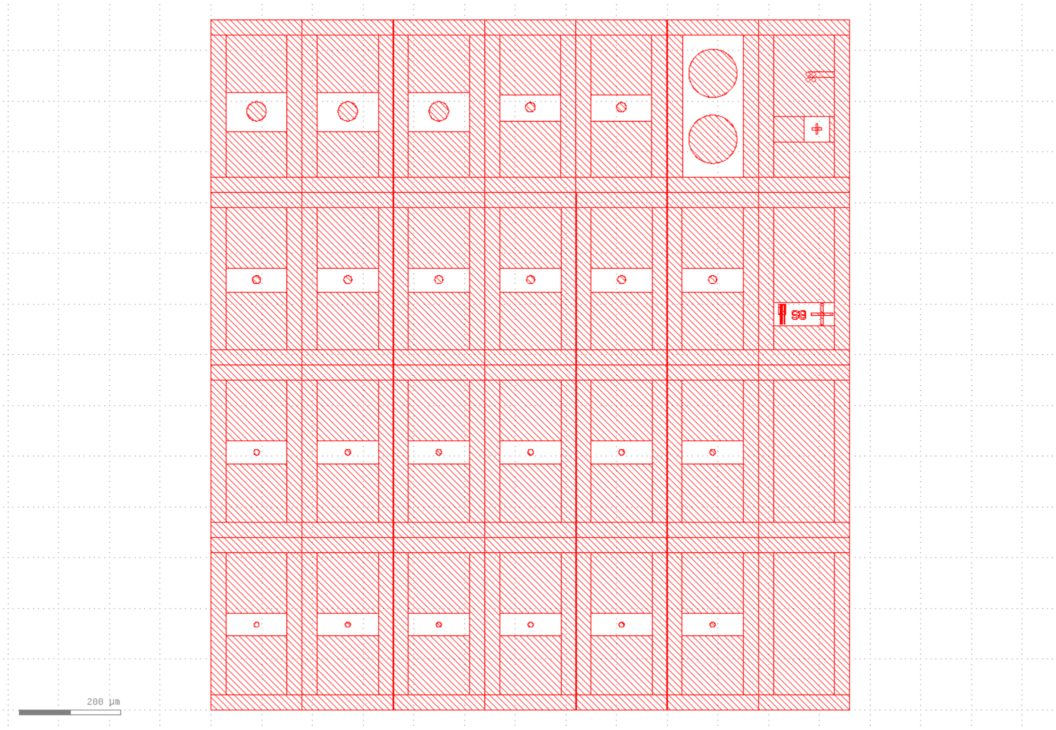


FIGURE B.8: Exposes SU-8, passivating the diode sidewalls.

B.2.3 Frame 3: AR Coating

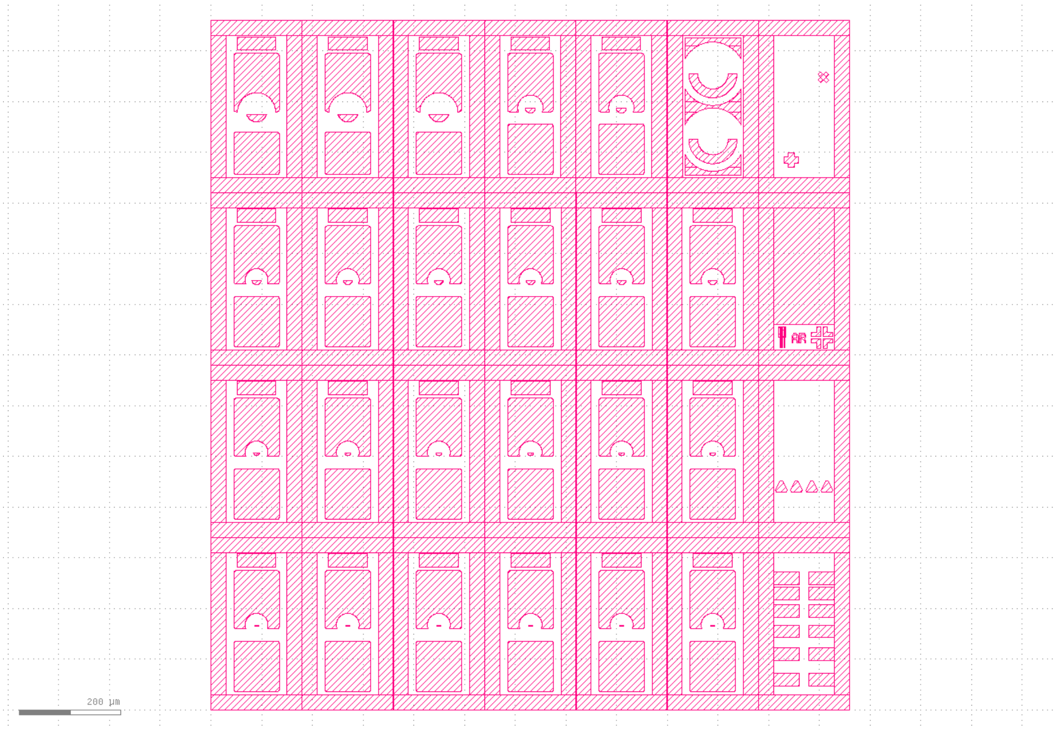


FIGURE B.9: Remove AR coating nitride from all areas other than the mesa windows.

B.2.4 Frame 4: Bond pad

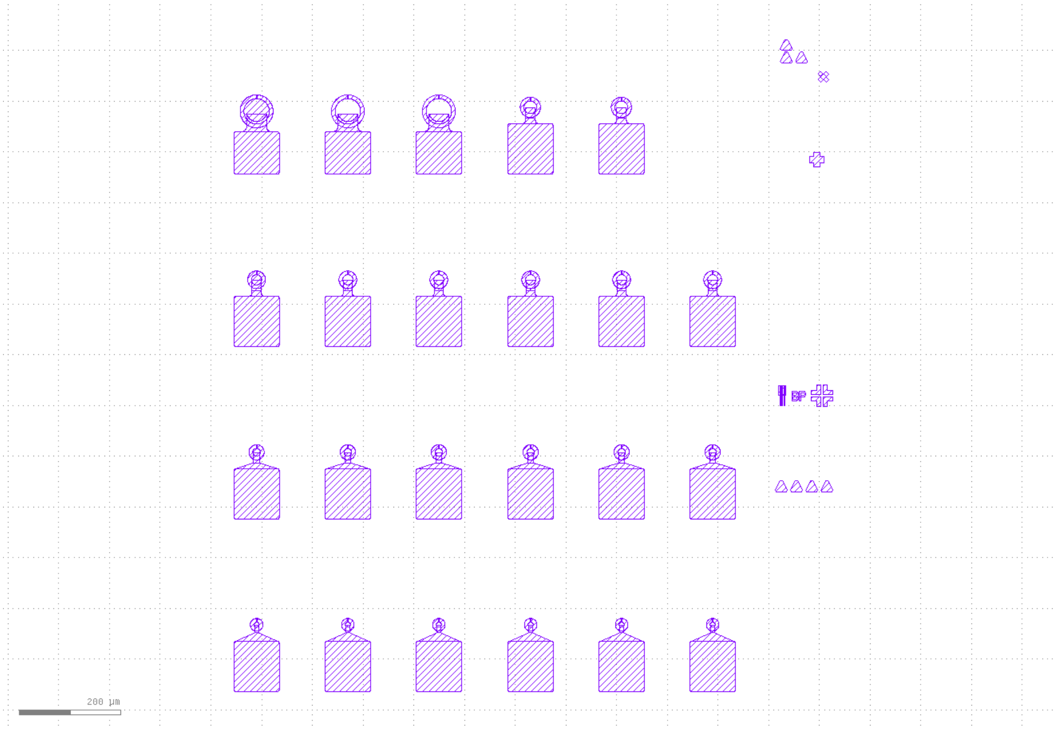


FIGURE B.10: Deposition layer for the top contact bondpad.

References

[1] R. H. Haitz, A. Goetzberger, R. M. Scarlett, and W. Shockley, "Avalanche Effects in Silicon p—n Junctions. I. Localized Photomultiplication Studies on Microplasmas," *Journal of Applied Physics*, vol. 34, no. 6, p. 1581, 1963.

[2] K. McKay, "Avalanche breakdown in silicon," *Physical Review*, vol. 94, no. 1953, p. 877, 1954.

[3] R. L. Batdorf, A. G. Chynoweth, G. C. Dacey, and P. W. Foy, "Uniform silicon p-n junctions. I. Broad area breakdown," *Journal of Applied Physics*, vol. 31, no. 7, pp. 1153–1160, 1960.

[4] A. Goetzberger, B. McDonald, R. H. Haitz, and R. M. Scarlett, "Avalanche Effects in Silicon p—n Junctions. II. Structurally Perfect Junctions," *Journal of Applied Physics*, vol. 34, no. 6, p. 1591, 1963.

[5] K. M. Johnson, "High-speed photodiode signal enhancement at avalanche breakdown voltage," *IEEE Transactions on Electron Devices*, vol. 12, no. 2, pp. 55–63, Feb. 1965.

[6] W. G. Oldham, R. R. Samuelson, and P. Antognetti, "Triggering phenomena in avalanche diodes," *IEEE Transactions on electron devices*, vol. 19, no. 9, pp. 1056–1060, 1972.

[7] R. J. McIntyre, "Theory of microplasma instability in silicon," *Journal of Applied Physics*, vol. 32, no. 6, pp. 983–995, 1961.

[8] R. McIntyre, "Multiplication noise in uniform avalanche diodes," *IEEE Transactions on Electron Devices*, vol. 13, no. 1, pp. 164–168, 1966.

[9] R. J. McIntyre, "The distribution of gains in uniformly multiplying avalanche photodiodes: Theory," *IEEE Transactions on Electron Devices*, vol. 19, no. 6, pp. 703– 713, 1972.

[10] J. Conradi, "The Distribution of Gains in Uniformlv Multiplying Avalanche Photodiodes: Experimental," *IEEE Transactions on Electron Devices*, no. 6, 1972.

[11] I. Wegrzecka *et al.*, “Design and properties of silicon avalanche photodiodes,” *Opto-Electronics Review*, vol. 12, no. 1, pp. 95–104, 2004.

[12] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*, 3rd ed. Wiley, 2007.

[13] J. Wang and S. Lee, “Ge-Photodetectors for Si-Based Optoelectronic Integration,” *Sensors*, vol. 11, no. 1, pp. 696–718, Jan. 2011.

[14] W. Y. Loh *et al.*, “Impact of Local Strain From Selective Epitaxial Germanium With Thin Si/SiGe Buffer on High-Performance p-i-n Photodetectors With a Low Thermal Budget,” *IEEE Electron Device Letters*, vol. 28, no. 11, pp. 984–986, Nov. 2007.

[15] J. Rothman, “Physics and Limitations of HgCdTe APDs: A Review,” *Journal of Electronic Materials*, vol. 47, no. 10, pp. 5657–5665, Oct. 2018.

[16] N. Vanamala, K. C. Santiago, and N. C. Das, “Enhanced MWIR absorption of HgCdTe (MCT) via plasmonic metal oxide nanostructures,” *AIP Advances*, vol. 9, no. 2, p. 025113, Feb. 2019.

[17] E. H. Steenbergen, C. P. Morath, D. Maestas, G. D. Jenkins, and J. Logan, “Comparing II-VI and III-V infrared detectors for space applications,” G. F. Fulop, C. M. Hanson, and B. F. Andresen, Eds. SPIE, May 2019, p. 47.

[18] W. E. Spicer *et al.*, “Band gap variation and lattice, surface, and interface “instabilities” in Hg $_{1-x}$ Cd $_{x}$ Te and related compounds,” *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, vol. 1, no. 3, pp. 1735–1743, Jul. 1983.

[19] E. Palik, *Handbook of Optical Constants of Solids*, 1st ed. Academic Press, 1985.

[20] A. Fox, *Optical Properties of Solids*, ser. Oxford master series in condensed matter physics. Oxford University Press, 2001.

[21] K. Nishida, K. Taguchi, and Y. Matsumoto, “InGaAsP heterostructure avalanche photodiodes with high avalanche gain,” *Applied Physics Letters*, vol. 35, no. 3, pp. 251–253, Aug. 1979.

[22] Y. L. Goh *et al.*, “Avalanche Multiplication in InAlAs,” *IEEE Transactions on Electron Devices*, vol. 54, no. 1, pp. 11–16, Jan. 2007.

[23] J. Campbell, A. Dentai, W. Holden, and B. Kasper, “High-performance avalanche photodiode with separate absorption ‘grading’ and multiplication regions,” *Electronics Letters*, vol. 19, no. 20, pp. 818 –820, 1983.

[24] F. Capasso, A. Cho, and P. Foy, “Low-dark-current low-voltage 1.3–1.6 μm avalanche photodiode with high-low electric field profile and separate absorption and multiplication regions by molecular beam epitaxy,” *Electronics Letters*, vol. 20, no. 15, pp. 635 –637, 1984.

[25] S. Tanaka *et al.*, “Burying type avalanche photodiode and fabrication method thereof,” Oct. 2003.

[26] B. S. White *et al.*, “High-Gain InAs Planar Avalanche Photodiodes,” *Journal of Lightwave Technology*, vol. 34, no. 11, pp. 2639–2644, Jun. 2016.

[27] M. Nada, Y. Muramoto, H. Yokoyama, T. Ishibashi, and H. Matsuzaki, “Triple-mesa Avalanche Photodiode With Inverted P-Down Structure for Reliability and Stability,” *Journal of Lightwave Technology*, vol. 32, no. 8, pp. 1543–1548, Apr. 2014.

[28] W. Lynch, *Method of fabricating a planar avalanche photodiode.* Google Patents, Jun. 1970, uS Patent 3,514,846.

[29] R. Deri and E. Kapon, “Low-loss III-V semiconductor optical waveguides,” *IEEE Journal of Quantum Electronics*, vol. 27, no. 3, pp. 626–640, Mar. 1991.

[30] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, “Quantum cryptography,” *Rev. Mod. Phys.*, vol. 74, no. 1, pp. 145–195, Mar. 2002.

[31] C. I. Rablau, “Lidar: a new self-driving vehicle for introducing optics to broader engineering and non-engineering audiences,” in *Fifteenth Conference on Education and Training in Optics and Photonics: ETOP 2019*, A.-S. Poulin-Girard and J. A. Shaw, Eds. Quebec City, Canada: SPIE, Jul. 2019, p. 138.

[32] H. Iams and B. Salzberg, “The Secondary Emission Phototube,” *Proceedings of the IRE*, vol. 23, no. 1, pp. 55–64, Jan. 1935.

[33] S. Fukuda *et al.*, “The Super-Kamiokande detector,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 501, no. 2-3, pp. 418–462, Apr. 2003.

[34] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, “Theory of Superconductivity,” *Phys. Rev.*, vol. 108, no. 5, pp. 1175–1204, Dec. 1957.

[35] S. P. Chockalingam, M. Chand, J. Jesudasan, V. Tripathi, and P. Raychaudhuri, “Superconducting properties and Hall effect of epitaxial NbN thin films,” *Physical Review B*, vol. 77, no. 21, Jun. 2008.

[36] A. J. Annunziata *et al.*, “Reset dynamics and latching in niobium superconducting nanowire single-photon detectors,” 2010, published: J. Appl. Phys. 108, 084507 (2010).

[37] K. M. Rosfjord *et al.*, “Nanowire Single-photon detector with an integrated optical cavity and anti-reflection coating,” *Opt. Express*, vol. 14, no. 2, pp. 527–534, Jan. 2006.

[38] F. Marsili *et al.*, “Detecting single infrared photons with 93% system efficiency,” *Nature Photonics*, vol. 7, no. 3, pp. 210–214, Mar. 2013.

[39] D. Fukuda *et al.*, “Titanium-based transition-edge photon number resolving detector with 98% detection efficiency with index-matched small-gap fiber coupling,” *Opt. Express*, vol. 19, no. 2, pp. 870–875, Jan. 2011.

[40] D. Rosenberg *et al.*, “Long-Distance Decoy-State Quantum Key Distribution in Optical Fiber,” *Physical Review Letters*, vol. 98, no. 1, Jan. 2007.

[41] A. Giudice *et al.*, “High-rate photon counting and picosecond timing with silicon-SPAD based compact detector modules,” *Journal of Modern Optics*, vol. 54, no. 2-3, pp. 225–237, Jan. 2007.

[42] A. Gulinatti *et al.*, “New silicon SPAD technology for enhanced red-sensitivity, high-resolution timing and system integration,” *Journal of Modern Optics*, vol. 59, no. 17, pp. 1489–1499, Oct. 2012.

[43] Z. Lu *et al.*, “Geiger-Mode Operation of Ge-on-Si Avalanche Photodiodes,” *IEEE Journal of Quantum Electronics*, vol. 47, no. 5, pp. 731–735, May 2011.

[44] R. E. Warburton *et al.*, “Ge-on-Si Single-Photon Avalanche Diode Detectors: Design, Modeling, Fabrication, and Characterization at Wavelengths 1310 and 1550 nm,” *IEEE Transactions on Electron Devices*, vol. 60, no. 11, pp. 3807–3813, Nov. 2013.

[45] M. A. Itzler *et al.*, “Advances in InGaAsP-based avalanche diode single photon detectors,” *Journal of Modern Optics*, vol. 58, no. 3-4, pp. 174–200, Feb. 2011.

[46] A. M. Pawlikowska, A. Halimi, R. A. Lamb, and G. S. Buller, “Single-photon three-dimensional imaging at up to 10 kilometers range,” *Optics Express*, vol. 25, no. 10, p. 11919, May 2017.

[47] L. C. Comandar *et al.*, “Gigahertz-gated InGaAs/InP single-photon detector with detection efficiency exceeding 55% at 1550 nm,” *Journal of Applied Physics*, vol. 117, no. 8, p. 083109, Feb. 2015.

[48] X. Meng *et al.*, “InGaAs/InAlAs single photon avalanche diode for 1550 nm photons,” *Royal Society Open Science*, vol. 3, no. 3, p. 150584, Mar. 2016.

[49] M. N. Ahmed, “Transimpedance Amplifier (TIA) Design for 400 Gb/s Optical Fiber Communications,” Ph.D. dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA, 2013.

[50] J. L. Gimlett, “A new low noise 16 GHz PIN/HEMT optical receiver,” in *1988 Fourteenth European Conference on Optical Communication, ECOC 88 (Conf. Publ. No.292)*, Sep. 1988, pp. 13–16 vol.2.

[51] M. A. R. Violas, D. J. T. Heatley, A. M. O. Duarte, and D. M. Beddow, “10 GHz bandwidth low-noise optical receiver using discrete commercial devices,” *Electronics Letters*, vol. 26, no. 1, pp. 35–36, Jan. 1990.

[52] I. Kimukin *et al.*, “InGaAs-based high-performance p-i-n photodiodes,” *IEEE Photonics Technology Letters*, vol. 14, no. 3, pp. 366–368, Mar. 2002.

[53] T. Ishibashi, S. Kodama, N. Shimizu, and T. Furuta, “High-Speed Response of Uni-Traveling-Carrier Photodiodes,” *Japanese Journal of Applied Physics*, vol. 36, no. 1, pp. 6263–6268, Oct. 1997.

[54] H. Ito, T. Furuta, S. Kodama, and T. Ishibashi, “InP/InGaAs uni-travelling-carrier photodiode with 310 GHz bandwidth,” *Electronics Letters*, vol. 36, no. 21, pp. 1809–1810, Oct. 2000.

[55] Y.-S. Wu, J.-W. Shi, P.-H. Chiu, and W. Lin, “High-Performance Dual-Step Evanescently Coupled Uni-Traveling-Carrier Photodiodes,” *IEEE Photonics Technology Letters*, vol. 19, no. 20, pp. 1682–1684, Oct. 2007.

[56] Q. Cheng, M. Bahadori, M. Glick, S. Rumley, and K. Bergman, “Recent advances in optical technologies for data centers: a review,” *Optica*, vol. 5, no. 11, p. 1354, Nov. 2018.

[57] J. C. Campbell, A. G. Dentai, W. S. Holden, and B. L. Kasper, “High-speed InP/InGaAsP/InGaAs avalanche photodiodes,” in *1983 International Electron Devices Meeting*, Dec. 1983, pp. 464–467.

[58] S. Xie *et al.*, “InGaAs/AlGaAsSb avalanche photodiode with high gain-bandwidth product,” *Optics Express*, vol. 24, no. 21, pp. 24 242–24 247, Oct. 2016.

[59] M. Nada, Y. Muramoto, H. Yokoyama, T. Ishibashi, and S. Kodama, “High-sensitivity 25 Gbit/s avalanche photodiode receiver optical sub-assembly for 40 km transmission,” *Electronics Letters*, vol. 48, no. 13, p. 777, 2012.

[60] M. Nada, H. Yokoyama, Y. Muramoto, T. Ishibashi, and H. Matsuzaki, “A 50-Gbit/s vertical illumination avalanche photodiode for 400-Gbit/s Ethernet systems,” *Optics Express*, vol. 22, no. 12, p. 14681, Jun. 2014.

[61] P. C. Becker, N. A. Olsson, and J. R. Simpson, *Erbium-doped fiber amplifiers: fundamentals and technology*, ser. Optics and photonics. San Diego: Academic Press, 1999.

[62] T. Nagatsuma and H. Ito, “High-Power RF Uni-Traveling-Carrier Photodiodes (UTC-PDs) and Their Applications,” 2011, dOI: 10.5772/14800.

[63] C. Caillaud *et al.*, “Integrated SOA-PIN Detector for High-Speed Short Reach Applications,” *Journal of Lightwave Technology*, vol. 33, no. 8, pp. 1596–1601, Apr. 2015.

[64] C. Caillaud *et al.*, “High sensitivity 40 Gbit/s preamplified SOA-PIN/TIA receiver module for high speed PON,” in *Optical Communication (ECOC), 2014 European Conference on.* IEEE, 2014, pp. 1–3.

[65] P. Angelini *et al.*, “High sensitivity SOA-PIN/TIA photoreceiver for 40 Gb/s applications and beyond,” *International Journal of Microwave and Wireless Technologies*, vol. 8, no. 3, pp. 437–445, May 2016.

[66] R. Borkowski *et al.*, “Real-time burst-mode operation of an integrated SOA-PIN/TIA receiver for 25 Gbit/s/λ and faster T(W)DM-PON,” in *2017 Optical Fiber Communications Conference and Exhibition (OFC)*, Mar. 2017, pp. 1–3.

[67] Mohd Azlishah Othman, Siti Nabilah Taib, Mohd Nor Husain, and Zul Atfyi Fauzan Mohammed Napiah, “Reviews on Avalanche Photodiodes for Optical Communication Technology,” *Journal of Engineering & Applied Sciences*, vol. 9, no. 1, pp. 35–44, 2014.

[68] S. C. Liew Tat Mun *et al.*, “A Theoretical Comparison of the Breakdown Behavior of inalas and InP Near-Infrared Single-Photon Avalanche Photodiodes,” *IEEE Journal of Quantum Electronics*, vol. 45, no. 5, pp. 566–571, May 2009.

[69] A. R. J. Marshall, J. P. R. David, and C. H. Tan, “Impact Ionization in InAs Electron Avalanche Photodiodes,” *IEEE Transactions on Electron Devices*, vol. 57, no. 10, pp. 2631–2638, Oct. 2010.

[70] P. J. Ker, A. R. J. Marshall, A. B. Krysa, J. P. R. David, and C. H. Tan, “Temperature Dependence of Leakage Current in InAs Avalanche Photodiodes,” *IEEE Journal of Quantum Electronics*, vol. 47, no. 8, pp. 1123–1128, Aug. 2011.

[71] R. McIntyre, “On the avalanche initiation probability of avalanche diodes above the breakdown voltage,” *IEEE Transactions on Electron Devices*, vol. 20, no. 7, pp. 637–641, Jul. 1973.

[72] R. J. McIntyre, “A new look at impact ionization-Part I: A theory of gain, noise, breakdown probability, and frequency response,” *IEEE Transactions on Electron Devices*, vol. 46, no. 8, pp. 1623–1631, Aug. 1999.

[73] S. R. Forrest, M. DiDomenico, R. G. Smith, and H. J. Stocker, “Evidence for tunneling in reverse-biased III-V photodetector diodes,” *Applied Physics Letters*, vol. 36, no. 7, pp. 580–582, 1980.

[74] S. J. Dimler, J. S. Ng, R. C. Tozer, G. J. Rees, and J. P. R. David, “Capacitive Quenching Measurement Circuit for Geiger-Mode Avalanche Photodiodes,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 13, no. 4, pp. 919–925, 2007.

[75] R. B. Emmons, “Avalanche-Photodiode Frequency Response,” *Journal of Applied Physics*, vol. 38, no. 9, pp. 3705–3714, Aug. 1967.

[76] S. Adachi, *Properties of semiconductor alloys: group-IV, III-V and II-VI semiconductors*, ser. Wiley series in materials for electronic & optoelectronic applications. Chichester, U.K: Wiley, 2009, oCLC: ocn268957354.

[77] D. K. Gaskill, N. Bottka, L. Aina, and M. Mattingly, “Band-gap determination by photoreflectance of InGaAs and InAlAs lattice matched to InP,” *Applied Physics Letters*, vol. 56, no. 13, pp. 1269–1271, Mar. 1990.

[78] Peng Sun, M. Hayat, B. Saleh, and M. Teich, “Statistical correlation of gain and buildup time in APDs and its effects on receiver performance,” *Journal of Lightwave Technology*, vol. 24, no. 2, pp. 755–768, Feb. 2006.

[79] D. Ong, Jo Shien Ng, M. Hayat, Peng Sun, and J. David, “Optimization of InP APDs for High-Speed Lightwave Systems,” *Journal of Lightwave Technology*, vol. 27, no. 15, pp. 3294–3302, Aug. 2009.

[80] D. S. G. Ong, M. M. Hayat, J. P. R. David, and J. S. Ng, “Sensitivity of High-Speed Lightwave System Receivers Using InAlAs Avalanche Photodiodes,” *IEEE Photonics Technology Letters*, vol. 23, no. 4, pp. 233–235, Feb. 2011.

[81] J. Beck *et al.*, “The HgCdTe electron avalanche photodiode,” *Journal of electronic materials*, vol. 35, no. 6, pp. 1166–1173, 2006.

[82] B. E. A. Saleh, M. M. Hayat, and M. C. Teich, "Effect of dead space on the excess noise factor and time response of avalanche photodiodes," *IEEE Transactions on Electron Devices*, vol. 37, no. 9, pp. 1976–1984, Sep. 1990.

[83] E. Jamil, M. M. Hayat, and G. A. Keeler, "Analytical Formulas for Mean Gain and Excess Noise Factor in InAs Avalanche Photodiodes," *IEEE Transactions on Electron Devices*, vol. 65, no. 2, pp. 610–614, Feb. 2018.

[84] J. S. Ng *et al.*, "Effect of dead space on avalanche speed APDs," *Electron Devices, IEEE Transactions on*, vol. 49, no. 4, pp. 544–549, 2002.

[85] R Piessens, *Quadpack: A subroutine package for automatic integration.* Berlin: Springer-Verlag, 1983.

[86] K. Brennan and K. Hess, "High field transport in GaAs, InP and InAs," *Solid-State Electronics*, vol. 27, no. 4, pp. 347 – 357, 1984.

[87] V. L. Dalal, "Hole velocity in *p* -GaAs," *Applied Physics Letters*, vol. 16, no. 12, pp. 489–491, Jun. 1970.

[88] A. R. Marshall, "The InAs electron avalanche photodiode and the influence of thin avalanche photodiodes on receiver sensitivity," PhD Thesis, University of Sheffield, 2009.

[89] H. Ikeda, T. Ohshima, M. Tsunotani, T. Ichioka, and T. Kimura, "An Auto-Gain Control Transimpedance Amplifier With Low Noise and Wide Input Dynamic Range for 10-Gb/s Optical Communication Systems," *IEEE Journal of Solid-State Circuits*, vol. 36, no. 9, pp. 1303–1308, Sep. 2001.

[90] ADSANTEC, *ASNT6123 25 Gbps Dual Transimpedance Aplifier*, Jan. 2010.

[91] S.-T. Chou, S.-H. Huang, Z.-H. Hon, and W.-Z. Chen, "A 40 Gbps optical receiver analog front-end in 65 nm CMOS," *IEEE International Symposium on Circuits and Systems*, May 2012.

[92] K. Shiba, T. Nakata, T. Takeuchi, T. Sasaki, and K. Makita, "10 Gbit/s asymmetric waveguide APD with high sensitivity of- 30 dBm," *Electronics Letters*, vol. 42, no. 20, pp. 1177–1178, 2006.

[93] Z. Huang *et al.*, "25 Gbps low-voltage waveguide Si–Ge avalanche photodiode," *Optica*, vol. 3, no. 8, p. 793, Aug. 2016.

[94] P. Angelini *et al.*, "Record -22.5-dBm Sensitivity SOA-PIN-TIA Photoreceiver Module for 40-Gb/s Applications," *IEEE Photonics Technology Letters*, vol. 27, no. 19, pp. 2027–2030, Oct. 2015.

[95] S. Takashima *et al.*, “40-Gbit/s receiver with -21 dBm sensitivity employing filterless semiconductor optical amplifier,” in *OFC 2003 Optical Fiber Communications Conference, 2003.*, Mar. 2003, pp. 471–472 vol.2.

[96] SiFotonics, “AP2005-12m3,” 2019. [Online]. Available: http://www.sifotonics.com/pd.jsp?id=35#_jcp=2 [Accessed: 2019-06-26]

[97] N. Yasuoka, H. Kuwatsuka, and A. Kuramata, “High-speed and high-efficiency InP/InGaAs waveguide avalanche photodiodes for 40 Gbit/s transmission systems.” IEEE, Feb. 2004.

[98] L. J. J. Tan, J. S. Ng, C. H. Tan, and J. P. R. David, “Avalanche Noise Characteristics in Submicron InP Diodes,” *IEEE Journal of Quantum Electronics*, vol. 44, no. 4, pp. 378–382, Apr. 2008.

[99] Y. L. Goh *et al.*, “Excess Avalanche Noise in inalas,” *IEEE Journal of Quantum Electronics*, vol. 43, no. 6, pp. 503–507, Jun. 2007.

[100] R. Quay, C. Moglestue, V. Palankovski, and S. Selberherr, “A Temperature Dependent Model for the Saturation Velocity in Semiconductor Materials,” vol. 3, no. 1-2, pp. 149–155, Mar. 2000.

[101] S. Adachi, *GaAs and Related Materials Bulk Semiconducting and Superlattice Properties.* World Scientific, Oct. 1994.

[102] A. R. J. Marshall, P. Vines, P. J. Ker, J. P. R. David, and C. H. Tan, “Avalanche Multiplication and Excess Noise in InAs Electron Avalanche Photodiodes at 77 K,” *IEEE Journal of Quantum Electronics*, vol. 47, no. 6, pp. 858–864, Jun. 2011.