Design and device fabrication of Silicon Single Photon Avalanche Diodes

Jonathan D. Petticrew

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

Department of Electronic and Electrical Engineering

The University of Sheffield

July 2020

Contents

Glossary v

List of Figures xi

List of Tables xv

Acknowledgements xvii

Abstract xviii

List of Publications xix

1 Introduction 1

1.1 Applications of Optical Detectors 1

1.1.1 LiDAR 1

1.1.2 QKD 2

1.1.3 TCSPC 3

1.2 Types of Optical Detectors 6

1.2.1 PMTs 6

1.2.2 MCP 7

1.2.3 APDs 7

1.2.4 SPADs 8

1.2.5 SiPMs 9

1.2.6 Nanowires 9

1.3 Thesis Outline 10

2 Background 11

2.1 APDs 11

2.1.1 Impact Ionisation 11

2.1.2 Performance Parameters 14

2.2 SPADs 14

2.2.1 Avalanche Breakdown 14

2.2.2 Equivalent Circuit 15

2.2.3 Quenching Circuitry 16

2.2.4 Performance Parameters 17

2.2.5 Silicon 21

2.2.6 InGaAs/InP 23

2.3 APD & SPAD Models 25

2.3.1 Local Model 25

2.3.2 Recurrence Equations 26

2.3.3 Random Path Length Model 27

2.3.4 Full Band Monte Carlo 28

2.3.5 Analytical Band Monte Carlo 28

2.3.6 Simple Monte Carlo 28

3 Experimental Details 31

3.1 Current-Voltage Measurements 31
3.2 Capacitance-Voltage Measurements 33
3.3 Avalanche Multiplication Measurements 33
3.4 Dark Count Rate Measurements 34
3.5 Photon Counting 37
3.5.1 Calculating SPDE 38
3.6 Silicon Device Fabrication 41
3.6.1 Wafer cleaning 41
3.6.2 Photolithography 42
3.6.3 Ion Implantation 44
3.6.4 Oxidation 44
3.6.5 Dry Etching 45
3.6.6 Metal Deposition 47
3.6.7 Secondary Ion Mass Spectrometry 48

4 A Simple Monte Carlo Model for Silicon SPADs 49

4.1 Introduction & existing State of SMC 49
4.2 SMC Simulation flow 50
4.3 Expansion of SMC model for SPAD simulation 53
4.4 Model Verification with Experimental Results 54
4.5 Effective Ionisation Coefficients 60
4.6 Simulated SPAD Characteristics 61
4.6.1 Breakdown Probability 62
4.6.2 Timing characteristics 62
4.6.3 P-on-N vs N-on-P 63
4.7 Chapter summary 66

5 InP SMC 67

5.1 Introduction 67
5.1.1 Motivation 67
5.2 Temperature Dependent InP SMC 68
5.3 Temperature Dependent Multiplication 73
5.4 Temperature Dependent Impact Ionisation Coefficients 74
5.5 Chapter summary 77

6 Mesa SPADs 79

6.1 Sample Fabrication 79
6.2 I-V 81
6.3 C-V 83
6.4 Gain 84
6.5 DCR 84
6.6 Photon Counting 86
6.7 Chapter Summary 88

7 Conclusion & Future Work 89

7.1 Conclusion 89
7.2 Future Work 90

References 92

A The Simple Monte Carlo Simulator 105

A.1 Architecture 106

B Electric Field Solver **107**
B.1 Calculating Simple Electric Fields 107
B.1.1 2 Regions 107
B.1.2 3 Regions 108
B.2 Infinite Electric Field Solver 109
B.2.1 PN junction finder 110
B.2.2 Find the number of regions 110
B.2.3 Finding the profile 111
C Obtaining Diffusion constant from $S(x,t)$ **113**
D Fitting InP PDFs **115**
E Mesa Fabrication Process **125**
E.1 Implantation 125
E.2 Sample Fabrication 126

Glossary

List of Abbreviations

ABMC .... Analytical Band Monte Carlo

APD .... Avalanche Photodiode

CCD .... Charge-coupled device

CMOS .... Complimentary Metal-Oxide-Semiconductor

CQC .... Capacitive Quenching Circuit

DCR .... Dark Count Rate

DOT .... Diffused Optical Tomography

DUT .... Device Under Test

FBMC .... Full-Band Monte Carlo

FLIM .... Fluorescence Lifetime Imaging Microscopy

FWHM .... Full-Width at Half-Maximum

HPMT .... Hybrid-Photomultiplier Tube

IPA .... Isopropyl Alcohol

LADAR .... Laser Detection and Ranging

LIA .... Lock-In Amplifier

LO Phonon .... Longitudinal Optical Phonon

LSA .... Loughborough Surface Analysis

LiDAR .... Light Detection and Ranging

MCP .... Microchannel plate

PDE .... Photon Detection Efficiency

PDF .... Probability Density Function

PMT .... Photomultiplier Tube

QKD .... Quantum Key Distribution

QSS .... Quantum Secret Sharing

RCA .... Radio Corporation of America

RE-SPAD .... Red Enhanced SPAD

RPL .... Random Path Length

SAM-SPAD .... Separate Absorption Multiplication-SPAD

SC-1 .... Standard Clean 1

SC-2 .... Standard Clean 2

SEM .... Scanning Electron Microscope

SIMS .... Secondary Ion Mass Spectrometry

SiPM .... Silicon Photomultiplier

SMC .... Simple Monte Carlo

SMU .... Source Measure Unit

SNR .... Signal to Noise Ratio

SNSPD .... Superconducting nanowire single-photon detector

SPAD .... Single photon Avalanche Diode

SRIM .... Stopping Ranges of Ions in Matter

TCSPC .... Time Correlated Single Photon Counting

TO Phonon .... Transverse Optical Phonon

TTS .... Transit-time spread

List of Symbols

α .... Electron impact ionisation coefficient

α^* .... Effective electron impact ionisation coefficient

β .... Hole impact ionisation coefficient

β^* .... Effective hole impact ionisation coefficient

c .... Speed of Light

C_{ii} .... Coefficient of impact ionisation

d .... Average gain of PMT dynode

D .... Diffusion coefficient

d^n .... Largest PMT gain

d_e .... Electron deadspace

d_h .... Hole deadspace

E .... Electric field strength

E_c .... Carrier energy

E_g .... Band Gap

E_{photon} .... Energy of photon

E_{th} .... Impact ionisation threshold energy

F .... Excess noise factor

F_{fill} .... Fill factor

γ .... softness factor of impact ionisation

Ω .... Phonon Energy

h .... Plancks Constant

$\hbar$ .... Reduced Plancks Constant ($\frac{h}{2\pi}$)

$\hbar\omega$ .... Phonon energy

$h_e(\xi)$ .... Electron impact ionisation probabilities

$h_h(\xi)$ .... Hole impact ionisation probabilities

$I_{anti-stokes}$ .... Intensity of anti-stokes scattered photons

I_{stokes} .... Intensity of stokes scattered photons

j .... Particle flux

k_b .... Boltzmann's constant

l .... Random path length

λ .... Wavelength

λ .... Mean free path

λ_0 .... Excitation wavelength

λ_I .... Emission wavelength

L .... Distance to object

M .... Avalanche gain (Mean)

m .... Avalanche gain from single event

m^* .... Effective carrier mass

M_e .... Pure electron avalanche gain

M_h .... Pure hole avalanche gain

N .... Phonon occupation factor

η .... External quantum efficiency

n .... Number of dynodes

n .... Number of carriers

ξ .... Impact ionisation path length

P_b .... Breakdown Probability

PDE_{array} .... PDE of an array.

q .... Electron charge

r .... Random number between 0 and 1

R_{ab} .... Rate of phonon absorption

R_{btb} .... Rate of Band to Band tunnelling

R_{em} .... Rate of phonon emission

R_{ii} .... Rate of impact ionisation

ΔT .... Time between optical emission and peak of detection signal

Δt .... Time taken for carrier to drift Δx

T .... Temperature of sample

V_b .... Breakdown voltage

ν .... Mean carrier velocity

$W_{forbidden}$ .... Width of forbidden region

w .... Depletion width

x .... Absorption position

Δx .... Minimum drift distance

$Y(x)$ .... Number of generated holes, used in Recurrence Equations

$Y_2(x)$ .... Mean squared of $Y(x)$

$Z(x)$ .... Number of generated electrons, used in Recurrence Equations

$Z_2(x)$ .... Mean squared of $Z(x)$

List of Figures

1.1 Schematic diagram of a PMT. The green arrows denote the electron beam path and the thickness denotes the beam intensity. 6

1.2 Schematic diagram of a HPMT where the width of the green region denotes the electron multiplication process. 7

1.3 Schematic diagram of a MCP. 7

2.1 Schematic of avalanche gain from a single free electron. 12

2.2 Example of hard dead space approximation (line) to simple monte carlo generated (symbols) impact ionisation probability densities, for holes in an 600 kV.cm^{-1} electric field, in InP at room temperature. 13

2.3 Example $M(V)$ indicating linear and Geiger mode operation with respect to the breakdown voltage (V_b) 15

2.4 Equivalent Circuit of a SPAD [44]. Where V_d, R_d, C_d represent the diode voltage, the resistance across the SPAD and the SPADs junction capacitance respectively. V_a, R_q represent the applied bias and the quenching resistor. 16

2.5 Schematic showing Band-to-Band tunnelling and Trap Assisted tunnelling mechanisms 19

2.6 Image showing a 128 pixel test sample with 256 aluminium wire bonds. . . 22

2.7 A schematic showing the structure of an $In_{0.53}Ga_{0.47}As$/InP SAM-SPAD where the orders of magnitude represent typical doping concentrations in dopants.cm^{-3} (top) along with the corresponding electric field profile (bottom). 24

3.1 Schematic of Avalanche multiplication setup. Light from a Helium-Neon (HeNe) laser is chopped at 180 Hz by an Lock-In Amplifier (LIA) controlled optical chopper before being focused, through an objective mounted in an objective turret, on to a reverse biased APD. The diode is reverse biased by an SMU and the LIA is able to measure a voltage proportional to avalanche gain across the sense resistor (typically 100 Ω). 34

3.2 Schematic showing the presence of capacitive transients on the SPAD signal (dashed line) from an applied electrical pulse (solid line). 35

3.3 Schematic of electrical connections for dark and photon counting measurements using the ST-500 Janis probe station. 36

3.4 Schematic showing the structure of HMDS (top) then showing how HMDS acts as an adhesion promoter on silicon oxide and silicon nitride. 43

3.5 Example SEM image, showing the side wall profile of dry etched Si 46

3.6 Example SEM image, showing surface roughness responsible for black silicon. . . . 46

4.1 Flow chart of device properties function for the Simple Monte Carlo Simulator . . . 51
4.2 Comparison of reported drift velocities (lines) Canali et al. [150], A.C. Prior et al. [151], Norris and Gibbons [152] to drift velocities simulated by the SMC model (symbols). . . . 55
4.3 Example $S(x,t)$ for 500000 electrons in a 200 kV.cm^{-1} field at different times after the initial carrier injection. . . . 55
4.4 Comparison between simulated SMC diffusion coefficients (red dots) and reported diffusion coefficients (black dots) for electrons (top) from Canali et al. [153] and holes (bottom) from Nava et al. [154]. . . . 57
4.5 Comparison of reported electron and hole ionisation coefficients (solid and dashed lines) from Massey et al. [155, 156] and Overstraeten and Man [157] to SMC simulations (symbols). . . . 58
4.6 Comparison of experimental (lines) [158] and simulated results (symbols) of $M(V)$ (top) and $F(M)$ (bottom) for Si diodes. . . . 59
4.7 Example of extraction of the experimental breakdown voltage by plotting the inverse gain versus reverse bias. The experimental data used for device A gave a breakdown voltage of 7 V. . . . 60
4.8 SMC simulated breakdown probabilities versus reverse bias characteristics for Si devices A-E (symbols). The results are consistent with the extracted V_b values in Table 4.3. . . . 60
4.9 Comparison of breakdown probability versus reverse bias characteristics for an ideal 1.0 µm p-i-n diode[3] from the SMC model (symbols) and recurrence equations (lines) [144] using the extracted $\alpha^*(E)$ and $\beta^*(E)$ (Equations (4.8) and (4.9).) . . . 62
4.10 Comparison between RPL simulated (lines) [143] and SMC simulated (symbols) mean time to breakdown (top) and standard deviation of time to breakdown (bottom) for a 1 µm ideal p-i-n diode. . . . 63
4.11 Electric field profiles of n-on-p and p-on-n designs at their breakdown voltage of 27 V. . . . 64
4.12 Comparison of simulated breakdown probability between the n-on-p and p-on-n designs . . . 64
4.13 Comparison between mean time to breakdown (top) and timing jitter, FWHM, (bottom) for the n-on-p and p-on-n designs. . . . 65

5.1 Comparison between published room temperature saturation velocities and room temperature drift velocities generated by the SMC. Windhorn et al.[174], Brennan and Hess [175]. . . . 69
5.2 Comparison between published room temperature ionisation coefficients and room temperature ionisation coefficients generated by the SMC. Cook et al. [167], Saleh et al. [168], and Tan et al. [169] . . . 70
5.3 Example C-V fit for device D with a radius of 200 µm [176]. . . . 71

5.4 Comparison of SMC and experimental $M(V)$ and $F(M)$ characteristics for InP at room temperature. 72
5.5 Temperature dependent gain of InP devices C-G, however device D was not measured at 100 K. Experimental results (symbols) from [98] 73
5.6 Examples of probability densities (symbols) fitted using the hard deadspace approximation (lines) at 290 K and 700 kV.cm^{-1} for electrons (top) and holes (bottom). 75
5.7 Simulated temperature gain, using the SMC generated effective ionisation coefficients in a recurrence model [102], compared to experimental gain results [98]. 76
5.8 Simulated room temperature excess noise factors, using the SMC generated effective ionisation coefficients in a recurrence model [102], compared to experimental results [169]. 77

6.1 Schematic of wafer epitaxial structure (left) and after ion implantation and annealing (right). 80
6.2 Microscope images of sample A (left) and sample B (right) 80
6.3 I-V results of sample A. Radius of diodes denoted by R in legend. Forward biased (left) and reversed biased (right). 82
6.4 I-V results of sample B diodes forward biased (left) and reversed biased (right). 82
6.5 (Top) Fit (symbols) used to calculate a series resistance of 120 kΩ from the forward bias of the sample B mesa devices, and the obtained ideality factor (bottom) 82
6.6 C-V data (symbols) and fitted (line) of 200 µm radius mesa from sample A. The fit used the adjusted doping profile shown in Figure 6.7. 83
6.7 Si doping profiles originally simulated in Sentaurus TCAD (solid) and adjusted to fit C-V data (dashed). 83
6.8 Gain data from 3 devices from sample B using 633 nm illumination. (Left) shows uniform gain to $M = 60$ with the addition of SMC simulated M using the simulated doping profile, whilst (Right) shows that the largest recorded M is 283. 84
6.9 Temperature stability of dark count measurements at 25.1 V with an additional 3 V pulse 85
6.10 0-40% DCR for 3 diodes from sample B. 86
6.11 Schematic showing optical paths for photon counting setup 86
6.12 Photon and Dark counts per second from 3 devices from sample B, when operated at 100 kHz, with an incoming optical power of 0.1 photons per pulse at 280 K. 87
6.13 SPDE from 3 devices from sample B, when operated at 100 kHz, with 0.1 photons per pulse for photon counting at 280 K. 87

A.1 Schematic of class dependencies of the Simple Monte Carlo Simulator . . . 106

B.1 Schematic of PN diode 108

B.2 Schematic of PIN diode 108
B.3 Flow chart outlining simplified process for the infinite electric field solver . 110

D.1 Example of probability density plot fitted using a linear y-axis (290 K holes at 600 kV.cm^{-1}) 116
D.2 PDFs of electron at 290 K in InP 117
D.3 PDFs of hole at 290 K in InP 118
D.4 PDFs of electron at 250 K in InP 119
D.5 PDFs of hole at 250 K in InP 120
D.6 PDFs of electron at 200 K in InP 121
D.7 PDFs of hole at 200 K in InP 122
D.8 PDFs of electron at 150 K in InP 123
D.9 PDFs of hole at 150 K in InP 124

List of Tables

2.1 Comparison of published Si SPAD performance 23

4.1 Table showing random number values required to trigger the interaction events from the SMC, where P_{abs}, P_{em}, and P_{ii} represent the probabilities of phonon absorption, phonon emission, and impact ionisation. 53
4.2 Diffusion coefficient values generated from the SMC model. 57
4.3 Si Devices A-E [158] used to validate $M(V)$ and $F(M)$ in Si SMC model. . 58
4.4 Structure of P-on-N and N-on-P devices. 63

5.1 InP SMC Parameters 69
5.2 InP device structures 71
5.3 α^* and β^* for InP at 150, 200, 250, and 290 K extracted from PDFs generated from the InP SMC 75

A.1 Simple Monte Carlo Simulator Class Descriptions 106

Acknowledgements

I wish to thank Prof. Jo Shien Ng and Prof. Chee Hing Tan for their continued guidance and supervision throughout my PhD. Special thanks is also due to Dr Simon Dimler, without his help my PhD project would have been incredibly more challenging!

I would like to thank all the current and past members of IIG, specially Vladimir Shulyak, Leh Woon Lim, Elizabeth Stark, Jonathan Taylor-Mew, Dr Lucas Pinel, Dr Salman Abdullah, Dr Benjamin White, Dr Xinxin Zhou, Dr Ian Sandall and Dr Xiao Meng.

I should also thank my colleagues and friends from the cleanroom facilities within the university, specially Dr Benjamin Royall, Dr Kenneth Kennedy, Saurabh Kumar, Dr. Timmothy Lewis-Roberts, Dr Kean Boon Lee and Dr. Rob Airey for the countless bits of advice and teaching me everything I know about device fabrication.

The support from the EPSRC in funding my PhD scholarship and placement at Faraday Scientific Ltd. is greatly appreciated.

Finally a special thanks to all my friends and family who have given me nothing but support over the duration of my studies.

Abstract

Silicon Single Photon Avalanche Diodes (SPADs) have become increasingly important due to a rise in applications requiring very sensitive, low level light detectors.

This thesis focuses on the development of a simple monte carlo simulator for the modelling of Si SPADs, along with the fabrication of a Si mesa SPAD. The simulator was validated against experimental and reported Si results. Simulations are performed to compare an n-on-p to a p-on-n SPAD design. These simulations find the n-on-p design offers better timing performance for a given breakdown probability, however the p-on-n design achieves a greater breakdown probability for a given bias.

A new temperature-dependent simple monte carlo parameter set is presented for InP APDs. This parameter set is extensively validated from 150-290 K, showing that the simulator is capable of temperature dependent modelling.

Finally, a Si mesa SPAD is demonstrated. This mesa SPAD suffers from a high dark count rate, however is still capable of achieving a 69% single photon detection efficiency at 633 nm when operated at 280 K.

Follow on work from this thesis could include further development of the simulator to add the simulation of external quenching mechanisms and the validation of the InP parameter set for Geiger-mode simulation. Fabrication of a planar Si SPAD using the same active device structure would allow for the direct comparison of dark current contributions due to the etching process.

List of Publications

Journal Papers

4. **J.D. Petticrew**, S.J. Dimler, C.H. Tan, J.S. Ng,"Modeling Temperature dependent avalanche characteristics of InP", IEEE Journal of Lightwave Technology, vol. 38, no. 4, February 2020. DOI: 10.1109/JLT.2019.2948072

3. L.W. Lim, C.H. Tan, J.S. Ng, **J.D. Petticrew**, A. Krysa, "Improved planar InAs avalanche photodiodes with reduced dark current and increased responsivity", IEEE Journal of Lightwave Technology, vol. 37, no. 10, pp. 2375-2379, May 2019. DOI: 10.1109/JLT.2019.2905535

2. **J.D. Petticrew**, S.J. Dimler, J.S.Ng, "Simple Monte Carlo Simulator for modelling Avalanche Photodiode & Single Photon Avalanche Photodiode Characteristics in C++", Journal of Open Research Software, vol. 6, no. 1, 17, May 2018. DOI: 10.5334/jors.212

1. **J.D. Petticrew**, S.J. Dimler, X. Zhou, A.P. Morrison, C.H. Tan, J.S. Ng, "Avalanche Breakdown Timing Statistics for Silicon Single Photon Avalanche Diodes", IEEE Journal of Selected Topics on Quantum Electronics, vol. 24, no. 2, 3801506, March 2018. DOI: 10.1109/JSTQE.2017.2779834

Conferences

9. L.W. Lim, **J. Petticrew**, A. Krysa, J.S. Ng, C.H. Tan, "Investigation of low energy Be Implantation and annealing conditions in InAs avalanche photodiodes", 44th Freiburg Infrared Colloquium, Freiburg, Germany, March 2019.

8. **J.D. Petticrew**, S.J. Dimler, C.H. Tan, J.S.Ng, "Modelling Timing Performance of InP Single Photon Avalanche Diodes", Photonics by the Lake, Windermere, UK, July 2018.

7. 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 AlGaAsSb avalanche photodiodes", Photonics by the Lake, Windermere, UK, July 2018.

6. S. Abdullah, S.J. Dimler, J.S. Ng, V. Shulyak, **J.D. Petticrew**, C.H. Tan, "Dark Counts and Their Timing Statistics in AlGaAsSb Single Photon Avalanche Diodes", UK Semiconductors, Sheffield, UK, July 2018.

5. **J. Petticrew**, S.J. Dimler, J.S.Ng, "Avalanche Gain & Excess Noise Factor Model for InP Avalanche Photodiodes", UK Semiconductors, Sheffield, UK, July 2018.

4. **J.D. Petticrew**, S.J. Dimler, J.S.Ng, "Modelling Timing Jitter in Silicon Single Photon Avalanche Diodes", SIOE, Cardiff, UK, March 2018.

3. **J. Petticrew**, S. Dimler, X. Zhou, A.P. Morrison, C.H. Tan, J.S. Ng, "Breakdown probability & timing simulation model for silicon single photon avalanche diodes", UK Semiconductors, Sheffield, UK, July 2017.

2. L.W. Lim, **J. Petticrew**, A. Krysa, C.H. Tan, "Investigation of low energy Be implantation in InAs photodiodes", UK Semiconductors, Sheffield, UK, July 2017.

1. L. Pinel, **J. Petticrew**, X. Zhou, S. Abdullah, C.H. Tan, J.S. Ng. "Improving Wet Etching of InGaAs/AlGaAsSb Avalanche Photodiodes", 19th International Conference on Molecular Beam Epitaxy, Montpellier, France, September 2016

Open Data & Open Source

4. (a) **J.D. Petticrew**, S.J. Dimler, C.H. Tan, J.S. Ng, "Data for Paper: Modeling Temperature dependent Avalanche Characteristics of InP", Figshare, 2019. DOI:10.15131/shef.data.7611380

   (b) **J.D. Petticrew**, S.J. Dimler, C.H. Tan, J.S. Ng, "Software for Paper: Modeling Temperature dependent Avalanche Characteristics of InP", Figshare, 2019. DOI:10.15131/shef.data.7611377

3. L.W. Lim, C.H. Tan, J.S. Ng, **J.D. Petticrew**, A. Krysa, "Improved planar InAs avalanche photodiodes with reduced dark current and increased responsivity", Figshare, 2019. DOI:10.15131/shef.data.6955037

2. (a) **J.D. Petticrew**, S.J. Dimler, J.S. Ng, "Simple Monte Carlo Simulator v1.0.0", Figshare, 2018. DOI: 10.15131/shef.data.5683939

   (b) **J.D. Petticrew**, S.J. Dimler, J.S. Ng, "Simple Monte Carlo Simulator", GitHub Repository, 2017. https://github.com/jdpetticrew/Simple-Monte-Carlo-Simulator

1. **J.D. Petticrew**, S.J. Dimler, X. Zhou, A.P. Morrison, C.H. Tan, J.S. Ng, "Data for Paper: Avalanche Breakdown Timing Statistics for Silicon Single Photon Avalanche Diodes", Figshare, 2017. DOI: 10.15131/shef.data.4823248

Chapter 1

Introduction

Optical detectors are devices designed to convert light signals into electrical signals. Their operation may be based upon the photovoltaic and photoelectric effects. The latter was explained by Einstein in 1905 [1] using wave-particle duality to treat light as both a wave and a particle. These particles, called photons, have energy (E_{photon}) described using,

$$E_{photon} = \frac{hc}{\lambda}, \tag{1.1}$$

where h is Plancks constant, c is the speed of light and λ is the wavelength of the photon. This chapter will briefly introduce several key applications of optical detectors and describe some currently available optical detector technologies before outlining the structure of this thesis. Though this chapter outlines a range of available detector technologies the focus of this thesis is highly sensitive Avalanche Photodiodes (APDs) and Single Photon Avalanche Diodes (SPADs).

1.1 Applications of Optical Detectors

1.1.1 LiDAR

Light detection and ranging (LiDAR), sometimes referred to as Laser detection and ranging (LADAR), has seen an increased popularity and increased research activity mainly for the advancement of self-driving cars. LiDAR systems are optical systems incorporating an optical emitter (usually a pulsed laser) and an optical detector. LiDAR operates using the principle of time of flight, which is an established technique for laser ranging. By measuring the time between the outgoing optical emission and the peak of the detected

return signal, ΔT, it is possible to calculate the distance to an object, L, in the direction the system is pointed, using

$$L = \frac{c\Delta T}{2}.$$

There is variation in ΔT due to timing jitter from several sources; including the pulsed laser, the detector, and atmospheric effects. For the detector, timing jitter is the variation in time lapsed between light being absorbed by the detector and a measurable electrical signal being produced. The timing jitter of a detector is often quoted as the Full Width at Half Maximum (FWHM) of multiple detection events.

Current LiDAR systems mainly work at 905 or 1550 nm with several trade-offs between the two wavelengths, including atmospheric absorption and laser safety. Atmospheric absorption includes obscurants such as water vapour and smoke. At the wavelength of 905 nm, there is less absorption by water vapour than at 1550 nm [2]. However, the 1550 nm light offers the benefit of having a higher optical power threshold cf. the 905 nm wavelength, whilst still maintaining eye-safety.

There is also research being done into a camera based alternative to LiDAR [3], which relies upon an algorithm to extract depth information from an image captured using a normal camera. Multiple cameras positioned at different angles are required to improve the depth information. One of the main motivations for a camera based alternative, is that cameras are significantly cheaper than the optical detectors used in LiDAR. However, the state of the art from the camera-based alternative is a 74% detection rate of correct object ranges at a range of 30 m [4]. The 74% accuracy was achieved using a computer vision method, trained using the KITTI stereoscopic dataset [5, 6]. The KITTI stereoscopic datasets contain images using two cameras pairs (colour and monochrome) spaced 0.54 m apart, each image in the dataset is from one of the camera pairs. Though an accuracy of 74% at 30 m ranges is impressive, and may result in the use of stereoscopic camera systems, the range in stereoscopic camera systems focal lengths is lacking. The focal length, and thus accuracy, of the system depends upon the camera used and the spacing between the two cameras. The camera system currently also suffer a large accuracy reduction in low light level environments.

1.1.2 QKD

Quantum Key Distribution (QKD) is a secure communication system that relies upon the detection of single photons [7]. QKD systems exploit the “no cloning principle” [8],

which says that it is impossible to replicate a quantum state, to secure the communication system. This also holds for Quantum Secret Sharing (QSS) [9]. QKD is between 2 parties, Alice and Bob, whilst QSS systems are for more than 2 parties. Both QKD and QSS systems are secure from eavesdropping because eavesdropping on the system requires the photons to be absorbed. The absorption removes the photons from the link between the parties, causing detectable errors [10]. In both the QKD and QSS schemes single photon detectors such as SPADs are required.

The maximum range of a QKD link depends upon the exact QKD scheme employed and the losses of the link (typically an optical fibre). Using ultra low loss optical fibres, without optical repeaters, a link distance of greater than 250 km has been demonstrated using standard QKD, whilst a twin-field QKD (TF-QKD) scheme has been demonstrated at ranges of over 500 km [11]. Unlike normal QKD, TF-QKD uses light pulses from two light sources that are encoded to obscure the true nature of the pulse [12].

1.1.3 TCSPC

Time Correlated Single Photon Counting (TCSPC) measurement systems are increasingly important for a wide variety of applications, with a few examples given here. Their operation relies on the accurate repeatable detection of very low light level signals, often down to near or at the single photon level. Without optical detectors with small timing jitter, these measurements would remain very challenging with a limited accuracy. However, the detectors used can't solely concentrate on the reduction of timing jitter, as this could come at the cost of other important operating characteristics, such as detection efficiency and false/dark count rates (more details in section 2.2.4).

Optical Tomography

Diffused Optical Tomography (DOT) is a non-invasive medical imaging technique used to trace water and haemoglobin in the human body [13]. DOT requires illumination of the body with near-IR light (650-1000 nm), which is mainly absorbed in the human body by water and haemoglobin. The non-absorbed near-IR light is scattered by the tissue, before emerging where it is detected. Using computer models the scattering of the near-IR light can be reconstructed to produce an image of the illuminated tissue. DOT can be used to detect breast cancer [14], without exposing the patient to ionising radiation used in X-ray mammograms. As DOT depends upon collecting highly scattered photons, the intensity of the optical signals can be very low, requiring the use of single photon detectors to achieve

a high resolution.

In addition to breast cancer screening (using 830 nm laser illumination [14]), DOT can also be used as a non-invasive method for monitoring oxygen levels in the brain during brain surgery by monitoring DOT from 687 and 832 nm laser illumination [15].

Raman Spectroscopy

Raman spectrometry is a technique for identifying materials using the Raman scattering of light by the material. Every material has a unique “fingerprint” that can be obtained through Raman spectrometry, which leads to it having many applications especially in the field of non-invasive medical diagnosis where it can be used to: diagnose bone disease; detect breast cancer; calculate cancer margins; and detect glucose [16]. Unfortunately, some Raman spectrometry measurements encounter a problem where the materials’ signature is drowned out compared to the background signal (or noise). The low signal to noise ratio (SNR) of Raman scattering is due to its low interaction cross-section of 10^{-30} cm^2 [17].

Raman scattering, first reported by Raman & Krishnan in 1928 [18], is an inelastic scattering process between optical phonons and electrons. Unique “fingerprinting” of materials is also possible through the measurement of Brillouin scattering (first measured in 1964 [19]), which is the inelastic scattering process between acoustic phonons and electrons. However, Brillouin scattering is much harder to measure due to the lower energy of acoustic phonons cf. optical phonons.

There are two types of Raman scattering, namely Stokes scattering and Anti-Stokes scattering. They are caused by the emission of a phonon and the absorption of a phonon, respectively. A typical Raman spectrum will show two peaks caused by the Stokes shift from the Transverse Optical (TO) and Longitudinal Optical (LO) phonons. Stokes shifted peaks are more likely to appear in a Raman spectra than Anti-Stokes peaks. The ratio of Stokes to Anti-Stokes intensities (I_{stokes} and $I_{anti-stokes}$) is

$$\frac{I_{anti-stokes}}{I_{stokes}} = \exp\left(-\frac{\hbar\Omega}{k_b T}\right),$$

where $\hbar$ is the reduced Plank’s constant, k_b is Boltzmann’s constant, Ω is the phonon energy, and T is the temperature of the sample. Raman spectra are normally plotted in wave number (cm^{-1}) to show the shift ($\Delta\omega$) between the excitation wavelength (λ_0) and the emission wavelength (λ_I). $\Delta\omega$ is calculated using $\Delta\omega = \left|\frac{1}{\lambda_0} - \frac{1}{\lambda_I}\right|$.

In [17] it is demonstrated that a significantly improved SNR can be achieved for Raman spectroscopy by changing the operating mode of the optical detector. [17] used a charge coupled device (CCD) camera in the study, which was changed from free-running mode to gated mode. In free-running mode operation the CCD camera is left continuously "armed" so any light that hits the detector will be observed. This is in direct contrast to gated operation where the detector is being switched between armed and unarmed by a gate signal. As an example, the gate signal could be a repeated signal from a pulse generator or a triggering signal from a pulsed laser generated when the laser emits a pulse. For Raman spectrometry the use of gated mode helps to overcome the issue of low SNR, caused by the low interaction cross-section, by reducing difference between the collection duration of background light and signal of interest. Whilst gated operation improves the SNR of Raman spectrometry, there is still room for further improvement as the low intensity signals can still be missed using a CCD camera. Combining single photon detectors with gated operation can further increase the SNR of Raman spectrometry.

Raman Spectrometry is able to use a wide range of wavelenegths, 488-1064 nm, with shorter wavelengths having higher interaction rates at the expense of increased possibility of photodecomposition [20]. The use of 1 dimensional SPAD array has been proposed for a compact time correlated Raman Spectrometry system [21].

FLIM

Fluorescence Lifetime Imaging Microscopy (FLIM) is a microscopy technique reliant on the emission and detection of photons. Unlike standard microscopy, FLIM images are based upon the excitation duration of the sample. FLIM works by exciting fluorescent dyes, which emit a photon when they relax. Using TCSPC systems to measure the photons emitted by the dye, an image can be built up overtime. TCSPC systems need accurate recordings of the emission time of the low intensity signals. One of the applications of FLIM is the study of DNA, where FLIM measurements with errors as low as ± 26 ps have been reported [22]. The operating wavelength of FLIM detectors depends upon the fluorescent dye used, so commercial FLIM measurement systems are available for wavelengths between 266 and 1064 nm [23].

1.2 Types of Optical Detectors

1.2.1 PMTs

Photomultiplier tubes (PMTs), which were first commercialised in 1936 [24], are vacuum tubes consisting of a photocathode, an electrode, and multiple dynodes. Typically PMTs are operated at high voltages, around 1000 V, and are capable of producing very large gains (10^6 - 10^7 are commercially available [25]). The largest gain that can be produced is d^n, where d is the average gain of a singular dynode, and n is the number of dynodes in the PMT. The photocathode is coated with a photoemissive material which emits a photoelectron when the photocathode is struck by an incident photon. A schematic of a typical PMT can be seen in Figure 1.1.

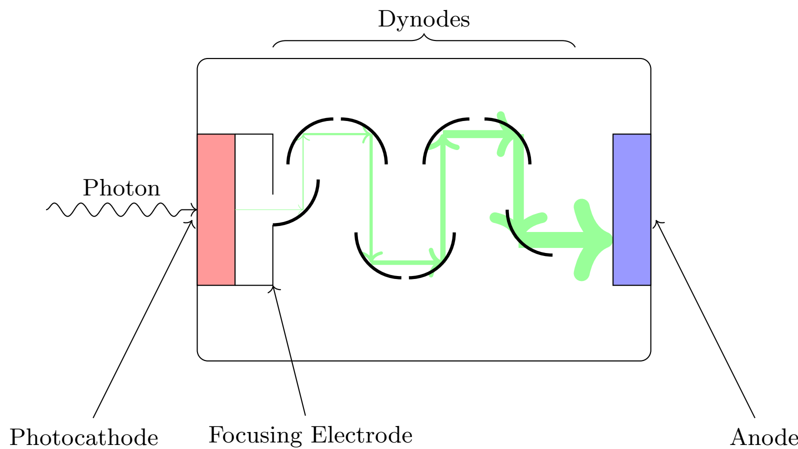

Figure 1.1: Schematic diagram of a PMT. The green arrows denote the electron beam path and the thickness denotes the beam intensity.

A significant disadvantage of PMTs are their large variations in their timing characteristics, which are dominated by their Transit-time spread (TTS), where the TTS of a PMT is the Full-Width at Half-Maximum of the time response. Commercial PMTs have a wide range of available time responses, with the fastest being 0.7 ns (with a TTS of 0.37 ns) [25] and the slowest having a time response of 20 ns (with a TTS of 18.5 ns) [26]. To improve the timing characteristics of PMTs, hybrid PMTs (HPMTs) were created. HPMTs, replace the dynodes from standard PMTs with an avalanche diode, as depicted in Figure 1.2. In HPMTs the main source of gain is from when the photoelectron impacts the bombardment region at the surface of the avalanche diode (∼1500), then the avalanche diode will provide additional gain (device dependent). The HPMTs from Hamamatsu have comparable gains to standard PMTs (1.2×10^5), with response rise times of only 400 ps and a significantly improved TTS of 50-130 ps depending on wavelength [27].

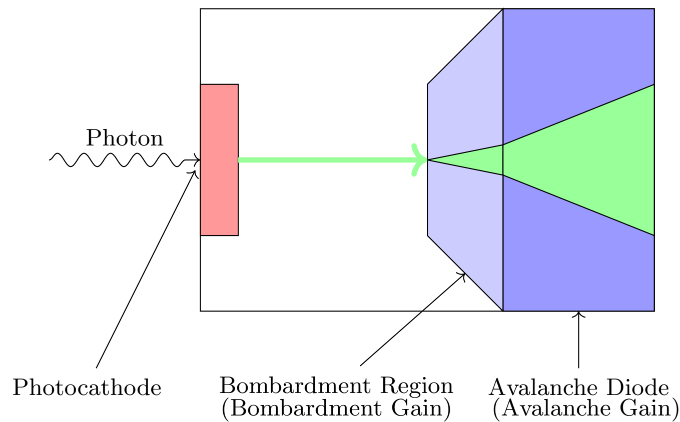

Figure 1.2: Schematic diagram of a HPMT where the width of the green region denotes the electron multiplication process.

1.2.2 MCP

Microchannel plates (MCPs), schematically shown in Figure 1.3, are glass plates with lots of capillaries (holes) running through them. The front and back of the MCPs are coated in metal to make the electrode and cathode whilst the inside of the capillaries is coated with a photoemissive material, similar to the dynodes in a PMT. Typically operated using a 1 kV supply voltage, commercial MCPs are capable of achieving gains 10^4 from a single MCP whilst combining 2 or 3 MCPs in stages can yield gains of 10^7 [28].

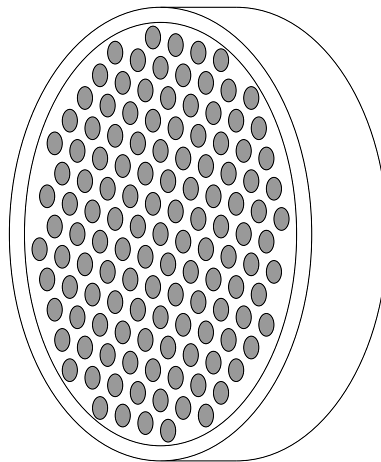

Figure 1.3: Schematic diagram of a MCP.

1.2.3 APDs

Avalanche photodiodes (APDs) are diodes operated at reverse bias, below their avalanche breakdown voltage (V_b). These devices exploit the impact ionisation process (more details in section 2.1.1) to produce avalanche gain. To fully exploit the impact ionisation process, the APDs require carefully designed structures to maintain a high electric field able to enable the impact ionisation process. In an APD the main performance criteria, explained more in section 2.1.2, are low dark current (the current flowing through the APD when it's

not illuminated), a low excess noise factor (F, defined as noise of avalanche multiplication) and a high avalanche multiplication.

1.2.4 SPADs

Single Photon Avalanche Diodes (SPADs) are APDs designed to operate above their avalanche breakdown voltage in the Geiger-mode. By operating above their avalanche breakdown voltage SPADs are capable of producing a measurable electrical signal from the absorption of a single photon. The difference between an APD and a SPAD may seem quite simplistic and it is true that some APDs can be operated in the Geiger-mode (as SPADs) and that some SPADs can be operated in the linear-mode (as APDs). It is however not ideal for APDs to be operated as SPADs.

There are several key differences in the design consideration of an APD and a SPAD. For a SPAD the key criteria are high photon detection efficiency (PDE), low timing jitter, and low dark count rate (DCR). The PDE of a SPAD is the product of the external quantum efficiency (η) and the breakdown probability of the SPAD (P_b). Timing jitter is defined as the FWHM of the time between a photon being absorbed and a measurable signal being generated, and DCR is the rate of false counts in a device. These performance parameters will be explored in more detail in section 2.2.4, however they lead to subtly different device designs [29] cf. APDs.

An example of the difference in designs between an APD and a SPAD would be the multiplication region thickness, where in an APD it is common to see a thin multiplication regions though this is mainly for high-speed applications. A thin multiplication region in an APD is sufficient to give a high avalanche gain, dependent on operating voltage, while keeping F low [30]. However, using a thin multiplication region in a SPAD tends to result in a high DCR due to the higher electric-field across the multiplication region enabling band-to-band tunnelling, which promotes valence electrons to the conduction band (more details in section 2.2.4). If the multiplication region is made thicker, it can reduce the DCR, by reducing the electric-field across the multiplication region. However, this may come at the expense of the PDE and increase the timing jitter. This means SPAD design has an extra trade-off leading to more complicated design than in APDs.

1.2.5 SiPMs

Silicon Photomultipliers (SiPMs) are arrays of Silicon SPADs each with their own quenching resistor and a common anode and cathode. Some suppliers like Hamamatsu offer arrays of SiPMs, known as Multi-Pixel Photon Counters. Currently the largest pixel count (number of individual SPADs) for a SiPM from Hamamatsu Photonics is 57600 (with a fill factor, F_{fill}, of 47%, where the fill factor is defined as the proportion of the array size capable of absorbing a photon) [31]. SiPMs were created as a solution to one of the main drawbacks of traditional Silicon SPADs where the DCR of a Silicon SPAD scales with the area of the SPAD, making large area SPADs less attractive. By producing large arrays of small SPADs, approximately 15-25 µm in diameter [32, 33], the SiPM can offer a higher effective detection area than a traditional SPAD but still keep DCR low. However, one of the main drawbacks to the SiPM concept is the pixilated nature of the SiPM. In the previous section the PDE of a SPAD was described as $\eta \times P_b$. However the PDE of an array of SPADs needs to account for the fill-factor of the array. This gives $PDE_{array} = F_{fill} \times \eta \times P_b$, which means the largest pixel count SiPM from Hamamatsu has a theoretical maximum PDE_{array} of 47%.

1.2.6 Nanowires

Nanowires have become a growing trend over recent years and detector wise can be broken down into 2 broad categories; Superconducting nanowire single-photon detectors (SNSPDs) and semiconductor nanowires.

SNSPDs

SNSPDs are fabricated as a long continuous meanders of a thin superconducting material. SNSPDs were first demonstrated by Semenov et al. in 2001 [34]. Due to the meandering nature, SNSPDs can offer a high density of photosensitive area covering a large proportion of the collection area, though they are limited to cryogenic operating temperatures of 2-5 K.

Semiconductor Nanowires

Semiconductor nanowires can either be fabricated as singular nanowires or arrays of nanowires. Singular nanowires offer advantages for on-chip coupling of light out of waveguides that see use in quantum computer circuitry [35]. These singular nanowires can be grown in situ as part of the waveguide fabrication process. Semiconductor nanowires have

been shown as an alternative to standard SPAD designs [36], as they are able to improve on the performance of each individual pixel.

1.3 Thesis Outline

Chapter 2 gives the background behind the key performance parameters of APDs and SPADs alongside several different models capable of modelling these performance parameters. The experimental methods used to measure these performance parameters are discussed in Chapter 3 alongside the key fabrication steps required for fabrication Si SPADs.

Chapter 4 demonstrates a simple monte carlo model for Si SPADs that was validated using experimental voltage dependent avalanche gain, $M(V)$, and avalanche gain dependent excess noise factor, $F(M)$, as well as reported silicon drift velocities, diffusion coefficients, and ionisation coefficients. This model was used to simulate SPAD parameters breakdown probability, time to breakdown, and timing jitter of 2 different SPAD designs to compare the effects of doping profile orientation.

Chapter 5 reports a new simple monte carlo parameter set for InP. This parameter set was validated against reported saturation velocities, impact ionisation coefficients, room temperature gain and excess noise data, as well as temperature dependent gain. The model has also been used to extract effective ionisation coefficients for the use with simpler models under electric fields of 400-800 kV.cm^{-1} at temperatures of 150, 200, 250, 290 K.

Chapter 6 reports on the results of Si mesa SPADs that achieve a single photon detection efficiency of 69%.

Chapter 2

Background

2.1 APDs

The theory of avalanche breakdown in semiconductors can be traced back to 1954 [37] with photon emission being observed from avalanche breakdown in 1956 [38]. The photon emission was observed from localised impact ionisation sites, at the time called micro-plasma sites. The photon emission was due to carrier recombination at these sites. Then in 1960 broad area sites, called macro-plasma sites, were observed [39], which were the beginning of avalanche regions as they are known today. An avalanche gain of 1000 (20 dB) was reported for Si photodiodes in 1964 [40], which was quickly followed by Si APD being used to enhance SNR in 1965 [41]. Commercial Si APDs are available from a number of suppliers worldwide.

2.1.1 Impact Ionisation

The operation of APDs and SPADs rely on the impact ionisation process. In an electric field, free electrons and holes are accelerated, gaining energy from the electric field. Impact ionisation occurs when an electron or hole with sufficient energy promotes a valence electron to the conduction band by impacting the crystal lattice and transferring energy to the valence electron. The promoted electron and generated hole, will then also gain energy from the electric field, allowing them to undergo impact ionisation before they leave the high electric field region. Figure 2.1 depicts chain of impact ionisation events occurring before the free carriers leave the high electric field region. In Figure 2.1, a single injected electron results in 4 impact ionisation events before all the carriers leave the active area, producing an avalanche gain of 5 for the event.

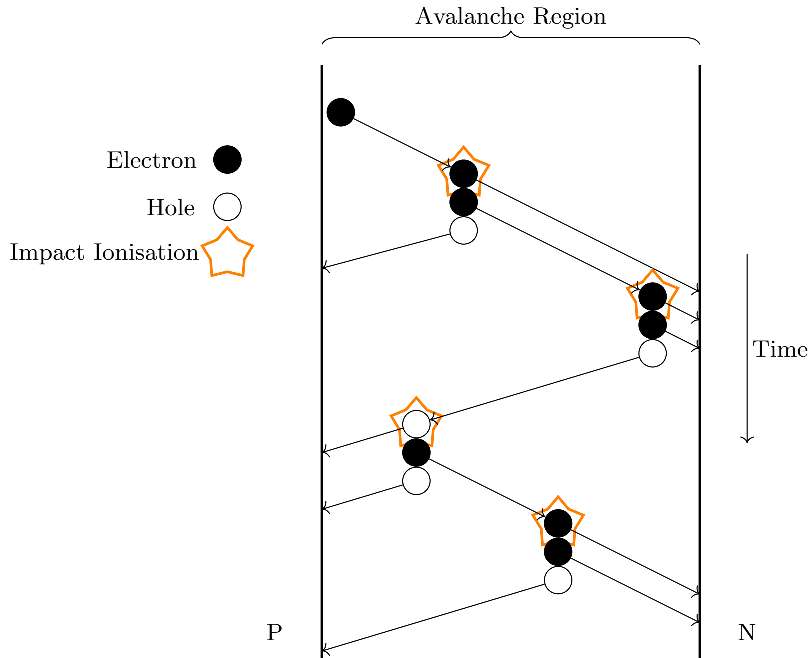

Figure 2.1: Schematic of avalanche gain from a single free electron.

The rate that electrons and holes can build up energy is material and electric field dependent. In addition to gaining energy from the electric field, there are several carrier interaction mechanisms that cause electrons and holes to gain/ lose energy in an APD. The main mechanisms are interactions with phonons, defects, and crystal lattice variations due to alloying multiple semiconductor materials together. Phonons are vibrations within the crystal structure, the two phonon interactions being phonon absorption and phonon emission. During phonon absorption a carrier absorbs a phonon and gains the phonon energy (∼10s meV), whilst during phonon emission a carrier emits a phonon and loses the phonon energy. Alloy scattering is an effect caused by the atomic non-uniformity of an alloy structure; alloy scattering effects cause the carriers to lose energy. Similarly, interactions with defects also cause the carriers to lose energy. The defects can take the form of crystal vacancies, substitution, dislocations, threading dislocations and traps. Traps are localised potential minima which carriers can fall into, until they gain enough energy to overcome the trap potential. Traps can also lead to trap assisted tunnelling (discussed later in section 2.2.4).

As part of the study of impact ionisation, electric field dependent impact ionisation coefficients were proposed, $\alpha(E)$ for electrons and $\beta(E)$ for holes. $\alpha(E)$ and $\beta(E)$, typically reported with units of cm^{-1} or m^{-1}, are the inverse of the mean distance between consecutive impact ionisation events for a single carrier in a uniform electric field of E. Deadspace effects were used to model avalanche gains and excess noise factors in 1990 [42], where the model was based on earlier non-local modelling by Okuto and Crowell [43]. Deadspace

theory proposes carriers must have a minimum energy to undergo impact ionisation, therefore, after carrier excitation or impact ionisation, the probability the carrier can undergo impact ionisation must be zero and the carrier must be accelerated for some distance to reach the energy required for impact ionisation to be possible. This distance is called deadspace, which is material and electric field dependent. In APD simulations that include deadspaces, there are two methods to include them: soft deadspace and hard deadspace. Both methods rely upon fitting to an impact ionisation probability density distribution. This distribution is constructed from the probabilities that a carrier will impact ionise after travelling a given distance and then normalised over all distances. The probabilities required to produce the probability density distribution can be calculated from complex models, which are discussed later.

An example of a hole impact ionisation probability density plot, fitted using a hard deadspace method, can be seen in Figure 2.2. As shown, using the hard deadspace method, the probability densities for a distance shorter than the deadspace (calculated as shortest path length with a probability half the peak probability density) are set to zero. The soft deadspace method requires fitting the entire probability density distribution which is significantly harder than fitting using hard deadspace.

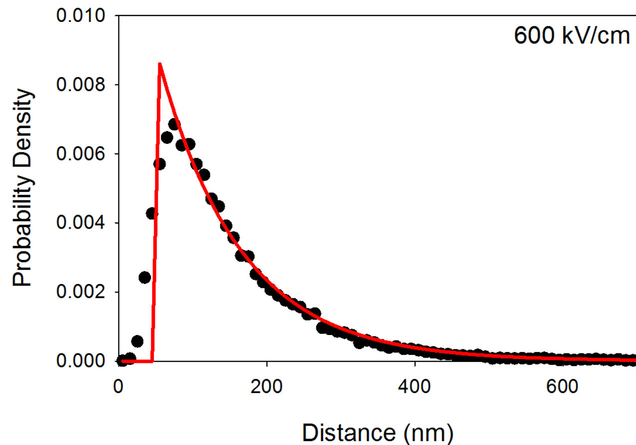

Figure 2.2: Example of hard dead space approximation (line) to simple monte carlo generated (symbols) impact ionisation probability densities, for holes in an 600 kV.cm^{-1} electric field, in InP at room temperature.

The hard deadspace approximation is used in this work and the hard deadspace equations are introduced later in Equations (2.18) and (2.19). Impact ionisation coefficients that need to be used with deadspace are called effective ionisation coefficients and used the notation $\alpha^*(E)$ and $\beta^*(E)$ rather than $\alpha(E)$ and $\beta(E)$ where,

$$\alpha^* = \frac{1}{\frac{1}{\alpha} - d_e}, \text{ and } \beta^* = \frac{1}{\frac{1}{\beta} - d_h}, \tag{2.1}$$

and d_e, d_h, represent the electron and hole deadspaces.

2.1.2 Performance Parameters

Gain

Avalanche multiplication (M), also referred to as Gain or Avalanche Gain, is a consequence of the impact ionisation process. For Figure 2.1 the resulting gain is 5 as, for the 1 electron injected, 4 additional electrons were generated, leading to a total output of 5 electrons; the same would be true if a hole initiated the impact ionisation chain. Quoted M values are mean multiplication values of the APD, rather than a guaranteed gain for an individual absorbed photon. The random nature of impact ionisation mean the gain from an absorbed photon will vary photon to photon. M can be described as

$$M = \langle m \rangle, \tag{2.2}$$

where m is the avalanche gain generated from a single photon being absorbed, and $\langle m \rangle$ represents the mean of m. The values of m, can vary quite significantly due to the random nature of impact ionisation. The measure of this variation/ noise is the excess noise factor (F).

Excess Noise

The excess noise factor (F) is a measure of the noise associated with avalanche gain. It is a measure of the variation in the avalanche gain for each absorbed photon. A lower F means that the gain is more uniform cf. a high F device. The equation,

$$F = \frac{\langle m^2 \rangle}{\langle m \rangle^2} \tag{2.3}$$

is a good description of excess noise but this equation better serves simulations compared to practical measurements as APDs operated in the linear mode (where F is relevant) are unable to produce large enough currents from a single absorbed photon to be measurable.

2.2 SPADs

2.2.1 Avalanche Breakdown

Whereas APDs operate below their breakdown voltage (V_b), in linear mode, a SPAD is operated above V_b in Geiger mode. An example $M(V)$ is shown in Figure 2.3 which

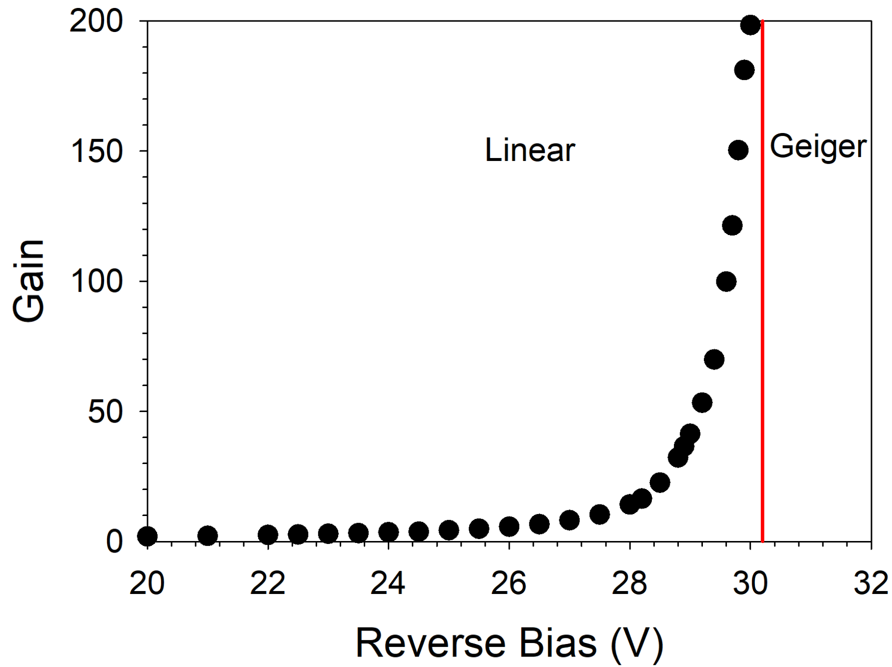

Figure 2.3: Example $M(V)$ indicating linear and Geiger mode operation with respect to the breakdown voltage (V_b)

indicates the breakdown voltage and transition from linear to Geiger mode operation.

In Geiger mode operation a single free electron or hole entering the multiplication/ avalanche region of a SPAD is able to trigger an avalanche breakdown event. An avalanche breakdown event is a self-sustaining event, meaning that most carriers will impact ionise at least once before leaving the avalanche region. In order to maintain avalanche breakdown both the electrons and holes must be able to impact ionise. As long as the current generated reaches the latching current it can continue to increase until the device is quenched by external quenching circuitry (see section 2.2.3). If the current does not reach the latching current then the avalanche event can stall, self-quenching itself, as the current is unable to be maintained.

2.2.2 Equivalent Circuit

Whilst the equivalent circuit of an APD can be thought of as a resistor and capacitor in parallel, [44] describe the equivalent circuit of a SPAD slightly differently. The equivalent circuit of a SPAD used by [44] can be seen in Figure 2.4. In their equivalent circuit C_d represents the junction capacitance and R_d represents the diode resistance. Triggering an avalanche event corresponds to closing the switch in the circuit. Initially while the switch is open, the capacitor C_d charges up until there is a voltage V_a across it, where V_a is the applied bias. The bias across the capacitor is called the diode voltage, V_d. When the

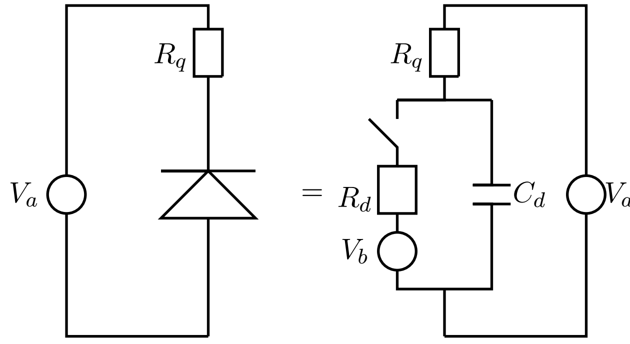

Figure 2.4: Equivalent Circuit of a SPAD [44]. Where V_d, R_d, C_d represent the diode voltage, the resistance across the SPAD and the SPADs junction capacitance respectively. V_a, R_q represent the applied bias and the quenching resistor.

switch is closed, by the triggering of an avalanche event, the capacitor begins to discharge creating a current, I_d, through the resistor R_d, where,

$$V_d = V_b + R_d I_d. \tag{2.4}$$

As the capacitor discharges, and V_d decreases, voltage must be dropped across the quenching resistor R_q, such that,

$$V_a = V_d + R_q I_q. \tag{2.5}$$

The current, I_d, is now made up of contributions from both the discharge of C_d and the current through R_q. Eventually C_d will discharge so that there is only a bias of V_b across it. At which point,

$$V_a = V_b + R_q I_q, \tag{2.6}$$

$$I_q = \frac{V_a - V_b}{R_q} = \frac{V_{ex}}{R_q}, \tag{2.7}$$

where V_{ex} is the applied overbias. The current of I_q will be sustained until either the SPAD is manually quenched by reducing V_a by more than V_{ex} or the instantaneous current is the SPAD drops below the latching current, due to the stochastic nature of impact ionisation, and the SPAD self-quenches. Either event will cause the switch to open.

2.2.3 Quenching Circuitry

In the operation of a SPAD quenching circuitry is really important. The quenching circuitry is used to stop the avalanche event, by bringing the SPAD below its breakdown voltage, before rebiasing the SPAD to enable it to detect again. The different quenching methods can be broken down into three broad categories: Passive, Gated, and Active [44].

The simplest of these is passive quenching with the use of a large ballast resistor (typically

100-200 kΩ) in series with the SPAD. As the current increases through the SPAD, the voltage dropped across the ballast resistor increases which reduces the voltage across the SAD to below V_b. The drawback of this technique is that it is always on, so even before the current has reached a detection threshold the SPAD is being partially quenched, which could increase the build up time required to hit the detection threshold.

Gated quenching works by applying a gate signal upon a fixed DC bias. This DC bias is set to slightly below the SPAD breakdown voltage and the gate signal will bring the SPAD above it's breakdown voltage for the duration of the gate period, before bringing it back below the breakdown voltage to quench the SPAD. Commonly the gate signal is a standard square wave (used in this thesis) though sinusoidal gating is also used to reduce the gate duration [45]. In gated mode the maximum counts from the SPAD is identical to the number of gate cycles, as each gate can only detect a single event. This is useful as it allows for easy time correlation of the event. However, there are several disadvantages to gated operation. The first of these is capacitive transient effects, which can distort the rising and falling edges of the gate pulse. The rising edge will overshoot the intended voltage before decaying back to the desired level. An example of transient effects and ways to mitigate them is discussed in the dark count rate measurements section (section 3.4). The second of these drawbacks is that each gate can only detect a single event, so if an event happens very early in the gate duration the rest of the gate period is wasted.

Finally, we have active quenching. In active quenching systems the SPAD is kept biased above its breakdown voltage until a avalanche current is detected. When the avalanche current is detected the bias across the SPAD is reduced to below the breakdown voltage to stop the avalanche. The two common methods to achieve this are to either superimpose a opposite polarity pulse onto the applied bias line to bring the applied bias to below the breakdown voltage [46], whilst the second is to directly adjust the low bias terminal which reduces the bias being dropped across the SPAD [47].

2.2.4 Performance Parameters

The performance of a SPAD is measured by three key performance parameters [48]. These parameters are measures of how many of the events are false events created by the device (dark count rate, DCR), the variation in the time of the detection signal (jitter) and how many incoming photons are detected/missed (photon detection efficiency, PDE).

DCR

Dark count rate (DCR) is a key performance parameter used for SPADs. A dark count occurs when the SPAD undergoes an avalanche event despite the lack of photon. Mechanisms behind dark count generation include: thermal excitation, band-to-band tunnelling, trap assisted tunnelling, optical crosstalk, and after-pulsing [49].

Thermal excitation is where the carriers are able to gain enough thermal energy to promote themselves from the valence band to the conduction band. Materials with smaller band gaps are more susceptible to thermal excitation. Itzler et al. [50] found with their $In_{0.53}Ga_{0.47}As$/InP SPADs that there was significant thermal excitation taking place in their $In_{0.53}Ga_{0.47}As$ absorption region with a band gap of ~0.8 eV cf. ~1.4 eV for InP.

Band-to-band tunnelling is where valence electrons are able to tunnel through the forbidden region to the conduction band [51]. The forbidden region is the region between the top energy state of the valence band and the bottom state in the conduction band, electrons are not able to stay in the forbidden region due to a lack of available energy states. A schematic illustrating band-to-band tunnelling is shown in Figure 2.5. Band-to-band tunnelling is an electric field dependent parameter, as it becomes easier under higher electric field strengths (E) due to the narrowing of the forbidden region with increasing electric fields, as can be seen in the equation for the thickness of the forbidden region W_f [52],

$$W_f = \frac{E_g}{qE}. \tag{2.8}$$

The rate of band-to-band tunnelling, R_{btb}, has previously been reported for Si APDs [53] as

$$R_{btb} = A_{btb}E^{2.5}B\exp\left(\frac{E}{C_{btb}}\right), \tag{2.9}$$

where B is set to 1 if band-to-band tunnelling is possible, or 0 if not, and the values of A_{btb} and C_{btb} are $4 \times 10^{14}\text{cm}^{-0.5}.\text{V}^{-2.5}.\text{s}^{-1}$ and $1.7 \times \text{V.cm}^{-1}$ respectively. The resulting value of R_{btb} is a rate that is also dependent upon the volume of the device (events/cm^3/s). Equation (2.9) can be used to obtain design guidelines for maximum permissible electric fields to achieve a particular DCR from Band-to-Band tunnelling. Assuming that all carriers that enter the depletion region trigger an avalanche breakdown event (breakdown probability of 1) and uniform electric field strengths, then for a 10 µm radius diode, with a 1.5 µm thick depletion region the maximum permissible electric field strengths to stay under 100 Hz and 1000 Hz DCR are 473 kV.cm^{-1} and 500 kV.cm^{-1} respectively.

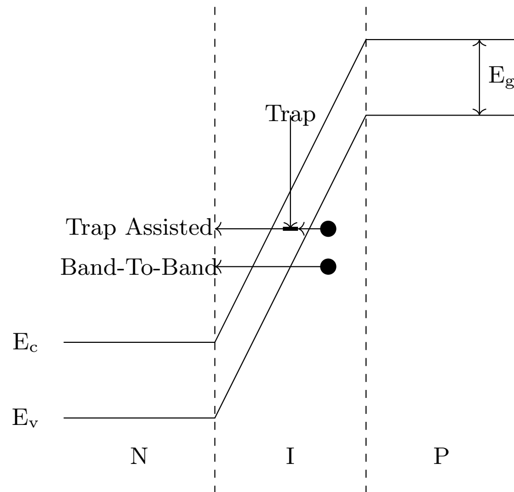

Figure 2.5: Schematic showing Band-to-Band tunnelling and Trap Assisted tunnelling mechanisms

Trap assisted tunnelling makes it easier for electrons to tunnel to the conduction band. The traps split the forbidden region, effectively creating two separate smaller barriers that can be overcome through phonon absorption or thermal effects [51]. Trap assisted tunnelling is illustrated in Figure 2.5. Recent modelling work for Si waveguide SPADs found that the rate of Band-to-Band tunnelling >> rate of trap-assisted [54].

Optical crosstalk is a mechanism that affects SPAD arrays and can contribute to DCR. A SPAD undergoing avalanche breakdown may emit a photon, which is then absorbed into a neighbouring SPAD, triggering an avalanche event in the neighbouring SPAD. To mitigate optical crosstalk SPADs are often separated with optical isolation, e.g. metal or dielectric coatings. These coating can mitigate crosstalk, however can't fully stop optical crosstalk as photons have been shown to reflect off the substrate of a device [55].

A further source of dark events is after-pulsing, where a single absorbed photon can lead to multiple avalanche events or detectable counts. Causes of afterpulsing include the relaxation of traps [56] and ineffective quenching of the SPAD leading to the retriggering of detection circuitry [57]. To combat after pulsing, dead/hold-off times are often used after a detector has undergone an avalanche event, however this reduces the duty cycle of the detectors. Dead times of under 1 µs [58] to 20 µs [59] have been reportedly used to suppress afterpulsing. A recent study on afterpulsing has claimed that afterpulsing is related more closely to the quenching circuitry rather than the SPAD [60].

Jitter

Timing jitter is the variation in time between the absorption of an optical signal and the subsequent detection of the resultant avalanche current. It describes how accurately the device can determine the arrival time of a photon. In the instance of time of flight applications, this would directly affect the error on the measured distance, with every 100 ps of jitter corresponding to a 1.5 cm error. To reduce this error it is common for systems to average over multiple measurements. For example, [61] used acquisition times of 3 ms per pixel using a 15.6 MHz source and a 30 ns gate on the detector.

Both the device and the external counting circuitry contribute to timing jitter. One contributor to device jitter is the injection position of the photon where differences in field profile between edge and centre of the device lead to differences in the build-up in avalanche current resulting in differences in the time taken to reach threshold currents [62]. These edge jitter contributions have been significantly reduced in modern SPAD designs through the incorporation of metal rings that cover the edge of the devices [63] preventing photon absorption at the device edge. Another cause of timing jitter variation arises from the lateral spreading of carriers within the device, though it has been reported to be fairly insignificant [62]. The photon absorption position and drift-diffusion also contribute to the timing jitter of the device.

PDE

As mentioned in chapter 1, the Photon Detection Efficiency (PDE) is different for SPAD arrays cf. stand alone devices. For a stand alone SPAD made of a single material the PDE can be described using,

$$PDE = \eta \times P_b \tag{2.10}$$

where η is the external quantum efficiency (i.e. the likelihood that an incoming photon will be absorbed into the SPAD and generate an electron-hole pair), and P_b is the breakdown probability (defined as the probability that an electron-hole pair will result in an avalanche breakdown event). For an avalanche breakdown event to have occurred, the SPAD must generate a self-sustaining avalanche current which grows to a measurable level, before being quenched by external circuitry. In this work a threshold detection level of 0.1 mA has been used [64] as this results in a 5 mV drop (across a 50 Ω resistance) which is routinely detectable experimentally [65]. The breakdown probability of a given SPAD increases as the applied bias is increased beyond its breakdown voltage. The difference between this applied bias and the breakdown voltage is called the over-bias.

Whilst for an array PDE_{array} is represented by,

$$PDE_{array} = F_{fill} \times \eta \times P_b \tag{2.11}$$

where F_{fill} is the fill factor indicating how much of the total array area is made up of the active (photosensitive) area of the devices. Equations (2.10) and (2.11) for PDE and PDE_{array} are true for Silicon and other single material SPADs but need slight modification for SPAD structures with their absorption region and multiplication region made from different materials. For these SPADs an additional parameter, P_t, is introduced to account for the probability that carriers will transition from the absorption to multiplication region. Thus PDE and PDE_{array} become,

$$PDE = P_t \times \eta \times P_b, \tag{2.12}$$

and

$$PDE_{array} = P_t \times F_{fill} \times \eta \times P_b. \tag{2.13}$$

2.2.5 Silicon

Silicon is a very widely used material in photodiodes, capable of absorbing light from 400-1100 nm. There are several different Silicon fabrication techniques, ranging from the older techniques using custom design to more modern complementary metal-oxide-semiconductor (CMOS) designs. Custom SPAD designs offer fully customisable electric field profiles, hence design freedom of depletion region widths. The custom designs can be heavily annealed to reduce the dark count rate by repairing defect sites caused by dislocations. Custom designs also exploit the gettering process to remove generation-recombination sites from the device active area to further reduce the dark count rate [66]. The gettering process is a mechanism for removing metallic impurities and crystal defects from the active area of a device by attracting them to gettering sites. Gettering sites can be created through ion implantation [67] or through dopant diffusion [68].

The first custom SPADs used the p-n diode design [69] originally proposed and developed in 1965 [70]. A timing resolution of 60 ps was reported in 1987 using this p-n diode design [71]. To increase the timing resolution the p-n diode design was replaced with a double epitaxial structure in 1989 [72]. The double epitaxial design achieved a timing resolution of only 30 ps [73], however the double epitaxial design suffered from a high dark count rate which limited the diode size. The active area section of the double epitaxial design is

still in use, with the addition of sinker contacts to improve the electric field properties of the active area and reduce the defect count in the active area by exploiting the gettering mechanism [74]. More recently there has been significant work on a red enhanced SPAD (RE-SPAD) from SPAD Lab [75], by increasing the SPAD absorption region to a thickness of ∼10 µm [76] to compensate for the low absorption coefficient of red and near-IR light in silicon. A PDE of 20% at 900 nm cf. 70% at 600 nm was reported [76]. A drawback of this approach is an increase in the breakdown voltage (a direct result of increased depletion region width) and has the potential to significantly affect the achievable timing jitter of the structure. Timing jitter results of the new RE-SPADs have not been published.

Unlike the custom designs, CMOS devices are limited to very thin structures, which limits their absorption efficiency [77]. Another drawback of CMOS processing is that it lacks a compatible high temperature annealing process capable of removing defects, such as dislocations, generated by ion implantation [78]. These implantation defects can act as generation sites, increasing the dark count rate of the SPAD. However, there are several attractive properties of CMOS SPADs. As CMOS SPADs are significantly thinner than custom SPAD designs they require a significantly smaller power supply to reach their breakdown voltage cf. thick SPADs such as the Excelitas C30902SH (with breakdown voltages typically ∼225 V) [79]. CMOS SPAD designs can also have external quenching circuitry fabricated directly on the same chip, eliminating the need for complicated wire bonding, such as the aluminium wedge bonding shown in Figure 2.6. The on-chip quenching circuitry may however be at the expense of array fill-factor and thus detection efficiency.

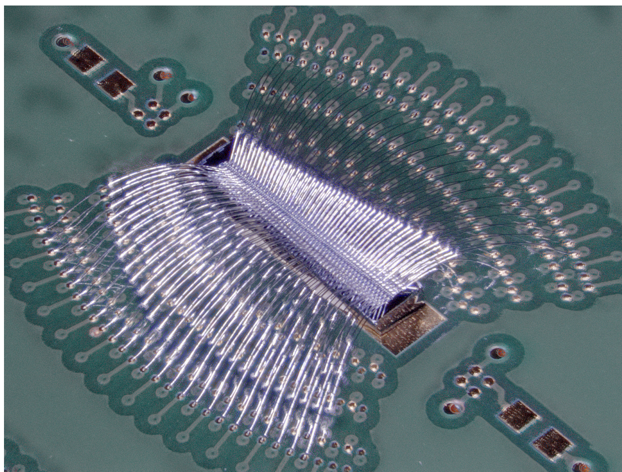

Figure 2.6: Image showing a test sample with 256 aluminium wire bonds. Sample has been wire bonded by Mr Simon Pyatt[1] onto a pcb designed in collaboration with Prof. Jon Lapington[2].

[1]Mr Simon Pyatt is currently associated with the Particle Physics Group, School of Physics and Astronomy, University of Birmingham.

[2]Prof. Jon Lapington is currently associated with the Space Research Centre, Department of Physics

An alternative to on-chip quenching circuitry is 3D-CMOS [80] which relies upon a two tier system. The two tiers are the SPAD tier and the circuitry tier, which are then joined together using either a flip-chip bonding technique (with back-illuminated Si SPADs) or Si vias (for top illumination SPADs) [81]. 3D-CMOS offers the advantages of CMOS SPADs without the reduction in fill factor caused by on-chip quenching circuitry and means that the tiers can be made using separate CMOS scales (where the CMOS scale is the minimum achievable feature size) if required [82], though it does add complexity due to alignment between tiers based on different CMOS scales. If the designs are limited to only using standard CMOS fabrication processes then the CMOS SPADs can be made using multiple different foundry services at a reduced cost cf. the custom SPAD designs.

CMOS SPADs fabricated using different CMOS scales are compared against custom Si SPADs in Table 2.1 to establish current state of the art.

Table 2.1: Comparison of published Si SPAD performance

Technology	Size (µm)	V_b (V)	Peak PDE (% @ wavelength (nm))	DCR (Hz)	Jitter (ps)	Ref
40 nm CMOS	Square: 5.4	–	34 @ 560	25	170	[83]
65 nm CMOS	Diameter: 20	9.9	8 @ –	2.8k/ µm^2	8	[77]
	Diameter: 8	9	5.5 @ 450	15.6k	235	[84]
0.16 µm CMOS	Diameter: 50	25.4	18 @ 490	400k	82	[85]
	Diameter: 50	26.3	36 @ 490	100k	78	[85]
0.18 µm CMOS	Square: 25	23.3	–	2.65k	–	[86, 87]
	Square: 14.4	20.5	62 @ 600	12.7k	–	[88]
0.35 µm CMOS	Diameter: 30	25	10 @ 420	117k	240-340	[89]
	Diameter: 50	25	27 @ 420	334k	240-340	[89]
	Square: 50	25	34 @ 420	503k	240-340	[89]
	Diameter: 30	25	55 @ 450	120	312	[90, 91]
	36 x 44	22	4 @ 810	200k	150	[92]
0.8 µm CMOS	Diameter: 5	27	0.62 @ 905	1k	–	[93]
	Diameter: ∼ 7	25.5	26 @ 460	350	115	[94]
Custom	Diameter: 50	32.7	68 @ 700	100s	83	[76, 75]
	Diameter: 15	31	–	∼5k		[95]

2.2.6 InGaAs/InP

Though Silicon is a very good material for a SPAD, it is unable to operate at either of the wavelength bands used for optical fibre communication systems (1.3 µm and 1.55 µm). The 1.55 µm wavelength is important for optical telecommunication due to the low absorp-

and Astronomy, University of Leicester.

tion loss of 1.55 µm photons in optical fibres, but is also a very promising wavelength for eye-safe free space applications such as for LiDAR. 1.55 µm wavelength light is considerd eye-safe as it can be operated at energies $\sim 10^5$ times greater than 633 nm wavelength light whilst still remaining eye safe assuming pulse durations between 1 ns-10 µs [96]. The most common material used for photon absorption at the telecommunication wavelengths is $In_{0.53}Ga_{0.47}As$ (latticed matched to InP). Its cut-off wavelength is ~1.65 µm (the cut-off wavelength is the longest wavelength that has enough energy, calculated using Equation (1.1), to overcome the band gap of the material). Separate Absorption Multiplication (SAM)-SPADs consisting of InP multiplication regions and $In_{0.53}Ga_{0.47}As$ absorption regions have been produced for 1.55 µm detection [97]. A schematic example of a $In_{0.53}Ga_{0.47}As$/InP SAM structure alongside the corresponding electric field diagram is shown in Figure 2.7.

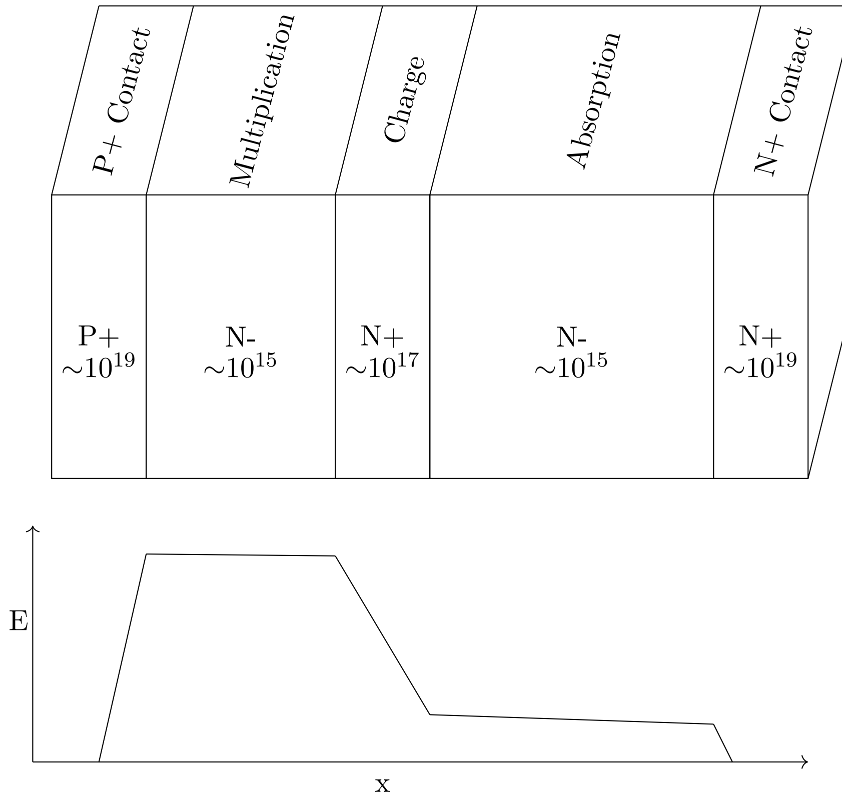

Figure 2.7: A schematic showing the structure of an $In_{0.53}Ga_{0.47}As$/InP SAM-SPAD where the orders of magnitude represent typical doping concentrations in dopants.cm^{-3} (top) along with the corresponding electric field profile (bottom).

Too avoid the SAM-structure, shown in Figure 2.7, containing multiple p-n junctions the intrinsically doped regions have been denoted as N- rather than intrinsic (alternatively the structure could be N+/P-/P+/P-/P+). When reverse bias is applied to the device it will start depleting from the P+/N- interface and gradually deplete the multiplication region as reverse bias is increased. Punch through will occur when the device also depletes the absorption region, and can be noted by a sudden drop in capacitance at that bias on

a capacitance-voltage measurement. The bias this occurs at is called the punch through voltage. With the careful control of the charge sheet doping, between the multiplication and absorption regions, the electric field strength in the multiplication region can be controlled.

An alternative to $In_{0.53}Ga_{0.47}As$/InP SAM-SPADs is $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ SAM-SPADs, which is also latticed matched to InP substrates. Among the reasons that a $In_{0.52}Al_{0.48}As$ multiplication regions are used as alternative is that $In_{0.53}Ga_{0.47}As$/ $In_{0.52}Al_{0.48}As$ devices have been shown to have a more temperature stable breakdown voltage [98], and the band gap of $In_{0.52}Al_{0.48}As$ is greater than InP which should help to reduce the dark currents and result in a lower DCR cf. $In_{0.53}Ga_{0.47}As$/InP. Studies on $In_{0.53}Ga_{0.47}As$/ $In_{0.52}Al_{0.48}As$ DCR have found that the system is dominated by band-to-band tunnelling in the multiplication region [99], they also found there was only a weak temperature dependence of dark count rate suggesting that the large band gap of 1.55 eV suppresses thermal generation of carriers.

2.3 APD & SPAD Models

A key part of APD and SPAD research has been to develop models capable of predicting the performance of diode designs. These models allow for designs to be developed and tailored before expensive and time consuming wafer growth and device fabrication are performed, so that the fabricated devices are more likely to achieve the desired performance parameters whilst minimising costly design iterations.

In later chapters Simple Monte Carlo modelling is used in chapters 4 and 5, with recurrence equation models being used in chapters 4 and 5, and a random path length model being used in chapter 4.

2.3.1 Local Model

The local model, proposed in 1966 [100], is a simple model for calculating avalanche gain and excess noise in an APD. The local model does not take into account the history of carriers and as a result does not account for deadspace, which was not significant for the thick APD structures at the time. From the local model it is possible to derive equations for pure electron gain, M_e, and pure hole gain, M_h, which only depend upon the ionisation

coefficients and the device width, w,

$$M_e = \frac{\alpha - \beta}{\alpha \exp\left(w\left(\alpha - \beta\right)\right) - \beta}, \tag{2.14}$$

$$M_h = \frac{\alpha - \beta}{\alpha - \beta \exp\left(w\left(\alpha - \beta\right)\right)}. \tag{2.15}$$

The equations for M_e and M_h can be rearranged so that the field dependent values $\alpha(E)$ and $\beta(E)$ can be calculated from experimentally measured M_e and M_h. The α and β values for a given field can be calculated from,

$$\alpha = \frac{1}{w}\left(\frac{M_e - 1}{M_e - M_h}\right)\ln\left(\frac{M_e}{M_h}\right), \tag{2.16}$$

and

$$\beta = \frac{1}{w}\left(\frac{M_h - 1}{M_h - M_e}\right)\ln\left(\frac{M_h}{M_e}\right). \tag{2.17}$$

Though the local model is very easy to use, the model is not suitable for devices with thin avalanche regions, where deadspace is significant compared to the avalanche region width [101].

2.3.2 Recurrence Equations

Recurrence equations for avalanche gain parameters were developed in [102]. These equations require probability densities of electrons and holes, $h_e(\xi)$ and $h_h(\xi)$, of impact ionisation path length,ξ. Typically, hard deadspace is assumed in $h_e(\xi)$ and $h_h(\xi)$, such that

$$h_e(\xi) = \begin{cases} 0, & \xi \leq d_e \\ \alpha^* exp(-\alpha^*(\xi - d_e)), & \xi > d_e \end{cases} \tag{2.18}$$

and

$$h_h(\xi) = \begin{cases} 0, & \xi \leq d_h \\ \beta^* exp(-\beta^*(\xi - d_h)), & \xi > d_h \end{cases} \tag{2.19}$$

Hayat et al. [102] provide coupled integrals that can be solved numerically to yield values for gain and excess noise factor for a given voltage. The gain and excess noise factors are calculated using,

$$M(x) = 0.5\left(Z(x) + Y(x)\right), \tag{2.20}$$

and

$$F = \frac{Z_2(x) + 2Z(x)Y(x) + Y_2(x)}{(Z(x) + Y(x))^2}, \tag{2.21}$$

where x is the absorption position of the device and $Z(x)$ is the number of generated electrons and $Y(x)$ is the number of generated holes. $Z(x)$ and $Y(x)$ can be described using,

$$Z(x) = 1 - (W - x)\int_{-\infty}^{x} h_e(\xi)d\xi + \int_{x}^{W} \langle 2Z(\xi) + Y(\xi)\rangle\, h_e(\xi - x)d\xi, \tag{2.22}$$

$$Y(x) = 1 - \int_{-\infty}^{x} h_h(\xi)d\xi + \int_{0}^{x} \langle 2Y(\xi) + Z(\xi)\rangle\, h_h(x - \xi)d\xi, \tag{2.23}$$

where the 1 - integral term represents the contribution if the carrier does not impact ionise and the remaining integral is the contribution if the carrier does impact ionise. $Z_2(x)$ and $Y_2(x)$ are the mean squared values of $Z(x)$ and $Y(x)$, represented as,

$$Z_2(x) = 1 - (W - x)\int_{-\infty}^{x} h_e(\xi)d\xi + \int_{x}^{W} (2Z_2(\xi) + Y_2(\xi) + 4Z(\xi)Y(\xi) + 2Z_2(\xi))h_e(\xi - x)d\xi, \tag{2.24}$$

and

$$Y_2(x) = 1 - \int_{-\infty}^{x} h_h(\xi)d\xi + \int_{0}^{x} (2Y_2(\xi) + Z_2(\xi) + 4Z(\xi)Y(\xi) + 2Y_2(\xi))h_h(x - \xi)d\xi. \tag{2.25}$$

2.3.3 Random Path Length Model

Another non-local model is the random path length model (RPL) [103], which uses the same hard deadspace approximation (Equations (2.18) and (2.19)) as the recurrence equations. The RPL is a monte carlo model that uses the impact ionisation probability densities from the hard dead space to simulate carrier impact ionisation. Random path lengths, l, are calculated using,

$$l = d_e - \frac{\ln(r)}{\alpha^*}, \tag{2.26}$$

for electrons, and,

$$l = d_h - \frac{\ln(r)}{\beta^*}, \tag{2.27}$$

for holes, where r is a random number between 0 and 1. Both recurrence equation and RPL models are favoured for wafer design as the models are quick to run and can be modified for different materials if parameters for the effective ionisation coefficients $\alpha^*(E)$, $\beta^*(E)$, and their corresponding threshold energies have been reported.

2.3.4 Full Band Monte Carlo

Full Band Monte Carlo (FBMC) models, such as [104] for Si or [105] for GaAs, are complex models that simulate the materials band structure and then the material or device properties required. FBMC models often simulate their band structures as a pseudopotential structure [106], which is a method of simulating the band structure without knowing the exact potential applied to the electron from the crystal lattice. Pseudopotentials are calculated from an effective potential due to the crystal lattice, and from the effective potential due to nearby valence electrons. The carrier drift velocities are calculated using deformation potentials (defined as the effective potential experienced by free carriers resulting from local deformation of the crystal lattice), though this is very computationally intensive so in some cases shortcuts are made and saturation velocity approximations will be used to save on computation times [107]. Normally FMBC modelling will consider scattering from ionised impurities; optical and acoustic phonons; and impact ionisation.

2.3.5 Analytical Band Monte Carlo

Analytical Band Monte Carlo (ABMC) models offer a reduced computational intensity cf. FBMC by using analytical models of the electron band structure and phonon dispersion relationships (if simulated) [108]. AMBC models are often used for modelling carrier transport, with models existing for Si, GaAs, GaSb, $Al_xGa_{1-x}As$ (for x of 0-0.35 and 0.5), InAs, and AlAs [109, 110, 111, 112, 113], though can also be used to model impact ionisation [114]. ABMC lends itself to carrier transport modelling as it can simulate the peak in carrier velocities that is observed in several materials, at low electric field strengths, whilst being quicker to run than FBMC.

2.3.6 Simple Monte Carlo

Simple Monte Carlo (SMC) modelling was originally developed by Plimmer et al. [115], in 1999, to simulate GaAs APDs. The method has since been expanded to include parameter sets for $In_{0.48}Ga_{0.52}P$ [116], $In_{0.52}Al_{0.48}As$ [117], InP [118], $Al_xGa_{1-x}As$ [119] and Si [120]. Unlike FBMC and ABMC, SMC uses a single parabolic band approximation and does not calculate the band structure at the start of the simulation. Instead of calculating the band structure, SMC calculates its energy dependent scattering rates using scattering rate equations for intervalley phonon absorption, intervalley phonon emission and impact ionisation (and alloy scattering for $In_{0.52}Al_{0.48}As$). These rate equations account for the rate of phonon absorption, R_{ab}, the rate of phonon emission, R_{em} and the rate of impact

ionisation, R_{ii} (which uses the Keldysh equation [121]),

$$R_{ii} = C_{ii} \left(\frac{E_c - E_{th}}{E_{th}} \right)^{\gamma} \tag{2.28}$$

$$R_{ab} = \frac{N}{\lambda(2N+1)} \sqrt{\frac{2(E_c + \hbar\omega)}{m^*}} \tag{2.29}$$

$$R_{em} = \frac{N+1}{\lambda(2N+1)} \sqrt{\frac{2(E_c - \hbar\omega)}{m^*}} \tag{2.30}$$

where C_{ii} is the coefficient of impact ionisation, E_c is the carrier energy, E_{th} is the impact ionisation threshold energy, γ is the impact ionisation softness factor, and N is the phonon occupation factor as defined in Equation (2.31) (where T is the temperature of the device), λ is the mean free path, $\hbar\omega$ ia the phonon energy, and m^* is the effective carrier mass.

$$N = \left(\exp\left(\frac{\hbar\omega}{k_b T} - 1 \right) \right)^{-1} \tag{2.31}$$

All SMC models reported have been validated against experimental avalanche gain and excess noise factor data over a wide electric field range, corresponding to a wide range of intrinsic region thicknesses from numerous devices.

Chapter 3

Experimental Details

3.1 Current-Voltage Measurements

Current-Voltage (I-V) measurements are usually the first type of measurements carried out in the electrical characterisation of APDs and SPADs. They are performed using a HP 4140B picoammeter, capable of measuring currents down to the picoamp level while providing DC voltages. I-V data are not only used for the comparison of dark currents, but also provide an indication of contact resistance, uniformity across the sample, and breakdown voltage.

The series resistance of the diodes can be calculated from the forward current of the diode, I_f, using,

$$I_f(V) = I_0 \left\{ \exp\left(\frac{qV}{nk_bT}\right) - 1 \right\}, \tag{3.1}$$

where I_0 is the forward current at 0 V applied bias, and n is the ideality factor of the diode. Presence of series resistance, R_s, would reduce the bias applied to the device, so I_f becomes,

$$I_f(V) = I_0 \left\{ \exp\left(\frac{q(V - I_f R_s)}{nk_bT}\right) - 1 \right\}. \tag{3.2}$$

It is challenging to solve for I_f using Equation (3.2) as I_f is on both sides of the equation. Previously [122], the approach has been to rearrange this to form,

$$V = \frac{nk_bT}{q}\ln(I_f) - \frac{nk_bT}{q}\ln(I_0) + I_f R_s, \tag{3.3}$$

using the approximation that $\frac{I_f}{I_0} >> 1$. Equation (3.3) can be differentiated with respect

to I_f, resulting in,

$$\frac{dV}{dI_f} = \frac{nk_bT}{qI_f} + R_s = \frac{nk_bT + qI_fR_s}{qI_f}. \tag{3.4}$$

Plotting $I_f\frac{dV}{dI_f}$ against I_f would yield R_s from the gradient (qR_s). However, in practice there is not always a linear relationship between $I_f\frac{dV}{dI_f}$ and I_f, preventing the extraction of R_s. An alternative approach, is to rearrange Equation (3.2), so that I_f only appears on one side of the equation. On the condition that $R_s \neq 0$, $\frac{nk_bT}{q} \neq 0$ and $I_0 \neq 0$, this can take the form of,

$$I_f(V) = \frac{nk_bT}{qR_s}W_0\left(\frac{qI_0R_s}{nk_bT}\exp\left[\frac{q}{nk_bT}\left(I_0R_s + V\right)\right]\right) - I_0, \tag{3.5}$$

where $W_0()$ is the zeroth solution to the Lambert W function. The zeroth solution is used as the other solutions, making up the solution set, are complex solutions which are not applicable in this use case of the Lambert W function.

Epitaxially grown layers can result in non-uniformity due to small variations in layer thicknesses. Non-uniformity can also be caused by variations in etching processes, misalignment between fabricated layers, or thin dielectric layers (when dielectric layers separate P/N doped materials or their corresponding contacts).

I-V measurements also allow for a rough indication of the breakdown voltage of the sample, which is further explored by performing avalanche multiplication measurements. If the sample shows a large variation in breakdown voltage then this indicates a problem with the sample, such as premature edge breakdown, or non-uniform layer thicknesses across the sample.

Fabricating diodes using the research group's new-pin mask set allows for this measurement to be carried out for up to 4 different sizes of circular mesa (25, 50, 100, 200 µm radius devices or 35, 60, 110, 210 µm radius depending on which etch pattern was used). Normalising the measured current of the different sizes against the cross-sectional area of the device and the perimeter of the device allows for the exploration of leakage mechanisms. If the measured currents scale uniformly with device area then bulk leakage is dominant in the device, which is ideal. If the currents scale with perimeter then that means that the current is surface leakage dominated.

3.2 Capacitance-Voltage Measurements

Data from Capacitance-Voltage (C-V) measurements enable the estimation of doping profiles, presence of lateral etching, and punch-through voltage (for SAM APDs). The C-V measurements in this thesis were performed using a HP 4275A multi-frequency LCR meter. The LCR meter applies a 50 mV, 1 MHz sinusoidal signal superimposed on a DC bias to the device under test, DUT. It measures impedance (L), which is then used to calculate the capacitance using a parallel resistance (R)- capacitance (C) assumption.

C-V measurements can also be used to determine how closely the electrically active doping profile reflects the design specifications. To fit the doping profile to C-V data a 1-dimensional Poisson equation solver is used, as described in appendix B. Capacitance should scale with device area, so lateral etching is assessed by comparing data from different sized devices.

C-V fitting is required for the development of material parameter sets for models, such as the SMC model. The C-V extracted doping profile is used to simulate the electric field conditions of the device under different bias conditions, to then simulate APD/SPAD characteristics.

3.3 Avalanche Multiplication Measurements

Avalanche multiplication measurements are performed using the setup shown schematically in Figure 3.1. The sample stage holds two probe positioners which are used to make electrical contacts to the DUT via its p- and n-contacts. Reverse bias to the DUT is supplied by a Keithley 236 source measure unit, SMU. The DUT is exposed to laser light that is fibre coupled and collimated into the setup. The laser light is chopped using an optical chopper controlled by a Lock-In Amplifier, LIA, (Chop-In Opus 1, Boston Electronics). The light is chopped at 180 Hz (chosen to allow extra filtering from the LIA available below 200 Hz and avoid stray light that tends to be at multiples of the mains frequency). The LIA is used to measure the photocurrent from the DUT, measured across a sense resistor (normally 100 Ω to avoid a significant voltage drop across the sense resistor).

The laser light reaches the DUT by being focused through an objective lens, mounted to an optical turret. A white light source and USB camera are also connected, enabling the sample to be aligned, and focused to the objective lens by adjusting the position of the sample stage.

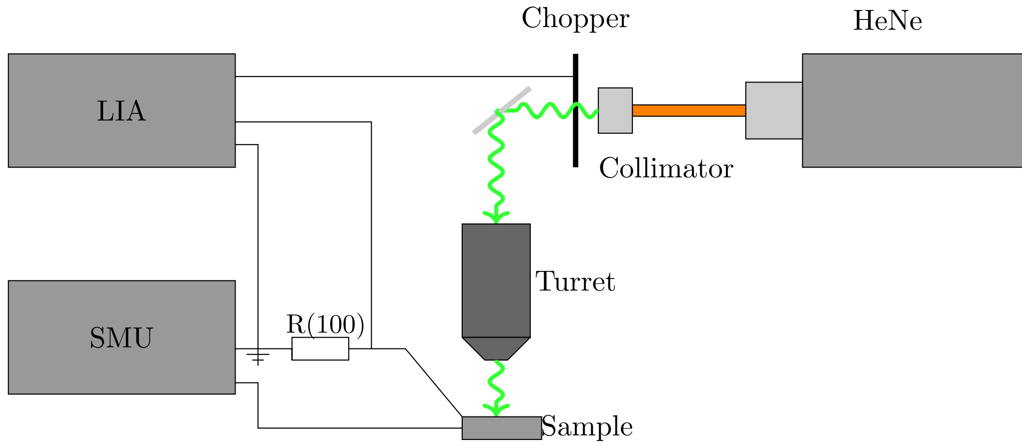

Figure 3.1: Schematic of Avalanche multiplication setup. Light from a Helium-Neon (HeNe) laser is chopped at 180 Hz by an Lock-In Amplifier (LIA) controlled optical chopper before being focused, through an objective mounted in an objective turret, on to a reverse biased APD. The diode is reverse biased by an SMU and the LIA is able to measure a voltage proportional to avalanche gain across the sense resistor (typically 100 Ω).

3.4 Dark Count Rate Measurements

Two different methods were used to carry out dark count rate (DCR) measurements. The first (preferred) method is only suitable for packaged devices, whilst the second method can be used for bare dies (unpackaged samples). The first method is performed in a black box setup using a capacitive quenching circuit, CQC, designed by Dr Simon Dimler [123]. This CQC operates the SPAD in gated mode by having an input pulse and a input bias. The magnitude of the inputs are chosen so that the SPAD is above its breakdown voltage during the pulse. The pulse applied to the SPAD causes transient artefacts superimposed on the output signal due to capacitance of the device, as shown in Figure 3.2. The CQC includes transient cancellation circuitry to remove the capacitive transients [44], which uses an adjustable capacitor to mimic the transients generated across the SPAD which are then subtracted from the SPAD output signal using a differential amplifier circuit. The CQC board also includes an active probe to measure the height of the pulses applied to the device. Since the active probe provides a value proportional to the pulse applied to the SPAD, careful calibration of the active probe is carried out using the process described in [124].

The packaged SPAD is placed into a copper holder cooled by a peltier element controlled by a Keithley 2510 TEC Source Meter. The cathode and anode of the SPAD are then placed into sockets on the CQC PCB. The output of the CQC is inspected using a 2 GHz Lecroy Waverunner 204Xi Oscilloscope whilst the DUT is pulsed using an Agilent 81101A Pulse Generator and the adjustable capacitor is trimmed to minimise the transients at this pulse height. The output of the CQC is then attached to a discriminator circuit with an

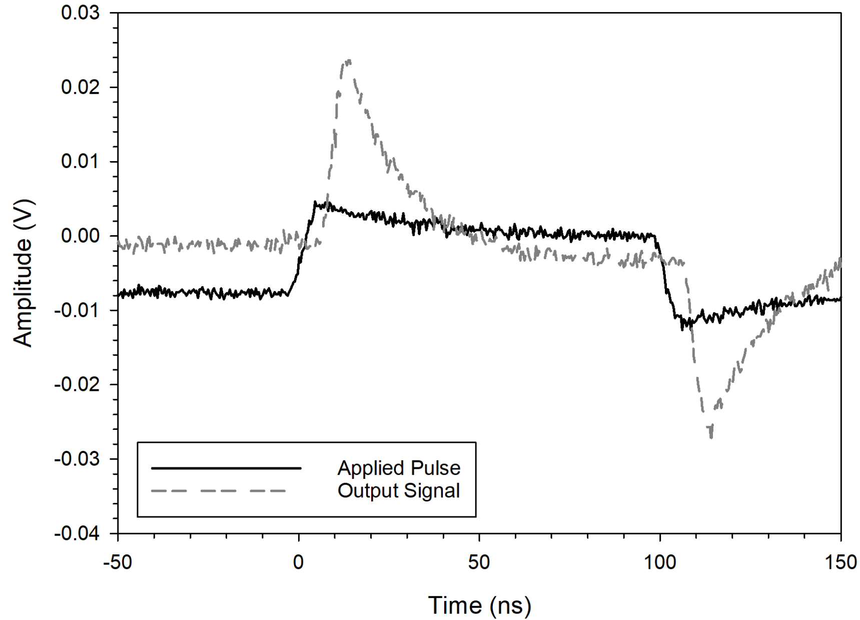

Figure 3.2: Schematic showing the presence of capacitive transients on the SPAD signal (dashed line) from an applied electrical pulse (solid line).

adjustable threshold level. The discriminator will trigger a nuclear instrumentation module (NIM) logic output which will in turn trigger a counter (Canberra Dual Counter/Timer 512), if the threshold has been exceeded. To set the threshold of the discriminator, the constant DC bias of the SPAD is set so DC bias+pulse is less than the breakdown voltage of the SPAD, the threshold level is increased from 0 until the counter is no longer being triggered by the residual transients. The setup is now ready to measure DCR.

The second method to measure DCR uses a Janis ST-500 low-temperature probe station, alongside electronic and optical equipment, as illustrated in Figure 3.3. To cancel/reduce the capacitive transients the electrical pulse from the pulse generator is applied to a 50 Ω matched signal splitter, with one output fed to a bias tee (Picosecond Pulse Labs 5530A), to be superimposed onto a DC voltage before being applied to the SPAD and the other output is connected to a parallel resistor capacitor (0.5 pF) circuit. The capacitor and resistor values (0.5 pF, 200 kΩ) produce similar transients to those produced by the SPAD. The mimicked transient is then attenuated using a variable attenuator (HP 355D), so its amplitudes match the SPAD transients. The SPAD transients and the mimicked transients are synchronised in time with ∼3 m of BNC cable. A differential amplifier then subtracts the mimicked transient from the SPAD output, producing an output signal dominated by avalanche pulse rather than capacitive transients. The differential amplifier output is fed to a discriminator which generates a NIM pulse if the signal passes above a user-adjustable threshold level. The discriminator threshold is adjusted using a potentiometer so that the

noise of the output does not trigger the discriminator. To perform this adjustment, the SPAD is reverse biased with combined DC and pulse that remain below its avalanche breakdown voltage, and the potentiometer is adjusted to raise the threshold above the noise floor at this point. During the measurement the threshold may be further increased to avoid double counting. Double counting is triggered if the noise in an avalanche pulses drops below and then comes back above the threshold level, causing multiple counts to be triggered. The CQC setup is simpler than using the Janis ST-500, as the signal splitter, bias tee, attenuator, and differential amplifier are not required as they are already integrated onto the CQC PCB.

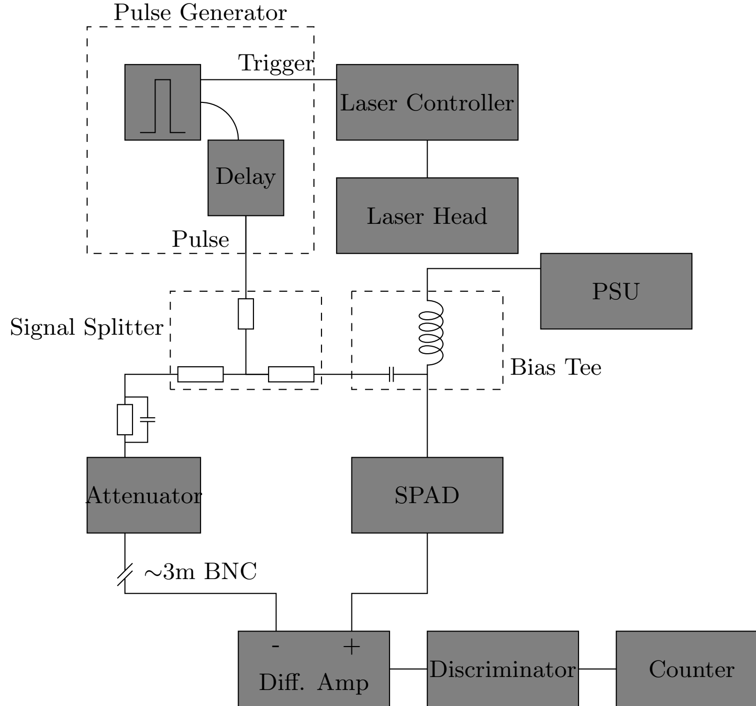

Figure 3.3: Schematic of electrical connections for dark and photon counting measurements using the ST-500 Janis probe station.

DCR measurements are taken using a fixed pulse height while modifying the DC bias, supplied by an Keithley 6517A Electrometer, to raise and lower the total applied bias above the SPADs' breakdown voltage. DCR measurements are recorded typically for 30 seconds or longer and this measurement is repeated multiple times, to build up a statistically significant result for DCR. The recorded count is normalised against the measurement duration (T_m), the pulse duration (T_p) and the pulse frequency (f) to produce comparable results to other devices that account for the effective on time of the measurement. This normalisation is described as,

$$\mathrm{DCR} = \frac{N}{T_m T_p f}, \tag{3.6}$$

where N is the average number of dark counts recorded for the measurements taken for a given set of test conditions.

3.5 Photon Counting

After performing dark count measurements, the next step in SPAD characterisation is to perform photon counting measurements, which can be performed in both of the setups used for dark count measurements. In both setups a PIL063SM PiLas pulsed laser head, with a centre wavelength of 634.1 nm and a spectral width of 3.43 nm, from Advanced Laser Diode Systems (recently acquired by NKT Photonics), is used along with a compatible controller. With the black box setup, the laser pulses are fibre coupled into the box where the optical spot is aligned to and focused on the DUT active area by adjusting motorised x, y, z stages. The x, y, z stage holds a small optical breadboard that contains the pulsed fibre input, a white light source and USB camera, all aligned to be on the same beam path. The white light source is used to illuminate the device and allow the initial alignment of the stage above the SPAD. Whilst, in the Janis ST-500 the pulse is also fibre coupled, however the fibre itself is directly aligned with the SPAD active area using a camera. For both systems, to aid alignment a 633 nm HeNe laser is normally used first to align rather than the pulsed laser, as the USB camera are unable to detect the very short pulses (even at high pump powers).

When using the pulsed laser head, the laser pulse needs to be synchronised so that it arrives at the SPAD while the SPAD is gated above its breakdown voltage. The pulse generators built-in delay feature is used to delay the generation of the over-bias pulse so it is produced a defined time after the output trigger synchronising the pulses. To set this delay time, the SPAD being tested is operated in the gated mode and the pulse laser is operated with a high tune power (typically 50%). By capturing the output of the SPAD using an oscilloscope, the delay from the output trigger can be adjusted until the large spike caused by the large incoming optical pulse can be seen in the output of the SPAD. Using the Agilent 81101A pulse generator and the PiLas pulsed laser controller with the 633 nm laser head the delay time was found to vary daily between 80-130 ns. This variation, is assumed to be caused by the PiLas laser controller, but was not an issue for the measurements presented in this work, as the required delay remained stable during operation only being prone to change if the controller was turned off.

3.5.1 Calculating SPDE

Within the SPAD research community, there are currently three main equations used to calculate the single photon detection efficiency (SPDE):

$$SPDE = \frac{1}{n}\ln\left(\frac{1 - P_d}{1 - P_t}\right), \quad (3.7)$$

$$SPDE = \frac{N_{total} - N_{dark}}{N_{photon}}, \quad (3.8)$$

$$SPDE = \frac{P_t - P_d}{P_{ph}}, \quad (3.9)$$

where n is the average number of photons per pulse and $N_{photons}$ is the total number of photons over the measurement duration. N_{total} is the total number of counts from either photons or dark carriers with P_t representing the corresponding count probability. N_{dark} and P_d are the dark carrier only equivalents to N_{total} and P_t. P_{ph}, discussed later in this section, is the probability of a pulse containing photons and defined as,

$$P_{ph} = 1 - e^{-n}. \quad (3.10)$$

Equation (3.7) is a rearrangement of eq. 1 from [125] which is based on the earlier works of Levine et al. [126, 127, 128, 129]. Equation (3.8) is from the verbal definition of SPDE used in [48, 49], while Equation (3.9) is a modification of Equation (3.8) that accounts for the Poissonian nature of a photon source [130]. Unfortunately, these equations can lead to very different values of SPDE. The difference in these equations will be shown to be due to two slightly different definitions of SPDE. These differing definitions are:

1. The SPDE is the photon detection efficiency where in a gate period events can be counted from a photon detection, a dark trigger, or both. (Inclusive probabilities)
2. The SPDE is the photon detection efficiency where in a gate period a single count is possible. This count can be from a photon or dark event. (Exclusive probabilities)

The first definition of SPDE implies that a SPAD can simultaneously detect a dark triggered avalanche event as well as a photon triggered avalanche event, which may be true for linear mode operated devices, capable of detecting low photon pulses, however is not true for a Geiger mode device. The second definition of SPDE, states that only one event may be detected at a time, which is true for Geiger mode operated devices. In the Geiger mode, a device either detects an event or it does not detect an event, there are no additional

levels (unlike linear mode, where the amplitude of the pulse may be used to quantify detection). This means that in the Geiger mode, as soon as a detection event starts, the device becomes blind to any subsequent events, until the device has been successfully quenched.

To check these equations, consider a SPAD being operated in gated mode. Assuming the dark carriers follow a poissonian distribution [44], then the probability that there are k carriers in a single gate, when the average carriers in a gate is n_d can be represented as

$$P(n_d, k) = \frac{n_d^k \times e^{-n_d}}{k!}. \tag{3.11}$$

For each k value the probability of a dark count in the SPAD is $1 - (1 - P_b)^k$, where P_b is the breakdown probability of the SPAD. By summing all k values, the total dark count probability, P_d can be determined:

$$P_d = \sum_{k=0}^{\infty} \left\{ \left[1 - (1 - P_b)^k \right] \times \frac{n_d^k \times e^{-n_d}}{k!} \right\}, \tag{3.12}$$

$$P_d = \sum_{k=0}^{\infty} \frac{n_d^k \times e^{-n_d}}{k!} - \sum_{k=0}^{\infty} \left\{ (1 - P_b)^k \times \frac{n_d^k \times e^{-n_d}}{k!} \right\}. \tag{3.13}$$

Rearranging and letting $x = n_d(1 - P_b)$

$$P_d = 1 - e^{-n_d \times P_b} \sum_{k=0}^{\infty} \frac{x^k e^{-x}}{k!}, \tag{3.14}$$

as $\sum_{k=0}^{\infty} \frac{x^k e^{-x}}{k!}$ converges to unity, P_d can be expressed as

$$P_d = 1 - e^{-n_d \times P_b} \tag{3.15}$$

Using a similar approach, the probability of detecting a photon, P_p can be represented as,

$$P_p = 1 - e^{-n \times SPDE}. \tag{3.16}$$

Inclusive Probabilities

In the inclusive case the total count probability, P_t, is defined as

$$P_t = P_p + P_d - P_p P_d \tag{3.17}$$

where the subtraction is required to avoid double counting the overlapping probabilities.

This leads to,

$$1 - P_t = (1 - P_p)(1 - P_d), \tag{3.18}$$

$$e^{-n \times SPDE} = \frac{1 - P_t}{1 - P_d}, \tag{3.19}$$

$$SPDE = \frac{1}{n} \ln\left(\frac{1 - P_d}{1 - P_t}\right) \tag{3.20}$$

Resulting in the first definition of SPDE (Equation (3.7)).

Exclusive Probabilities

In the exclusive case the total count probability can be expressed as,

$$P_t = P_d + P_p, \tag{3.21}$$

implying

$$P_p = P_t - P_d, \tag{3.22}$$

$$e^{-n \times SPDE} = 1 + P_d - P_t, \tag{3.23}$$

$$-n \times SPDE = \ln(1 + P_d - P_t). \tag{3.24}$$

Using the taylor expansion for $ln(1 + x) \sim x$ where $x \ll 1$,

$$-n \times SPDE = P_d - P_t, \tag{3.25}$$

$$SPDE = \frac{P_t - P_d}{n} \tag{3.26}$$

Which can be written as

$$SPDE = \frac{N_t/N_{gate} - N_d/N_{gate}}{n} = \frac{N_t - N_d}{N_{gate} n}. \tag{3.27}$$

Where gate is the number of gates passed for the counts of N_d and N_t. $N_{gate}n$ results in the number of photons measured, $N_{photons}$, resulting in the derivation of the second equation for SPDE (Equation (3.8)).

Currently, one of the assumptions made is that the photon source either emits a photon or not. This is not necessarily true, as there is no reason the photon source can not emit multiple photons whilst still giving the average of n. If the photon source emission follows

a Poisson distribution, then:

$$P(\text{No Photons}) = \frac{e^{-n}n^0}{0!} = e^{-n}, \tag{3.28}$$

suggesting

$$P(Photons) = 1 - e^{-n}, \tag{3.29}$$

which is the same form as P_{ph} (Equation (3.10)). From this, using the second SPDE equation (Equation (3.8)), the third SPDE equation (Equation (3.9))

$$SPDE = \frac{P_{total} - P_d}{P_{ph}} \tag{3.30}$$

can be shown.

3.6 Silicon Device Fabrication

A key part of the work for this thesis has been the device fabrication of Silicon devices within the cleanroom facilities at the University of Sheffield and through external contractors where equipment was not available internally. This section covers the techniques used during the fabrication process.

3.6.1 Wafer cleaning

Normally with the III-V semiconductors, a 3 solvent cleaning process is used with n-butyl acetate, acetone, and isopropyl alcohol (IPA). The first solvent used is the n-butyl acetate which will remove most surface contaminants, however it is prone to leaving a residue on the surface of the sample being cleaned. The second solvent, acetone, will remove the residue from the n-butyl acetate and some contaminants that the n-butyl acetate struggles to remove. Unfortunately the acetone evaporates very quickly from the sample leaving drying stains on the surface, which can interfere with further processing. The third solvent, IPA, will clear the acetone drying stains and also has a much slower evaporation rate meaning that it can be blown off using a nitrogen gas gun before it evaporates, so no drying marks are left on the sample.

Working with silicon, the 3 solvent cleaning process can be used but there is also the Radio Corporation of America (RCA) cleaning process [131] based on hydrogen peroxide solutions. The RCA process is a 2 stage cleaning process comprising of standard clean 1 (SC-1) and standard clean 2 (SC-2). SC-1 is used to remove organic and particulate

contamination from the wafer surface, whilst SC-2 targets metallic contamination. When working with the RCA process pure quartz/silica glassware should be used rather than standard lab glass (borosilicate), because SC-1 etches the native oxide from the surface of the silicon and redeposits new oxide on the surface. Silica based glassware will therefore be slightly etched by SC-1, leading to the release of boron from borosilicate glassware potentially causing inadvertent doping of the silicon surface. The original recipe for SC-1 was a 1:1:5 solution of hydrogen peroxide: aqueous ammonia: deionized water, though in this work a diluted 1:1:20 solution [132] was used. The original SC-2 recipe used a 1:1:6 solution of hydrogen peroxide: hydrochloric acid: deionized water, again diluted to a 1:1:20 solution for this work [133].

3.6.2 Photolithography

Photolithography plays an important role in any semiconductor fabrication process. For this work two main photoresists have been used, these were the Dow resists SPR 220 and SPR 350 (more information about these resists is available from their data sheets [134, 135]). When using these photoresists, first the samples were cleaned, then baked at 100 °C for at least 1 minute to make sure that any solvent residue has evaporated, preventing left over trace solvents from interfering with the photolithography process. Next, the sample was mounted on a spin coater, coated with the resist of choice then spun typically at 4000 rpm for 30 seconds. The spin speed and duration can be adjusted to change the thickness of the resulting photoresist film. Spinning the sample spreads the photoresist over the sample allowing for an even coating and uniform thickness. After resist deposition the sample was baked on a hotplate at 100 °C, for 1 minute for SPR 350, or 1.5 minutes for SPR 220 to cure the photoresist. If the sample was coated with silicon oxide or silicon nitride then an adhesion promoter, such as hexamethyldisilazane (HMDS) should be used, applied before the photoresist. HMDS was applied to the surface with a disposable pipette, allowed to soak for 30 seconds, then spun at 4000 rpm for 30 seconds before proceeding with the resist deposition. The spin coating of HMDS is not recommended by the manufacturer, instead the recommended procedure is to heat the sample in an oven at 100 °C whilst flowing gaseous nitrogen, that has been bubbled through HMDS, across the sample. The HMDS improves the wettability of the surface by reacting with the dangling OH bonds, to leave CH_3 bonds, nitrogen and hydrogen gas [136], as shown in Figure 3.4.

After resist deposition and baking, the sample is then placed in a mask aligner with a 300 nm UV source, and the appropriate mask plate is loaded into the mask aligner. The

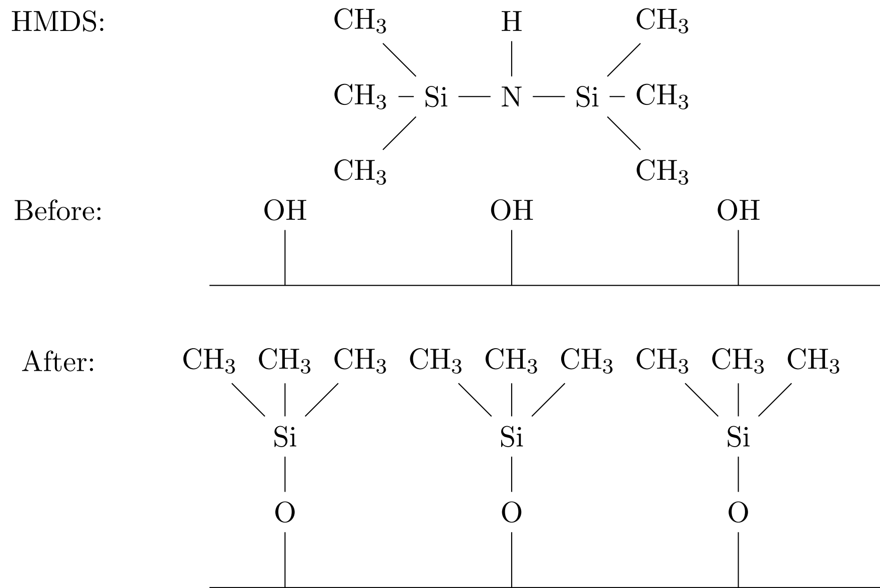

Figure 3.4: Schematic showing the structure of HMDS (top) then showing how HMDS acts as an adhesion promoter on silicon oxide and silicon nitride.

mask plates are 5x5" quartz plates, with a chromium pattern. These mask plates are used in a contact printing setup, where the sample surface is brought into direct contact with the mask plate. After aligning the sample to the mask plate, using the sample stages 4 degrees of freedom (x,y,z, and rotational), the sample is exposed to the UV source for the desired exposure time. Under exposing the sample will result in a pattern that is not fully formed and could still contain photoresist in the exposed feature areas after development, whilst over exposure can result in the broadening of features with the potential loss of detail in the small feature areas after development. Once the exposure was completed the sample was removed from the mask aligner and submerged in the corresponding developer for the SPR resists (MF26A) for 1 minute for SPR 350, or 1.5 minutes for SPR 220, to remove the exposed positive resist (working with negative resists, the non-exposed resist would be removed), then rinsed with deionised water. The alignment and exposure quality was checked under a microscope, to verify the alignment and confirm the exposure marks were correctly exposed. The exposure marks are quality control features added to the mask plates, consisting of a bunch of "fingers" separated by 5-10 µm (or less if the minimum feature size is smaller). Inspecting these fingers can help identify under/over exposure. Depending on the severity of the under/over exposure the sample may need to be stripped in acetone and IPA before repeating the photolithography process using a compensated exposure time.

3.6.3 Ion Implantation

Ion implantation is one of the key ways of forming localised doping regions and heavily graded doping profiles. In this work ion implantation was performed externally by Cutting Edge Ions, Anaheim (CA), USA. For the ion implantation the conditions were first simulated using the Stopping Ranges of Ions in Matter (SRIM) software [137] and Sentaurus TCAD [138]. The samples were patterned with either, SPR220 photoresist, silicon dioxide, or an aluminium pattern depending upon the beam energy and dose of the implant. Originally 190 keV high dose Boron implants were performed using an SPR 220 mask, however the SPR 220 was completely carbonised and could not be removed from the samples (despite several hours of HF treatment to remove the screening oxide and undercut the resist). For the second attempt at 190 keV boron implants, an aluminium mask was used, which was easily removed. The first choice of metal masking material was nickel, as it is very easy to remove with nitric acid, however Cutting Edge Ions requested nickel was not used to avoid chamber contamination from sputtering.

3.6.4 Oxidation

Two different types of oxidation have been used as part of the fabrication in this work; Plasma Enhanced Chemical Vapour Deposition (PECVD) and wet oxide growth. PECVD is a plasma process that was used to deposit SiO_2 onto the sample, but can also be used to deposit SiN and poly-Si (amorphous Si). PECVD deposition of SiO_2 was performed using a blend of SiH_4:N_2O:N_2. From this gas blend the SiH_4 and the N_2O are used to form SiO_2 for deposition, whilst the additional N_2 is used as a diluting gas [139]. Diluting gasses are used in PECVD to manipulate the available gas interactions. In PECVD the SiO_2 is deposited onto the surface, whereas with wet oxide growth the oxide is grown from the surface. In wet oxide growth, the samples are heated in an oxidation chamber and water vapour is added to the chamber. The water vapour reacts with the surface of the sample, leading to the growth of SiO_2 with the Si provided by the sample, resulting in consumption of the surface of the sample. For SiO_2 to be grown using a wet oxide process 46% of the growth thickness will be below the original surface of the sample surface, for example if 1 µm of SiO_2 oxide is grown then the top 0.46 µm of the Si will be consumed. Due to the high temperature required for wet oxide growth ($\sim$ 1000°C), the process can also be used for simultaneous annealing of samples.

3.6.5 Dry Etching

For this work, dry etching was used to pattern Si and SiO_2. Dry etching uses a plasma tool with different gas recipes for etching different materials. Unlike wet chemical etching, which produces angled side walls, dry etching is capable of producing vertical side walls making it ideal for the etching of tightly packed features, and deep features. An example of a vertical side wall etch can be seen in Figure 3.5. To etch Si, a blend of SF_6:O_2:Ar was used, while a blend of CHF_3:Ar was used to etch SiO_2. In the Si etch, the SF_6 is a source of fluorine radicals that perform the etching, the oxygen passivates the side walls, and the argon acts to reduce the formation of black silicon [140]. Black Silicon is rough Silicon where the surface has become rough enough to change the refractive index of the material, a scanning electron microscope (SEM) image of a black Silicon surface is shown in Figure 3.6. The example shows a cross section of a etched mesa where black Si has formed on both the etch surface and the top of the mesa. Black Si formation on the top of the mesa was due to the sample being very over etched, to the point that the photoresist at the top of the mesa was completely etched away. In contrast to the Si etch, the SiO_2 etch uses CHF_3 as its source of fluorine radicals and argon to increase the anisotropic etching (non-uniformity of etching, preferentially in the vertical plane out of the wafer rather than across the wafer), due to the large size of argon atoms cf. fluorine atoms (mass of 39.95u cf. 19.00u, where u is the unified atomic mass unit).

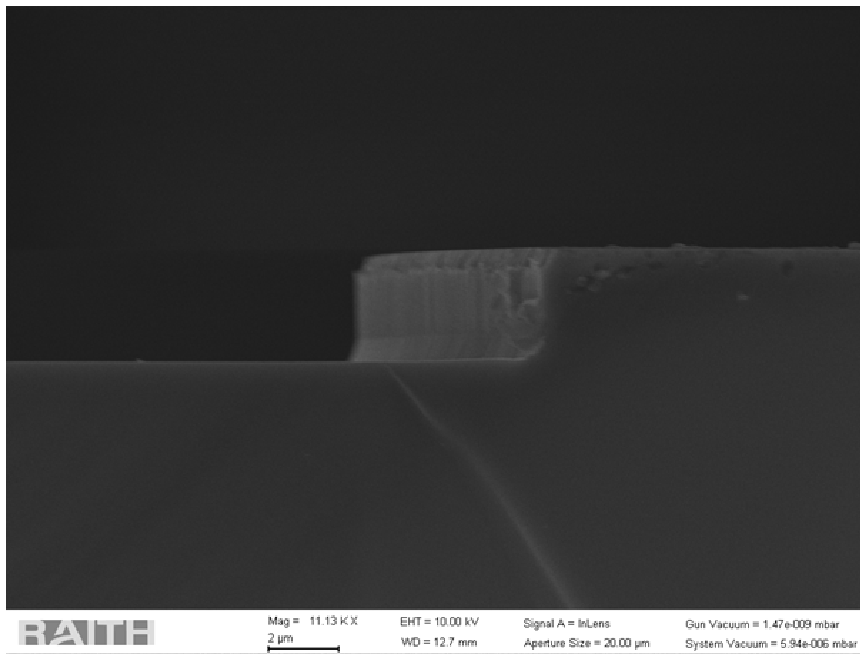

Figure 3.5: Example SEM image, showing the side wall profile of dry etched Si

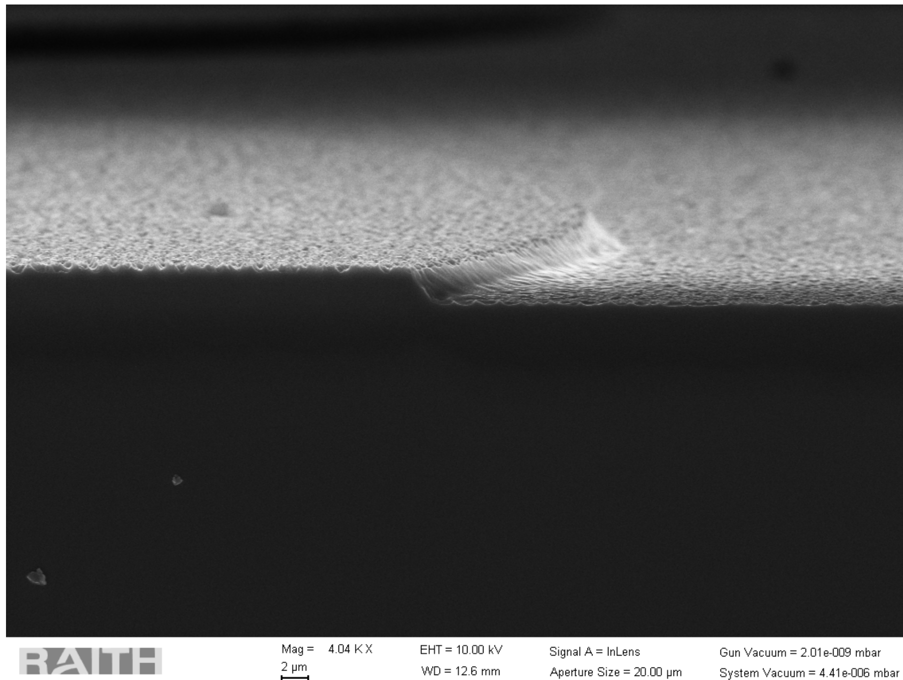

Figure 3.6: Example SEM image, showing surface roughness responsible for black silicon.

A Reactive-Ion etching (RIE) system was used to etch the SiO_2 and an Inductively coupled plasma (ICP)-RIE system was used to etch the Si. A RIE system has one RF generator, which is used to create the plasma and bombard the sample being etched with radicals. While, an ICP-RIE system, contains two RF sources, the ICP source, which is used to generate the plasma, and the normal RF source which is used to bombard the sample. To achieve a uniform plasma across the sample, high density plasmas are required (requiring high power RF sources), however smooth etched surfaces are also required (requiring low RF power to minimise the crater size from each radical). These two factors are contradictory, meaning that a standard RIE system can often struggle to produce large scale uniformity, unlike an ICP-RIE system. The ICP-RIE was used to etch the Si at 20 nm/min, which provided a uniform etch with near vertical side walls. However, this results in very long process times especially when etching 3.5 µm high Si features! The etch rate of SiO_2 was similar in the ICP-RIE, however side wall uniformity was less critical, so an RIE process, capable of etching SiO_2 at 45 nm/min, was used instead.

3.6.6 Metal Deposition

In this work, 99.99% aluminium was used as the contact material, deposited using thermal evaporators. To achieve the desired 800 nm thick contacts, 4 small tungsten coils each loaded with ~20 cm of 0.5 mm diameter Al wire were used in a thermal evaporator. The coils were loaded 12 cm above the sample to avoid heat radiated from the evaporation coils damaging the photoresist, which would significantly hinder the lift-off process. Due to equipment limitations the loaded tungsten coils were placed in parallel in the evaporator so 2 coils were being evaporated together, then the next set. This method worked to achieve the contact thickness but meant that if one of the coils in the pair failed then the other coil would also break after radiating a significant amount of heat into the evaporator, with the potential to burn the sample even from the 12 cm height.

After evaporating aluminium onto the samples, the samples were placed in Dimethyl Sulfoxide (DMSO). The DMSO acts as a photoresist stripper, dissolving the photoresist pattern on the sample, removing any aluminium coating the photoresist in the process, leaving just the contacts on the samples. Originally acetone was tried for the lift-off process, however this was found to be a very slow process, possibly due to the design of the mask. As using acetone as the lift-off agent was quite a slow process, some of the aluminium that had been undercut had time to sag down and touch the unprotected sample, preventing its removal. The EKC830 and Microposit Remover 1165 [141] positive

resist strippers were also tried, however they both contain n-methyl-2-pyrrolidone which we found etched aluminium forming a blistered surface.

3.6.7 Secondary Ion Mass Spectrometry

Secondary Ions Mass Spectrometry (SIMS) has been performed by Loughborough Surface Analysis (LSA). SIMS is performed to build up accurate information about the doping profile of the samples. The staff at LSA load the sample into one of their SIMS machines, then fire a focused beam of ${O_2}^+$ or Cs^+ at the sample. This beam acts as a mill to slowly etch away the sample and the exhaust material is fed through a mass spectrometer which allows the LSA staff to determine what materials are present in the sample. The signal readout of the SIMS system is initially doping (arb.) versus time (arb.), which needs to be converted into more usable units. To convert time (arb.) into depth (nm) a profilometer is used to measure the total etched depth then a direct conversion is performed. This conversion is valid as during the SIMS measurement the ion beam energy and flux are kept constant to maintain a constant etch rate. A doping calibration run is performed using a reference sample with a known doping concentration. It is important to note that the SIMS profile contains all the material in the sample, but not all of the material may be electrically active. Combining the doping profile from SIMS with experimentally measured C-V data will provide a lot of detail about the sample structure.

Chapter 4

A Simple Monte Carlo Model for Silicon SPADs

4.1 Introduction & existing State of SMC

This chapter describes how the SMC approach has been improved to enable simulation of Silicon SPADs using the Si parameter set from Zhou et al. [120]. The work included a new Simple Monte Carlo Simulator [142] with the detailed implementation outlined in appendix A. This new SMC simulator was used to explore the effects of doping orientation (p-topped or n-topped SAM-SPAD structures) on Si SPAD timing characteristics [64].

SMC offers many advantages for device simulation compared to other simulation techniques. The main simple simulation techniques are Random Path Length (RPL) models [143] and models that use the recurrence equations from McIntyre [144] or Hayat [102]. Though these simple methods offer significantly faster computation times, when compared to the SMC model they often struggle with accuracy for device structures with rapidly changing electric field profiles [145], which are often found in Separate Absorption Multiplication (SAM)-APDs and SPAD designs. The lack of accuracy surrounding non-uniform electric field profiles also leads to the use of the saturation drift velocity approximation (i.e. the carriers are assumed to be travelling at their saturation velocities at all times in the device).

Two other monte carlo modelling methods are Full-Band Monte Carlo (FBMC) [146] and Analytical Band Monte Carlo (ABMC) [108]. As with SMC, FBMC and AMBC are capable of simulating device structures with rapidly changing electric field profiles.

However, both FMBC and ABMC suffer from drawbacks that make them unsuitable for practical simulation of device characteristics. FMBC models simulate the entire band structure of the material which can lead to very accurate models, but it also leads to very slow, computationally intensive, models which are not ideal when simulating a number of structures to compare performance. A less computationally intensive alternative is the ABMC which uses analytical equations to approximate the band structure of the material. Whilst ABMC requires less computational resources than FBMC, both models require various experimental input parameters which can be very hard to obtain experimentally. One of the key advantages of FBMC and ABMC over SMC is that SMC is unable to produce the low-field peaks observed in the electric field dependent carrier drift velocities [147]. However, this low-field peak in drift velocity is not in the relevant electric-field range for APD or SPAD operation. The expanded SMC model, of this work, focuses on the timing characteristics caused by the stochastic nature of the impact ionisation process. Other mechanisms contributing to the timing characteristics are described in section 2.2.4.

4.2 SMC Simulation flow

The SMC process used in this thesis is based upon the rate equations for phonon absorption, phonon emission, and impact ionisation, as discussed in section 2.3.6. The rate equations are repeated below for convenience,

$$R_{ii} = C_{ii}\left(\frac{E_c - E_{th}}{E_{th}}\right)^{\gamma} \tag{2.28 repeated}$$

$$R_{ab} = \frac{N}{\lambda(2N+1)}\sqrt{\frac{2(E_c + \hbar\omega)}{m^*}} \tag{2.29 repeated}$$

$$R_{em} = \frac{N+1}{\lambda(2N+1)}\sqrt{\frac{2(E_c - \hbar\omega)}{m^*}}. \tag{2.30 repeated}$$

Each SMC material parameter set adjusts these rate equations for the material being simulated. As part of the work for this thesis a new simulator had to be written, from scratch, to re-establish the existing position of SMC modelling within our research group before the model could be expanded upon as outlined later in section 4.3. The structure of the new simulator, and an outline of what each class does, is detailed in appendix A. Here a process flow of the device properties simulation is outlined which follows the flow chart shown in Figure 4.1.

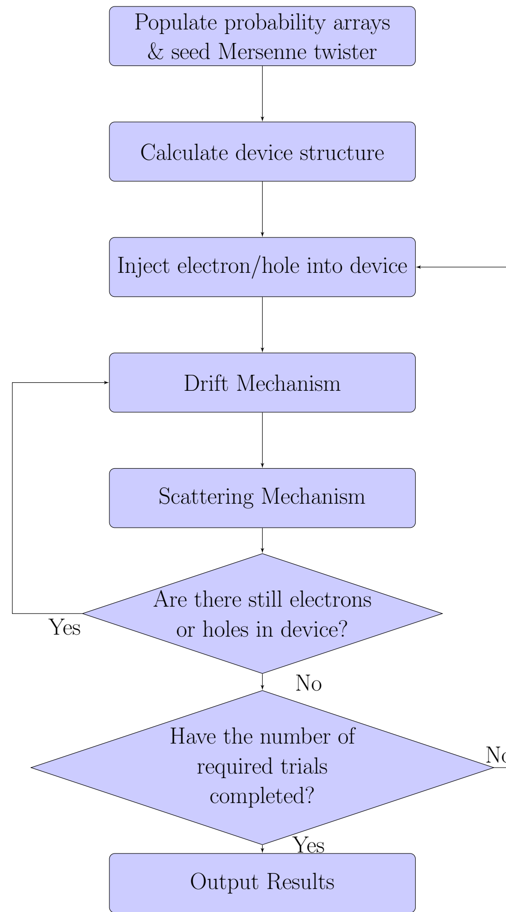

Figure 4.1: Flow chart of device properties function for the Simple Monte Carlo Simulator

To start the simulation process, after the doping structure, bias list, and user parameters, have been read in, the relative energy dependent probabilities of each type of interaction event have to be calculated. In the simulator there are 4 different possible events: phonon absorption, phonon emission, impact ionisation, and not interacting. The rate equations, Equations (2.28) to (2.30), are calculated in 1 meV increments from 0 to 6 eV. To convert these rates into relative probabilities, the rate of each interaction at 6 eV are summed, then assuming that there is a 100% chance of interaction at a carrier energy of 6 eV the relative probabilities are calculated by dividing through by the total interaction rate. This process is carried out separately for electrons and holes.

Next the pseudo random number generator is initialised, in this simulator a Mersenne Twister was used. For debugging purposes the Mersenne Twister seed was kept constant. Following this the electric field profile is calculated for the first bias on the bias list using the read in doping profile and the infinite region electric field solver outlined in appendix B.2. The electric field solver uses the 1D Poisson equation, shown in Equation (4.1),

$$\frac{dE}{dx} = \frac{qXN}{\epsilon}, \tag{4.1}$$

where dE/dx is the electric field gradient within a doped region that has a doping concentration of N, a thickness of X, and a relative permittivity of ϵ_r where $\epsilon = \epsilon_r \epsilon_0$, and ϵ_0 represents the permittivity of free space. P-type dopants are considered positive, while N-type dopants are considered negative, for simulation purposes.

To begin the monte carlo part of the simulation, either a single electron or hole is injected into the end of the P-type or N-type region to initiate either pure electron or pure hole injection (as selected by the user). After all generated carriers have left the device and the result of the injection trial calculated, the injection event is repeated. Typically this will be repeated at least 10000 times to build up statistically significant results.

Each injection trial begins by drifting the injected carrier for a random drift time. This random drift time, T_d, is calculated using,

$$T_d = \frac{-\log(R_{0-1})}{R_T}, \tag{4.2}$$

where R_{0-1} is a random number between 0 and 1, and R_T is the total interaction rate of the carrier at 6 eV. The carriers change in momentum and energy are then calculated, using the electric field strength from the position carrier drift was initiated. Finally the carrier position is updated to end the drift cycle. Next a second random number between 0 and 1 is generated to decide on the interaction event, using Table 4.1. If the carrier undergoes phonon absorption then it will gain the phonon energy, while if it undergoes phonon emission it will loose the phonon energy. Phonon emission is only possible if the carrier has more energy than it would loose by emitting a phonon. If the carrier undergoes impact ionisation, a new electron and hole are generated at the same position as the impact ionisation event. The energy each of the carriers have after the impact ionisation event, E_c^o, is calculated as

$$E_c^o = \frac{1}{3}(E_c^i - E_{th}), \tag{4.3}$$

where E_c^i was the energy of the impact ionising carrier. From Equation (4.3) it is possible to see that carriers can only undergo impact ionisation if their energy exceeds the threshold energy of impact ionisation (E_{th}).

This drift and scattering process is repeated for all carriers in the device, as a loop, until

Table 4.1: Table showing random number values required to trigger the interaction events from the SMC, where P_{abs}, P_{em}, and P_{ii} represent the probabilities of phonon absorption, phonon emission, and impact ionisation.

Interaction Event	Minimum Random Number	Maximum Random Number
Phonon Absorption	0	P_{abs}
Phonon Emission	P_{abs}	$P_{abs} + P_{em}$
Impact Ionisation	$P_{abs} + P_{em}$	$P_{abs} + P_{em} + P_{ii}$
Nothing	$P_{abs} + P_{em} + P_{ii}$	1

they have all drifted out of the device. At this point the next injection trial is started, which repeats until the requested number of trials has been completed.

The drift velocity and impact ionisation coefficient modes work similarly to this. However, they use a simplified process as they only require an infinite length uniform strength electric field for the tracking of a single electron/hole in time and position. As they are only required to track a singular carrier, the generation of addition carriers is disabled.

In the drift velocity mode, the simulator tracks the distance travelled and time taken for the carrier to undergo one million drift events to calculate the mean drift velocity at that field strength. For the impact ionisation coefficient mode, the simulator tracks the distance travelled between 20000 consecutive ionisation events, which is used to produce ionisation coefficients from the inverse of the mean distance between impact ionisation events.

4.3 Expansion of SMC model for SPAD simulation

The original silicon SMC model from [120] calculated the gain values for a given trial using gain = *number of impact ionisation events + 1*. In this work, the SMC model was modified to record the instantaneous current, $I(t)$, for each trial, building up a distribution of $I(t)$ over many trials. The contribution to instantaneous current was calculated after each carrier had drifted a minimum distance, Δx, using Ramo's theorem [148],

$$I = \frac{q\Delta x}{w\Delta t}, \tag{4.4}$$

where w is the depletion region width for the simulated bias condition, and Δt is the time taken for the carrier to travel the distance Δx. To verify the implementation of the instantaneous current tracking the SMC model was also modified to calculate the gain of each trial by integrating the $I(t)$ distribution of each trial, to determine the total charge, and dividing the result by q to determine the number of electrons. This value was then compared with the gain value calculated from the number of impact ionisation events,

good agreement between the values implied successful implementation.

Recording $I(t)$ enabled the model to determine if the device had undergone a breakdown event by defining an instantaneous current threshold, which when exceeded, would be counted as an avalanche breakdown event. In this work an instantaneous current threshold of 0.1 mA was used, which corresponds to a minimum detectable signal of 5 mV (into a 50 Ω termination) [65, 149].

4.4 Model Verification with Experimental Results

Further verification of the new SMC model involved benchmarking the simulated results of several key characteristics. These include drift velocities, drift diffusion coefficients, impact ionisation coefficients, avalanche gains, and excess noise factors.

To generate the SMC drift velocities, carriers were injected into a uniform strength electric field, with infinite length. The distance the carriers travelled, and time taken, for the carriers was recorded for 1 million individual drift events. Dividing the distance by the travel time results in the average drift velocity of the carrier for the given electric field strength. By repeating this simulation for different electric fields the average values of the electric field drift velocity for electrons and holes were obtained as a function of electric field strength. The simulated electric field dependent drift velocity compared well to reported data [150, 151, 152], as shown in Figure 4.2.

The SMC was used to simulate the drift diffusion coefficient, D, under weak electric field conditions. For each simulation condition 500000 electrons and 500000 holes were injected into an infinite length, uniform field strength electric field. Positions of these carriers was recorded with respect to time to yield $S(x,t)$, the accumulated probability distribution of the carrier position x at time t. It is the normalised probability that the carrier will have travelled less than a distance x from the point of injection (horizontal axis) by time t. Examples of $S(x,t)$ are shown in Figure 4.3.

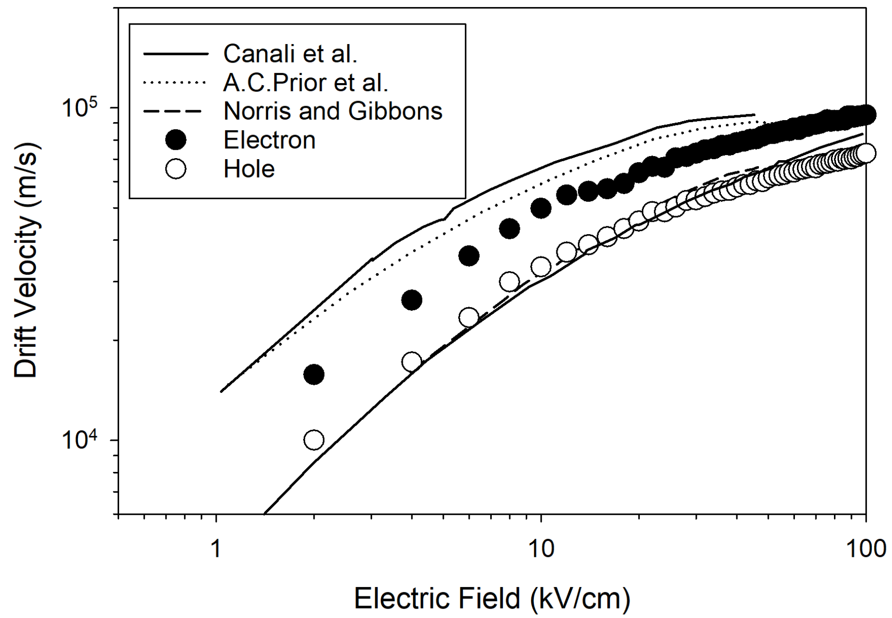

Figure 4.2: Comparison of reported drift velocities (lines) Canali et al. [150], A.C. Prior et al. [151], Norris and Gibbons [152] to drift velocities simulated by the SMC model (symbols).

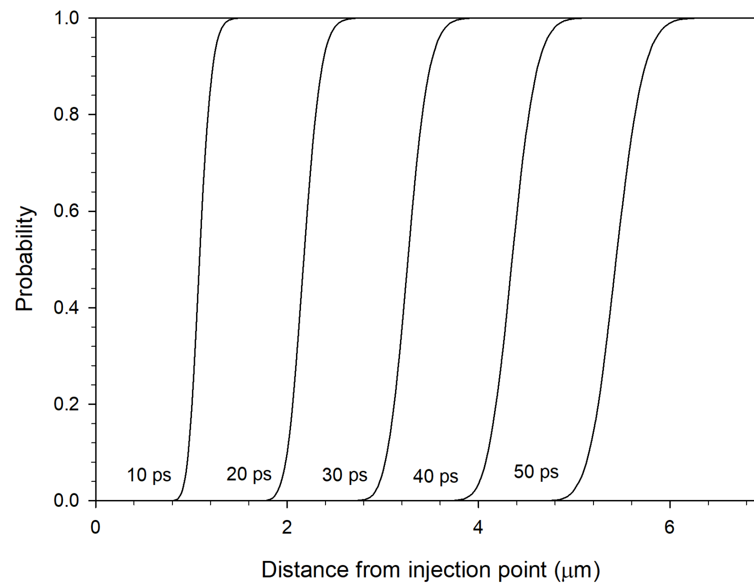

Figure 4.3: Example $S(x, t)$ for 500000 electrons in a 200 kV.cm^{-1} field at different times after the initial carrier injection.

$S(x,t)$ represents the error function with the horizontal axis scaled by $\sqrt{4Dt}$, and shifted by νt, where t is the drift time, and ν is the mean carrier velocity. The value of x corresponding to $S = 0.5$ is the mean distance travelled by the carriers in the electric field. To model the S-curve, the help of Emeritus Professor Graham Rees was enlisted. The derivation of $S(x,t)$ (Equation (4.5)) is shown in appendix C.

$$S(x,t) = \frac{1}{\sqrt{4\pi Dt}}\exp\left(-\frac{(x-vt)^2}{4Dt}\right), \tag{4.5}$$

The format of Equation (4.5) shows the linear region of the $S(x,t)$ has a gradient,m_p, of

$$m_p = \frac{1}{\sqrt{4\pi Dt}}, \tag{4.6}$$

which leads to,

$$D = \frac{1}{4\pi {m_p}^2 t}. \tag{4.7}$$

Using Equation (4.7) simulated on $S(x,t)$, diffusion coefficients were calculated at 10, 20, and 40 kV.cm^{-1}, as summarised in Table 4.2. The results are compared to reported results [153, 154], in Figure 4.4. The agreement shows that the SMC can simulate drift-diffusion under low electric field conditions, hence the model can be used for investigation of timing jitter.

As shown in Figure 4.4, the diffusion coefficients are electric field dependent. Normally, when considering diffusion you would not consider it to be electric field dependent as unlike drift, diffusion currents are possible without the presence of an electric field. Assuming diffusion is electric field independent is only true under weak field conditions, not a condition that APDs or SPADs are operated in.

The presence of an electric field gives the carriers energy, with a greater field strength giving more energy. Carriers with more energy are more likely to interact with their surrounding material. As the interaction rate increases, with increasing carrier energy, the variation in time between interaction events reduces. This reduction in time variation in turn reduces the random spread of carriers and hence the carriers diffusion coefficient. As the carrier energies required for the different interaction mechanisms vary between semiconductor materials the electric field dependency also varies between materials.

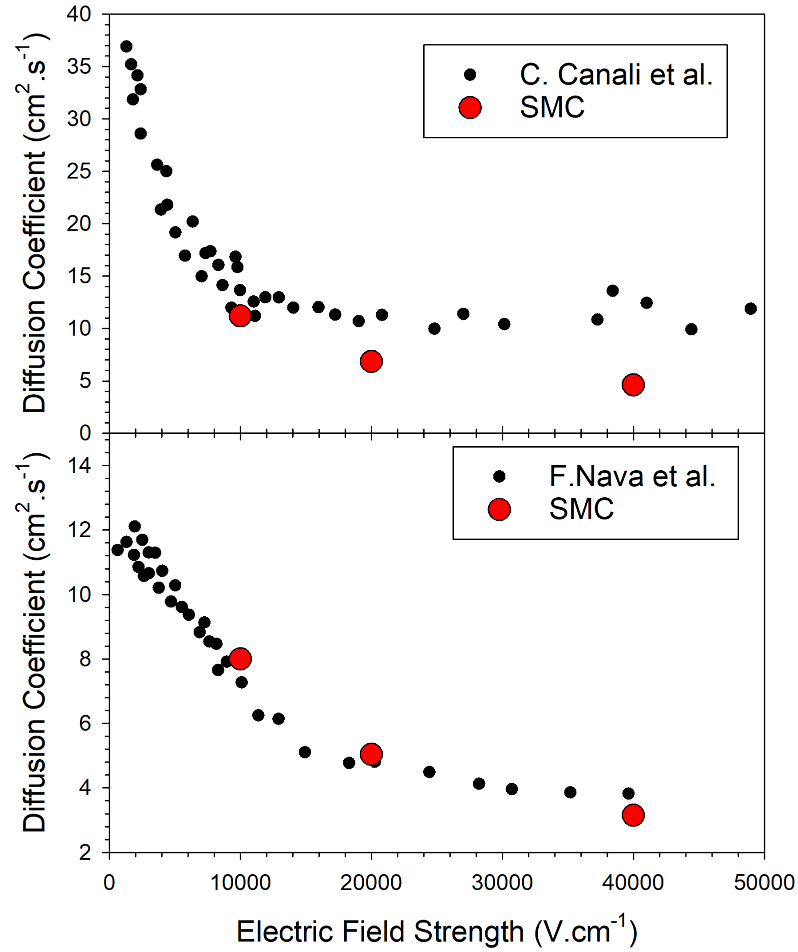

Figure 4.4: Comparison between simulated SMC diffusion coefficients (red dots) and reported diffusion coefficients (black dots) for electrons (top) from Canali et al. [153] and holes (bottom) from Nava et al. [154].

Table 4.2: Diffusion coefficient values generated from the SMC model.

Electric Field Strength kV.cm^{-1}	Electron Diffusion Coefficient cm^2.s^{-1}	Hole Diffusion Coefficient cm^2.s^{-1}
10	11.19	8.00
20	6.85	5.04
40	4.61	3.15

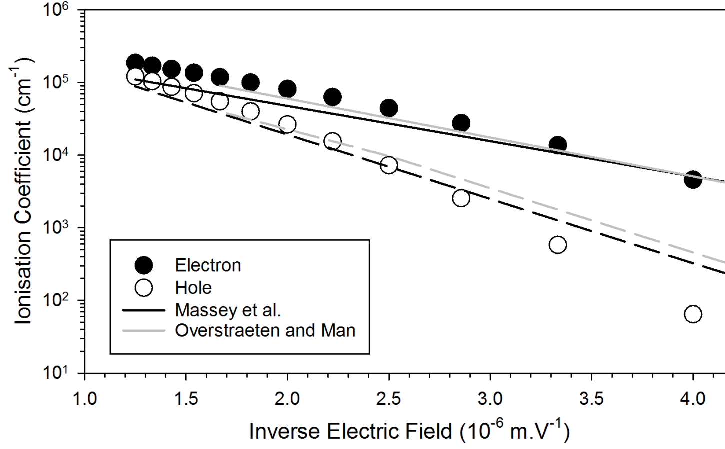

Figure 4.5: Comparison of reported electron and hole ionisation coefficients (solid and dashed lines) from Massey et al. [155, 156] and Overstraeten and Man [157] to SMC simulations (symbols).

To simulate the impact ionisation coefficients, carriers were again injected into an infinitely long structure with a uniform electric field. Then the distance travelled to complete 20000 impact ionisation events was recorded. The simulated impact ionisation coefficients were calculated as the inverse of the mean distance between consecutive impact ionisation events. The simulated impact ionisation coefficients are in good agreement with reported results [155, 156, 157], as shown in Figure 4.5.

Table 4.3: Si Devices A-E [158] used to validate $M(V)$ and $F(M)$ in Si SMC model.

Device (Layer)	Structure	Intrinsic Region (µm)	N_p ($\times 10^{18}$ cm^{-3})	N_i ($\times 10^{16}$ cm^{-3})	N_n ($\times 10^{18}$ cm^{-3})	V_b (V)
A (6B115)	P-I-N	0.082	2.5	2	4	7.0
B (6B118)	P-I-N	0.13	3	2	3	8.8
C (6B119)	P-I-N	0.26	3	2	3	13.3
D (8A3F31)	N-I-P	0.82	6	1.8	0.5	29.5
E (8A5F31)	P-N	-	7.5	-	0.075	17.0

The validation of the new simulator included avalanche multiplication characteristics of five devices, consisting of three P-I-Ns, a N-I-P and a P-N diode [158]. $M(V)$ and $F(M)$ of these devices, whose structures are described in Table 4.3, were simulated to compare to experimental data. This comparison can be seen in Figure 4.6.

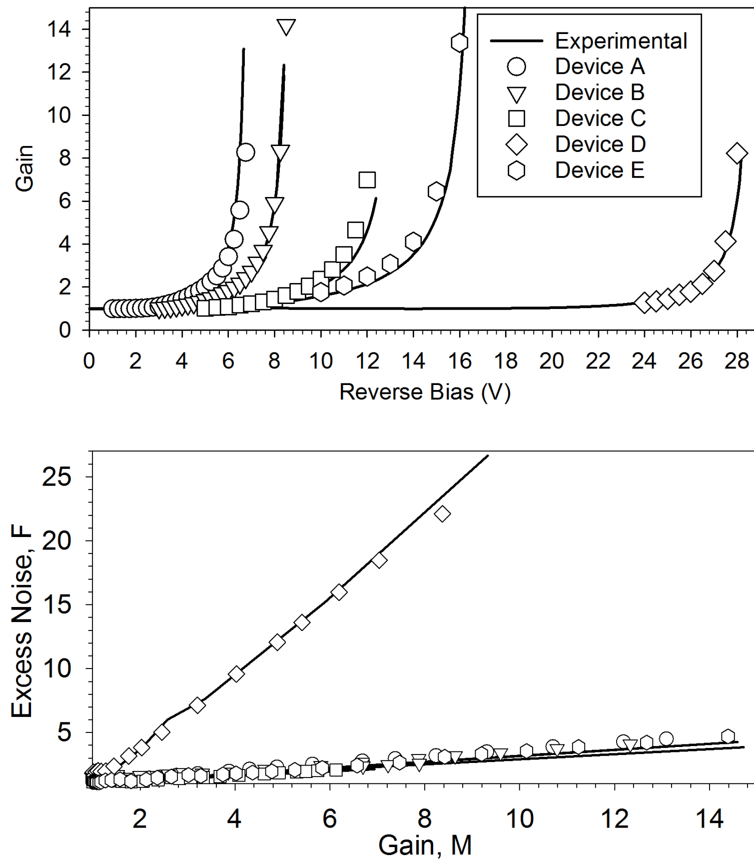

Figure 4.6: Comparison of experimental (lines) [158] and simulated results (symbols) of $M(V)$ (top) and $F(M)$ (bottom) for Si diodes.

Breakdown voltages, V_b, of devices A-E were extracted from their experimental gain data [158], by plotting the inverse gain against reverse bias, as in Figure 4.7. By extrapolating the linear part of the inverse gain curve to the x-axis a reasonable approximation for breakdown voltage can be achieved. The extracted V_b values are shown in Table 4.3. Figure 4.8, which plots the SMC simulated breakdown probability versus reverse bias, shows that the simulated results remain zero until the devices reach approximately their extracted V_b. As reverse bias increases beyond V_b, P_b steadily rises and eventually approaches unity.

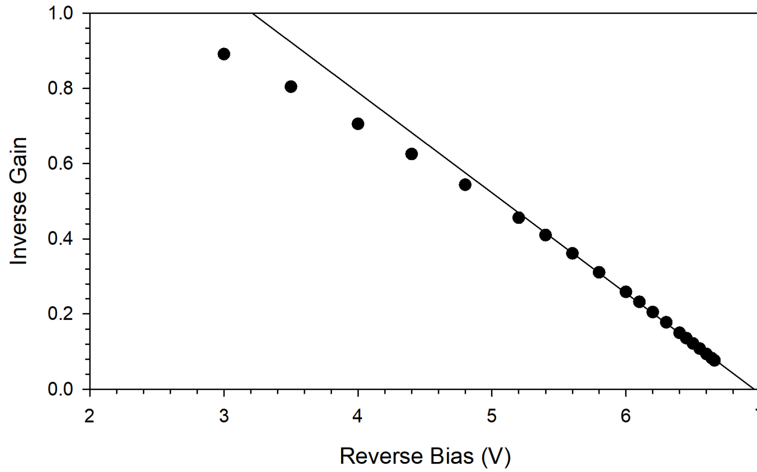

Figure 4.7: Example of extraction of the experimental breakdown voltage by plotting the inverse gain versus reverse bias. The experimental data used for device A gave a breakdown voltage of 7 V.

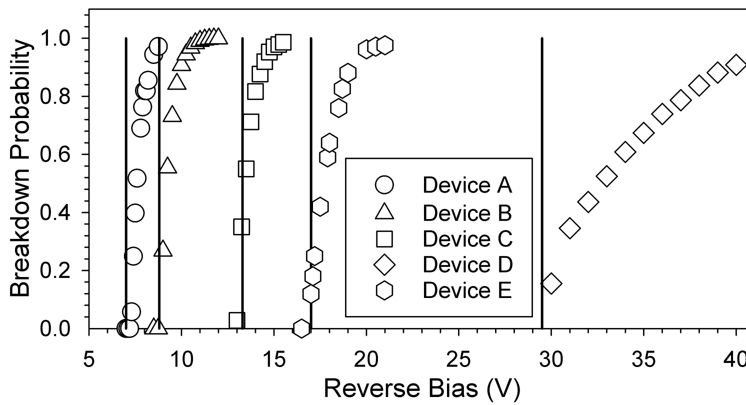

Figure 4.8: SMC simulated breakdown probabilities versus reverse bias characteristics for Si devices A-E (symbols). The results are consistent with the extracted V_b values in Table 4.3.

4.5 Effective Ionisation Coefficients

Before simulating SPAD characteristics for actual devices using the SMC model the SPAD characteristics should be benchmarked against previously reported models. This benchmarking is to check that the breakdown probabilities and timing characteristics predicted by the model reflect the characteristics predicted by existing models and experimental results. To compare the SMC model generated breakdown probabilities and timing char-

acteristics against other models, only the models, and not the input parameters should be compared. The RPL model [143] and recurrence equation model [144] used for comparison require, different inputs cf. SMC models, namely the effective ionisation coefficients, $\alpha^*(E)$ and $\beta^*(E)$. If previously published effective ionisation coefficients are used then there is the potential for the discrepancies between the results from the SMC model and the RPL or recurrence equation models to be as a result of the impact ionisation coefficients rather than the modelling methods. Therefore, effective impact ionisation coefficients require generation from the SMC model.

Impact ionisation path lengths for 20000 consecutive impact ionisation events were calculated for each simulated electric field strength, then these impact ionisation path lengths were used to create probability densities. Fitting these probability densities, using a hard deadspace approximation, specific values, for each electric field, of α^*, β^* , electron deadspace (d_e), and hole deadspace (d_h) can be generated. The best fits were obtained using effective impact ionisation coefficients of

$$\alpha^* = 5.0 \times 10^7 \exp\left(-\left(\frac{8.0 \times 10^7}{E}\right)^{1.28}\right), \tag{4.8}$$

and

$$\beta^* = 4.4 \times 10^7 \exp\left(-\left(\frac{9.4 \times 10^7}{E}\right)^{1.65}\right). \tag{4.9}$$

To obtain the required impact ionisation threshold energies, deadspace is plotted against the inverse electric field, then the extracted gradients are the values of threshold energy of 2.18 and 3.41 eV for electrons and holes. Threshold energies allow for the deadspace to be calculated at any required electric field strength rather than them having to be manually inputted or stored in a lookup table in advance.

4.6 Simulated SPAD Characteristics

The validated Si SPAD SMC model was used to simulate breakdown probability and timing characteristics of an ideal 1 µm device which was subsequently compare against results produced by previously reported models [144, 143] using the SMC generated effective ionisation coefficients. The Si SPAD SMC model was then used to simulate the breakdown probability and timing characteristics of an Si SAM-SPAD structure to explore the effect of doping orientation by comparing SPAD characteristics of a p-topped and an n-topped SAM-SPAD structure.

4.6.1 Breakdown Probability

Simulated breakdown probability versus reverse bias of an ideal silicon P-I–N, with a 1 µm intrinsic region, are compared to results from recurrence equations [144] where the previously extracted $\alpha^*(E)$ and $\beta^*(E)$ (Equations (4.8) and (4.9)) were used. A comparison of this can be seen in Figure 4.9. The maximum variation between the two models is 21% for an injected electron-hole pair at the junction of the p/i regions (pure electron injection) or the i/n regions (pure hole injection). At any reverse bias above the breakdown voltage the breakdown probability for pure electron injection is higher than for pure hole injection, as expected due to $\alpha^* > \beta^*$ in Silicon.

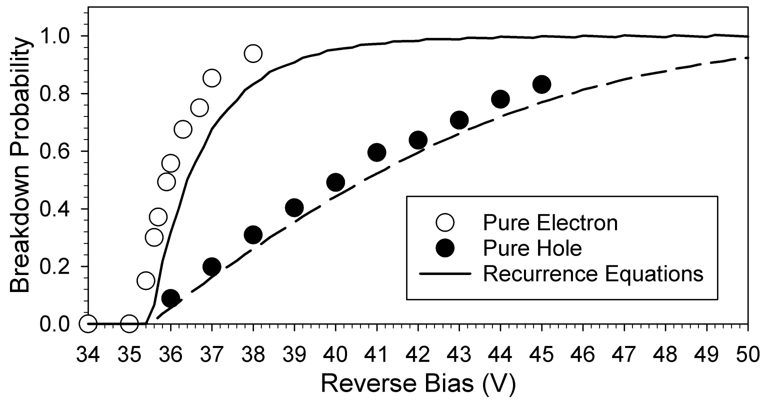

Figure 4.9: Comparison of breakdown probability versus reverse bias characteristics for an ideal 1.0 µm p-i-n diode[3] from the SMC model (symbols) and recurrence equations (lines) [144] using the extracted $\alpha^*(E)$ and $\beta^*(E)$ (Equations (4.8) and (4.9).)

4.6.2 Timing characteristics

Timing characteristics, mean time to breakdown and timing jitter, generated by the SMC model were compared to an existing RPL model [143]. Again, an ideal p-i-n with a 1.0 µm intrinsic region as well as Equations (4.8) and (4.9) were used. For this comparison timing jitter was calculated as the standard deviation of time to breakdown rather than the FWHM to enable direct comparison with the values from [143]. The data from [143] assumes that the carriers are travelling at a constant drift velocity of 1×10^5 m.s^{-1}. As can be seen in Figure 4.10 although results from both models are similar the RPL overestimates the mean time to breakdown. This overestimation could be due to the constant drift velocity approximation used in RPL models, which underestimates the velocity for the impact ionising carriers [159].

[3] An ideal p-i-n diode has a uniform electric field strength across the intrinsic (i) region and no depletion into either the p or n regions.

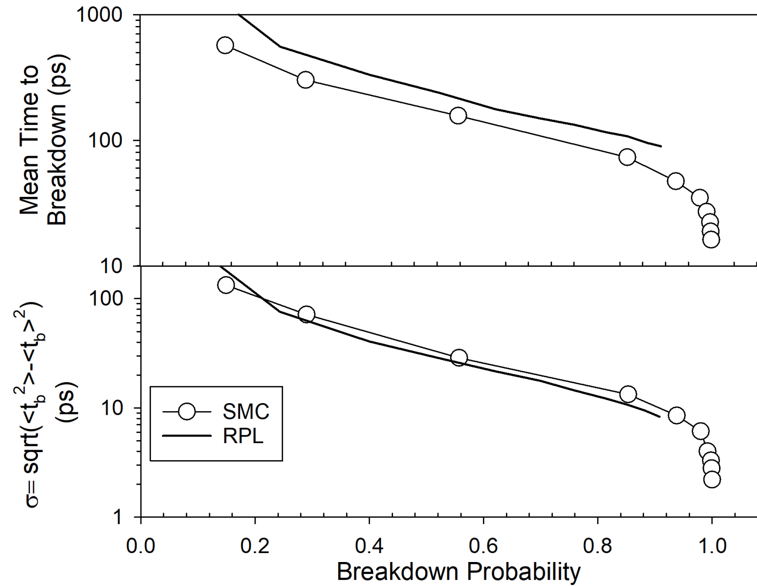

Figure 4.10: Comparison between RPL simulated (lines) [143] and SMC simulated (symbols) mean time to breakdown (top) and standard deviation of time to breakdown (bottom) for a 1 µm ideal p-i-n diode.

4.6.3 P-on-N vs N-on-P

Finally, the new SMC model was used to predict avalanche breakdown characteristics from different SPAD designs. Two Si SPAD structures based on an n^+-p-p^--p^+ design, which are either n^+-topped or inverted to become p^+-topped were investigated. The doping profile of the structures are shown in Table 4.4. The n^+-topped (n-on-p) is the same orientation as commercial SPADs, whilst the p^+-topped (p-on-n) is a potential alternative to the commercial standard. The electric field profiles of both designs are shown in Figure 4.11 at their breakdown voltage of 27 V.

Table 4.4: Structure of P-on-N and N-on-P devices.

(a) N-on-P

Doping Concentration (cm^{-3})	Thickness (µm)
-0.5×10^{18}	0.5
1×10^{16}	0.1
5×10^{16}	0.5
5×10^{15}	0.5
6×10^{18}	0.5

(b) P-on-N

Doping Concentration (cm^{-3})	Thickness (µm)
6×10^{18}	0.5
5×10^{15}	0.5
5×10^{16}	0.5
1×10^{16}	0.1
-0.5×10^{18}	0.5

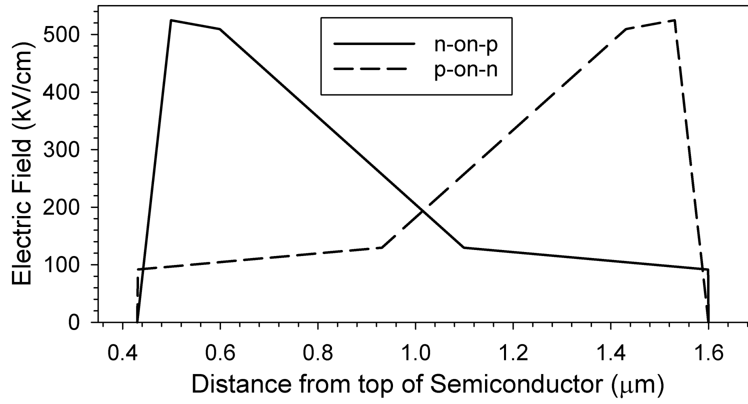

Figure 4.11: Electric field profiles of n-on-p and p-on-n designs at their breakdown voltage of 27 V.

In the simulations, all incident photons are assumed to be absorbed in the uppermost layer, corresponding to pure hole injection in the n-on-p design, and pure electron injection in the p-on-n design. Simulated breakdown probability, P_b, are in Figure 4.12. It is apparent, that P_b rises more rapidly with reverse bias for the p-on-n design cf. the n-on-p design. This is expected because the p-on-n design is reliant on pure electron impact ionisation and $\alpha > \beta$ in Si. The lower P_b for the n-on-p design is reflected in the typical SPAD operating conditions, where the reverse bias is normally kept $> 1.2 \times \mathrm{V_b}$ [160].

Although the p-on-n design offers more desirable breakdown probability characteristics its timing characteristics are worse than that of the n-on-p design, as shown in Figure 4.13.

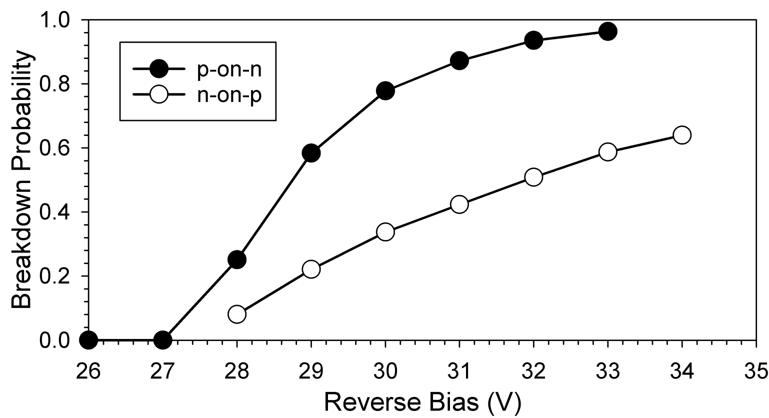

Figure 4.12: Comparison of simulated breakdown probability between the n-on-p and p-on-n designs

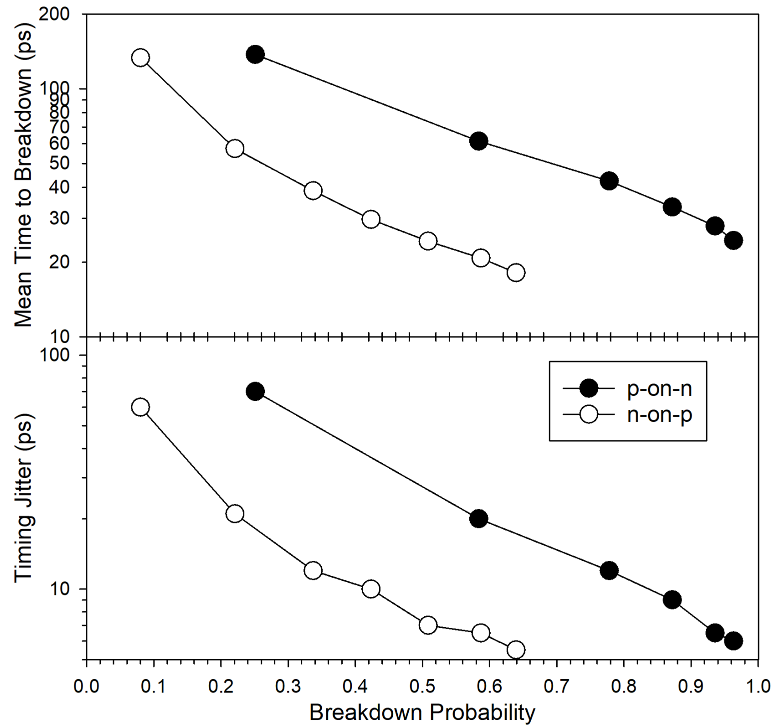

Figure 4.13: Comparison between mean time to breakdown (top) and timing jitter, FWHM, (bottom) for the n-on-p and p-on-n designs.

The impact of the alternating doping orientations can be explored by comparing characteristics at a particular P_b. To achieve $P_b = 0.5$, the n-on-p design requires a reverse bias of 32.0 V, corresponding to 1.19 V_b and peak electric field of 563 kV.cm^{-1}. The corresponding values for the p-on-n design are lower at 28.7 V, 1.06 V_b and 535 kV.cm^{-1}. Due to the lesser electric field in the p-on-n design, carriers need to travel greater distances to build up the required energy to undergo impact ionisation. The carriers should also be more susceptible to scattering events other than impact ionisation which would further delay impact ionisation, increasing the impact ionisation path length, as they have a lesser energy.

The carrier scattering susceptibility was tested by recording the number of scattering events for both designs at P_b of 0.5 and 0.6. At both breakdown probabilities the p-on-n design underwent 2.7× phonon emission, 3.0× phonon absorption and 1.6× impact ionisation events cf. the n-on-p design.

4.7 Chapter summary

In this chapter, an updated SMC for the simulation of Silicon SPADs has been presented. It has been validated using experimental $M(V)$ and $F(M)$ as well as reported silicon drift velocities, diffusion coefficients, and impact ionisation coefficients. The SMC model can also simulate breakdown probability, time to breakdown, and timing jitter.

The model was then used to investigate Si SPAD designs, focusing on the comparison between a typical n-on-p design with a p-on-n design. For a given breakdown probability the n-on-p design offers better timing performance (i.e. smaller timing jitter), while the p-on-n design requires a lesser applied bias.

Chapter 5

InP SMC

5.1 Introduction

InP has been one of the commonly used avalanche materials for APDs for decades. As for Si, the avalanche characteristics of InP can be simulated using an SMC model, if a validated parameter set for InP can be developed. In addition, it is desirable to extend the temperature range of the SMC below room temperature, down to 200 K, at which $In_{0.53}Ga_{0.47}As$/InP APDs are often operated [29, 50, 161].

In this chapter a SMC parameter set for InP capable of simulating devices between 150-300 K is presented. This parameter set has also been used to generate equations for the effective impact ionisation coefficients, $\alpha^*(E)$ and $\beta^*(E)$, which can be used in quicker models such as a Recurrence equation model [162] or a Random path length model [163].

5.1.1 Motivation

Several papers have previously reported on sub-room temperature impact ionisation in InP [164, 165, 166], which were obtained from diffused InP junctions, leading to significant uncertainties surrounding the electric field profiles of the devices used. Further to this, the impact ionisation coefficients were not actually presented in [164] and, as their doping profile was not presented, it is not possible to extract the ionisation coefficients from their results. Comparing the room temperature ionisation coefficients from [165, 167, 168, 169], values from [165] were significantly higher than the rest. Considering that [165] used only a single InP structure to determine their ionisation coefficients and have treated mixed injection measurements as pure hole injection, reported values from [165] may not be sufficiently accurate.

5.2 Temperature Dependent InP SMC

The SMC simulator used for the Silicon work in chapter 4 was extended in this work to incorporate temperature dependence in the same form as [119]. In this temperature dependent version of the SMC the intervalley phonon absorption, and intervalley phonon emission rate equations, R_{ab} and R_{em}, take the form

$$R_{ab}(T) = \frac{N(T)}{\lambda(T)(2N(T)+1)}\sqrt{\frac{2(E_c+\hbar\omega)}{m^*}}, \tag{5.1}$$

$$R_{em}(T) = \frac{N(T)+1}{\lambda(T)(2N(T)+1)}\sqrt{\frac{2(E_c-\hbar\omega)}{m^*}}, \tag{5.2}$$

where $N(T)$ is the temperature dependent phonon occupation factor,

$$N(T) = \left(\exp\left(\frac{\hbar\omega}{k_bT}\right)-1\right)^{-1}, \tag{5.3}$$

and $\lambda(T)$ is the temperature dependent mean free path between scattering events. As $\lambda(T) \sim (1+2N(T))^{-1}$ [170], $\lambda(T)$ can be expressed as

$$\lambda(T) = \lambda(300K) \times \frac{2N(300K)+1}{2N(T)+1}. \tag{5.4}$$

The impact ionisation rate equation remains unchanged, and uses the Keldysh equation [121] (see Equation (2.28), repeated below). Therefore the temperature dependent nature of the SMC simulator is implemented through $N(T)$ and $\lambda(T)$.

$$R_{ii} = C_{ii}\left(\frac{E_c-E_{th}}{E_{th}}\right)^{\gamma} \tag{2.28 repeated}$$

For the InP parameter set, a phonon energy of 42 meV was chosen, based upon reported values of the longitudinal optical phonon frequency of 345 cm^{-1} [171]. In the SMC, one of the ways of obtaining the threshold energy, E_{th}, is to calculate it as a weighted average of the 3 band gaps (where E_{τ} is the gamma band gap energy, E_x is the X band gap energy, and E_L is the L band gap energy), using Equation (5.5) originally used for Diamond in 1959 [172],

$$E_{th} = \frac{1}{8}\left(E_{\tau}+3E_x+4E_L\right). \tag{5.5}$$

A more accurate E_{th} could be achieved from the two parameter equation, presented by Chadi and Cohen [173], unfortunately the values of these two points can not be experimentally measured. The 3-band approximation method gave an E_{th} value of 1.95 eV but this

Table 5.1: InP SMC Parameters

	Electrons	Holes
Prefactor of impact ionisation rate, $C_{ii}(\mathrm{s}^{-1})$	3.5×10^{12}	8.5×10^{12}
Threshold Energy, E_{th} (eV)	1.55	1.55
Softness Factor, γ	0.7	0.7
Phonon Energy, $\hbar\omega$ (meV)	42	
Mean Free Path at 300 K, $\lambda(300K)$(Å)	41	42
Effective mass, m^* (kg)	$0.62\ m_0$	$0.63\ m_0$
Relative Permittivity	12.5	

value produced simulated drift velocities that were several times higher than the reported values and did not yield experimentally correct gain and excess noise factors. So, instead, the combined scattering rates were fitted against reported rates from [147], resulting in a E_{th} value of 1.55 eV. This forms part of the parameter set presented in this chapter. The key parameters for InP are summarised in Table 5.1.

The electron and hole effective masses were adjusted so that the SMC simulator produced total scattering rates and saturation velocities in line with reported values, [147, 174, 175]. A comparison of saturation velocities (6.8×10^6 cm.s^{-1} for electrons [174] and 7.0×10^6 cm.s^{-1} for holes [175]) is shown in Figure 5.1. Effective masses of $0.62m_0$ and $0.63m_0$ were used for electrons and holes, respectively.

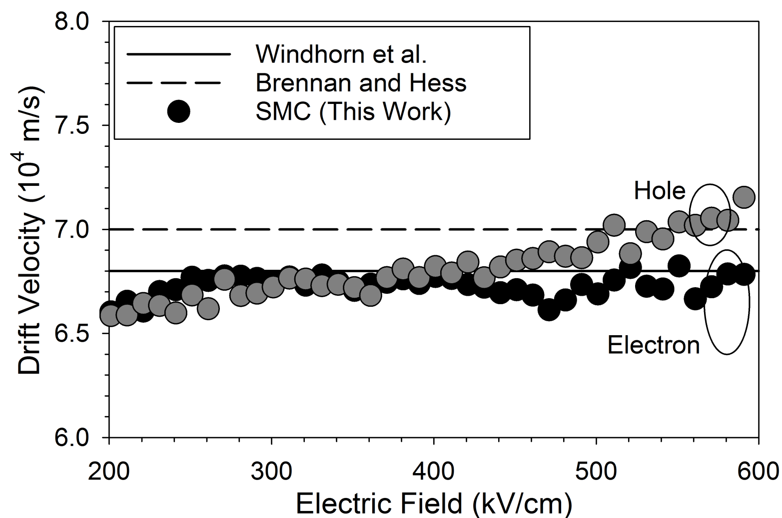

Figure 5.1: Comparison between published room temperature saturation velocities and room temperature drift velocities generated by the SMC. Windhorn et al.[174], Brennan and Hess [175].

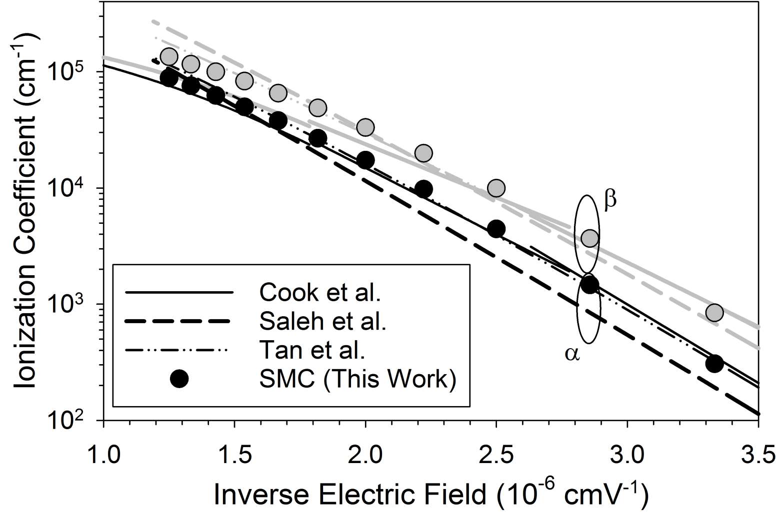

Figure 5.2: Comparison between published room temperature ionisation coefficients and room temperature ionisation coefficients generated by the SMC. Cook et al. [167], Saleh et al. [168], and Tan et al. [169]

The next step in producing the InP SMC parameter set was to adjust the values of C_{ii}, $\lambda(300K)$, and γ so that the impact ionisation coefficients (the average of the inverse path length between consecutive ionisation events) agreed with the published coefficients [167, 168, 169] as shown in Figure 5.2. Whilst making this adjustment, care was taken to ensure the drift velocity and combined scattering rates remained in agreement with the reference values.

Though characterisation data for the devices used in this work have previously been published in [169, 98], capacitance-voltage fitting has been repeated to ensure data accuracy. In this work, the entire electric-field profile is considered including the rapidly increasing field close to the intrinsic region and the graded nature of the intrinsic region. This is distinct from the approximation of uniform field with abrupt boundaries commonly used in simpler models.

C-V data for devices A-G, shown in Table 5.2, was obtained from [176]. Using the Poisson equation (appendix B) as well as doping densities and layer thicknesses, C-V characteristics were obtained. An example of the C-V fit for device D is shown in Figure 5.3 which shows good agreement between the measured C(V) and simulated C(V) of a 200 µm radius circular diode.

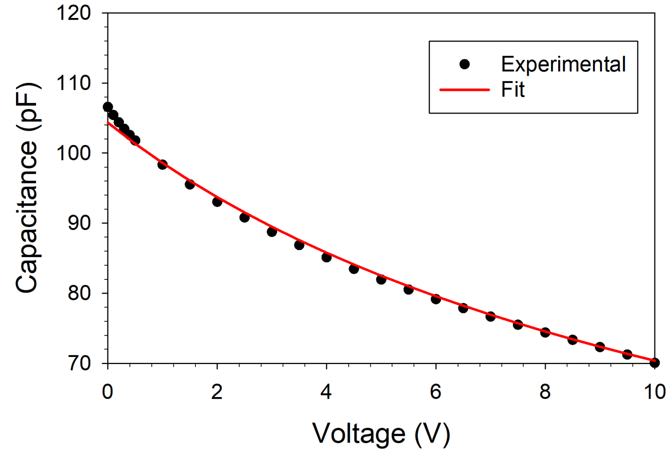

Figure 5.3: Example C-V fit for device D with a radius of 200 µm [176].

Table 5.2: InP device structure details for devices used as part of the model validation. InP Devices A-F used for room temperature validation. InP Devices C-G used in temperature dependent validation

Device (Layer)	Intrinsic region width (µm)	Structure	P/N doping ($10^{18}cm^{-3}$)	I doping ($10^{15}cm^{-3}$)
A (MR987) [169]	2.5	P-I-N	0.6	0.3
B (MR1776) [169]	1.25	P-I-N	0.7	2
C (VN80) [169]	0.545	P-I-N	1.1	1
D (VN76) [169]	0.125	P-I-N	1.3	10
E (MR1104) [169]	0.8	N-I-P	0.8	0.8
F (MR1105) [169]	0.23	N-I-P	0.9	4
G (MR2267) [98]	1.63	P-I-N	1.8	1

Finally, to produce the InP parameter set, fine adjustment was performed to the parameter set so the SMC simulator could produce accurate $M(V)$ and $F(M)$ results for the device A-F. The simulated and experimental results are compared in Figure 5.4.

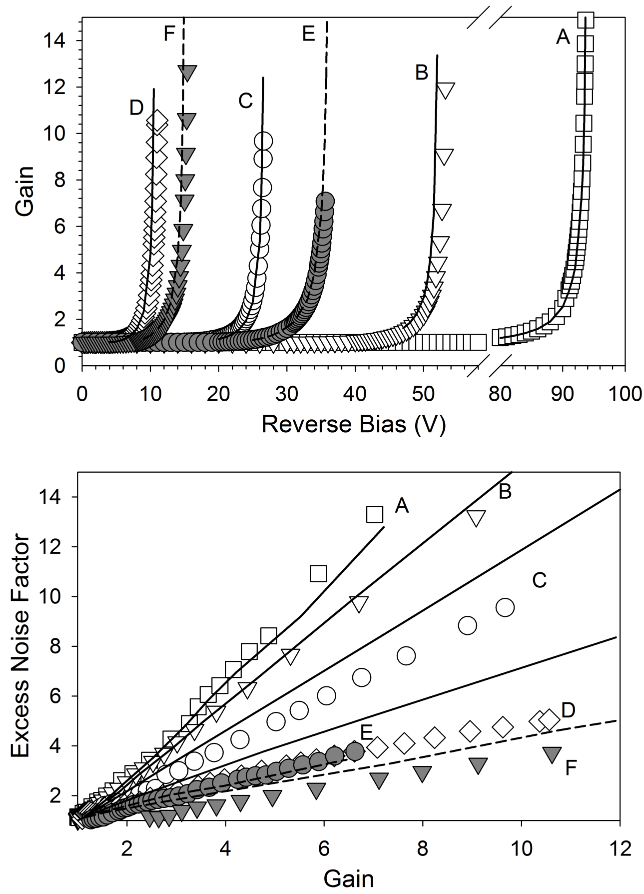

Figure 5.4: Comparison of (top) $M(V)$ and (bottom) $F(M)$ simulated by the InP SMC (lines) with data from [169] (symbols) for 4 P-I-N devices (A-D in Table 5.2) and 2 N-I-P devices (E-F in Table 5.2) at 300K.

5.3 Temperature Dependent Multiplication

Once the parameter set had been established, it was possible to simulate the temperature dependent gain results for devices C-G (Table 5.2) from [98], as shown in Figure 5.5, using the scaling nature of Equations (5.3) and (5.4). Good agreement is observed down to 150 K, with simulations at 100 K and 77 K underestimating $M(V)$. The underestimation of $M(V)$ could be due to reduced active doping densities at low temperatures or that our approximation of $\lambda(T)$ and $N(T)$ as the only temperature dependent parameters is no longer valid below 150 K. Alternatively, it could be that the limit of the capabilities of the single parabolic band approximation has been reached, secondary band interactions could become more important at these low temperatures.

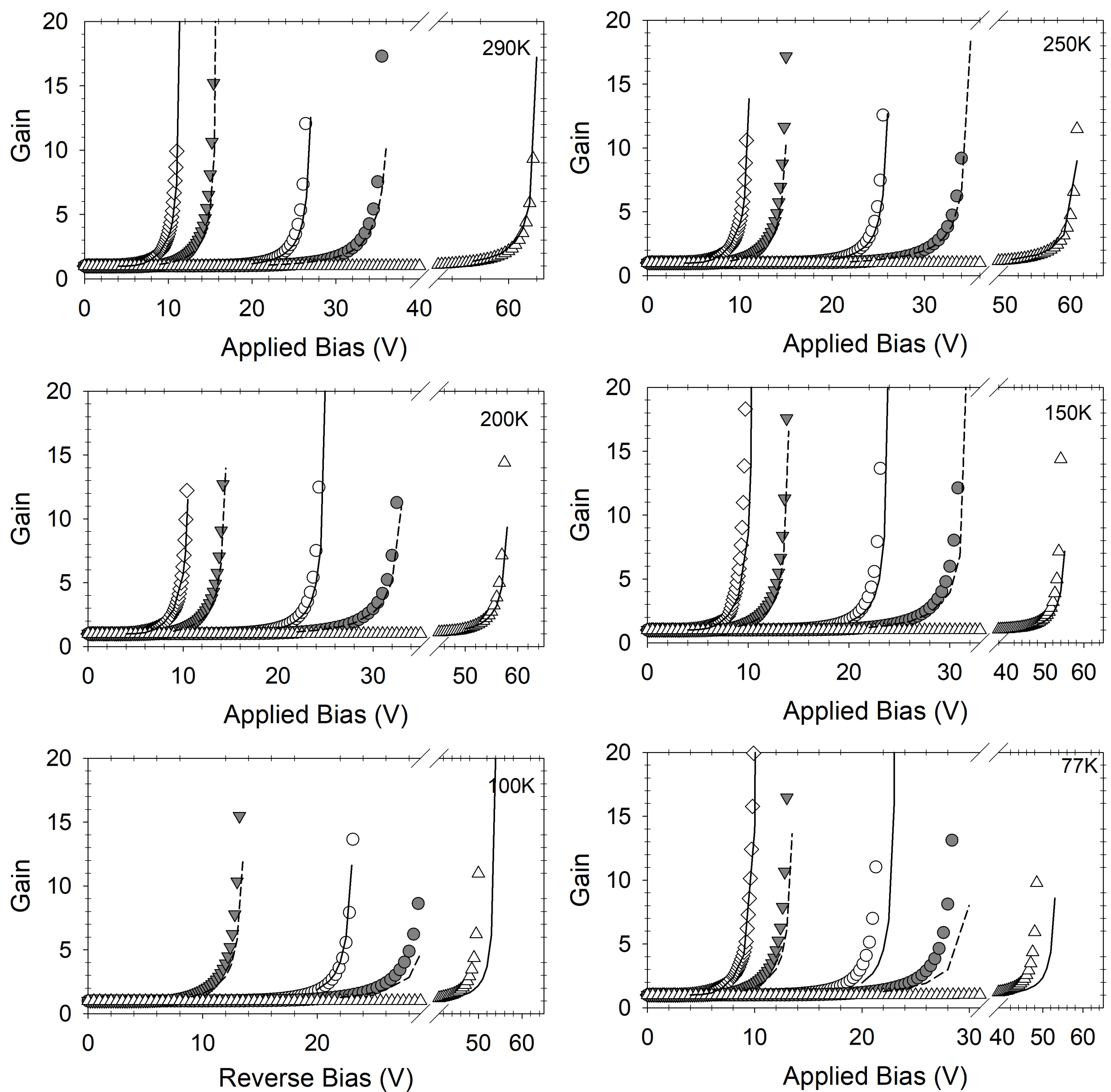

Figure 5.5: Temperature dependent gain of InP devices C-G, however device D was not measured at 100 K. Experimental results (symbols) from [98]

5.4 Temperature Dependent Impact Ionisation Coefficients

Though impact ionisation coefficients can be generated from the SMC as previously discussed, these impact ionisation coefficients are not useful for simpler models such as the RPL model and recurrence equation models (section 2.3). Both of these simpler models require ionisation coefficients in the form of the effective ionisation coefficients $\alpha^*(E)$ and $\beta^*(E)$ which can be extracted from probability density functions, PDFs, of impact ionisation path lengths. The impact ionisation path lengths can be generated from the SMC, then can be processed to form probability densities. The probability densities can be fitted to $h_e(x)$ for electrons and $h_h(x)$ for holes, where $h_e(x)$ and $h_h(x)$ are shown in Equations (2.18) and (2.19) repeated below.

$$h_e(\xi) = \begin{cases} 0, & \xi \leq d_e \\ \alpha^* exp(-\alpha^*(\xi - d_e)), & \xi > d_e \end{cases} \qquad \text{(2.18 repeated)}$$

and

$$h_h(\xi) = \begin{cases} 0, & \xi \leq d_h \\ \beta^* exp(-\beta^*(\xi - d_h)), & \xi > d_h \end{cases} \qquad \text{(2.19 repeated)}$$

Examples of fitted probability density functions are shown in Figure 5.6 and the full set of PDFs used to generate $\alpha^*(E)$ and $\beta^*(E)$ parameters for 150, 200, 250, 290 K in Table 5.3 can be found in appendix D. Figure 5.6 shows that the electron PDF extends over a greater distance than the hole PDF, for a given combination of temperature and electric field conditions. This is consistent with $\alpha < \beta$ in InP. The α^* and β^* parameters are expressed as,

$$\alpha^*(E) \text{ or } \beta^*(E) = A_f exp\left[-\left(\frac{B_f}{E}\right)^{C_f}\right], \qquad (5.6)$$

where A_f, B_f, and C_f are the fitting parameters presented in Table 5.3. For this work the published impact ionisation threshold energies of 2.8 eV for electrons and 3.0 eV for holes [169] were used. Using the relationship between deadspace, threshold energy and electric field, E,

$$d_e \text{ or } d_h = \frac{E_{th}}{qE} \qquad (5.7)$$

where q is the charge of an electron.

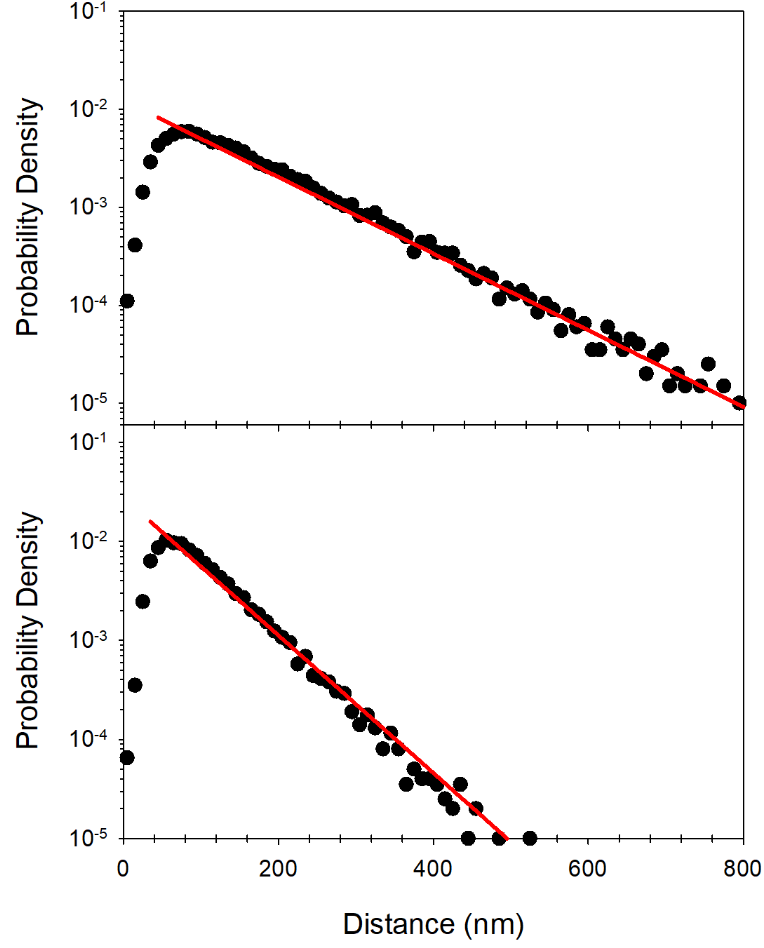

Figure 5.6: Examples of probability densities (symbols) fitted using the hard deadspace approximation (lines) at 290 K and 700 kV.cm-1 for electrons (top) and holes (bottom).

Table 5.3: α^* and β^* for InP at 150, 200, 250, and 290 K extracted from PDFs generated from the InP SMC

Temperature (K)		A_f ($10^8 \times \text{m}^{-1}$)	B_f ($10^8 \times \text{V.m}^{-1}$)	C_f
290	α^*	2.55	2.15	1.08
	β^*	0.98	1.00	1.60
250	α^*	4.70	2.96	0.94
	β^*	1.43	1.14	1.48
200	α^*	1.70	1.60	1.21
	β^*	6.30	2.74	0.90
150	α^*	4.03×10^2	7.73×10^1	0.44
	β^*	1.68	1.15	1.39

To test the accuracy of the effective ionisation coefficients of Table 5.3, they were used with a recurrence model (implemented using [102]), for devices C-G at 150, 200, 250, and 290 K. The simulated $M(V)$ characteristics were in agreement with the experimental results [98], as shown in Figure 5.7. Room temperature $F(M)$ characteristics were also simulated for devices A-F and compared to experimental results from [169], as shown in Figure 5.8. There is good agreement except for device A with a 2.5 µm intrinsic region, which is rarely used for APD/SPAD designs. This discrepancy for device A is likely to have arisen from inaccuracy of the extracted α^* and β^* expressions at very low electric fields.

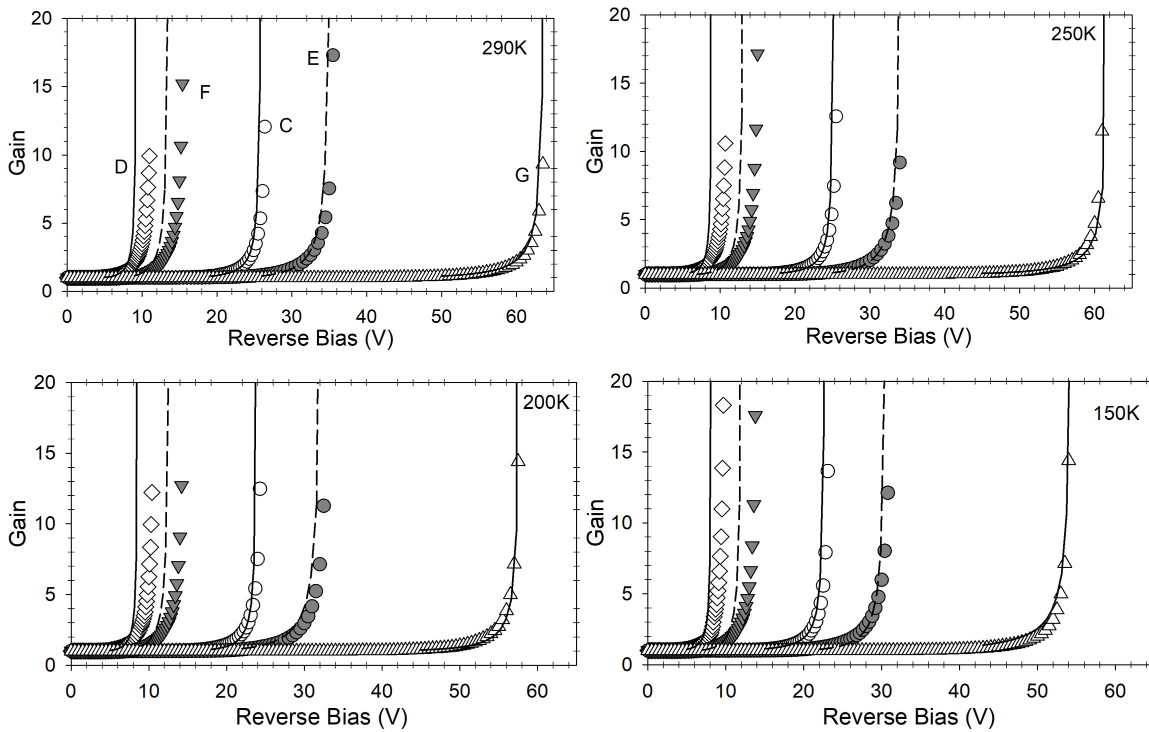

Figure 5.7: Simulated temperature gain, using the SMC generated effective ionisation coefficients in a recurrence model [102], compared to experimental gain results [98].

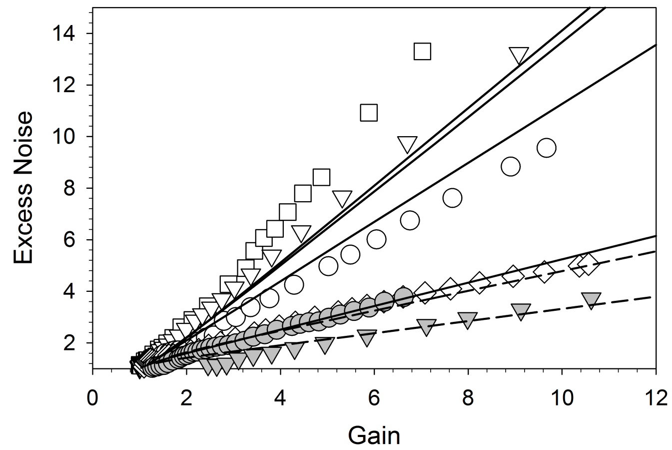

Figure 5.8: Simulated room temperature excess noise factors, using the SMC generated effective ionisation coefficients in a recurrence model [102], compared to experimental results [169].

5.5 Chapter summary

A new SMC parameter set for InP has been validated against reported saturation velocities, impact ionisation coefficients, room temperature gain and excess noise data, as well as temperature dependent gain data. Using the InP parameter set, expressions for effective ionisation coefficients have been extracted for use with simpler models (e.g. RPL or recurrence equations) for electric fields at 400-800 kV.cm^{-1} at 150, 200, 250, and 290 K.

Chapter 6

Mesa SPADs

This chapter presents results from Si Mesa diodes designed, and fabricated for SPAD operation. These mesa Si SPADs show higher DCR than commercial planar Si SPAD structures, however they can still compete in terms of SPDE. Mesa fabrication is worse cf. planar devices due to high DCR caused by sidewall damage. However, mesa fabrication processes allow for a more versatile use of material as diode shapes are not limited by the first ion implantation performed on the wafers, instead they are decided during the etching phase much later in the fabrication process.

6.1 Sample Fabrication

A brief outline of the fabrication process is presented here, whilst a fuller description is outlined in appendix E. The wafers used in this chapter were epitaxially grown by IQE Silicon, Cardiff, UK, onto a <100> Si substrate. The epitaxial layers include a 1 µm thick P+ layer with a resistivity of 0.02 Ω.cm (3.3×10^{18} atoms.cm^{-3}) and a 3.25 µm thick P- layer with a resistivity of 15 Ω.cm (8.8×10^{15} atoms.cm^{-3}) as shown in Figure 6.1 (left).

Device fabrication began with boron and phosphorus implantation. Post-implantation thermal annealing and oxidation steps were used to diffuse and activate the dopants as well as remove the implantation damage. Ion implantation was performed by Cutting Edge Ions, Anaheim, USA, and the high temperature annealing was performed by Ion Beam Services, Peynier, France. After the annealing the sample should have a structure closer resembling the schematic shown in Figure 6.1 (right) rather than Figure 6.1 (left). Before the Si substrates were purchased the fabrication process was modelled by Simon Dimler

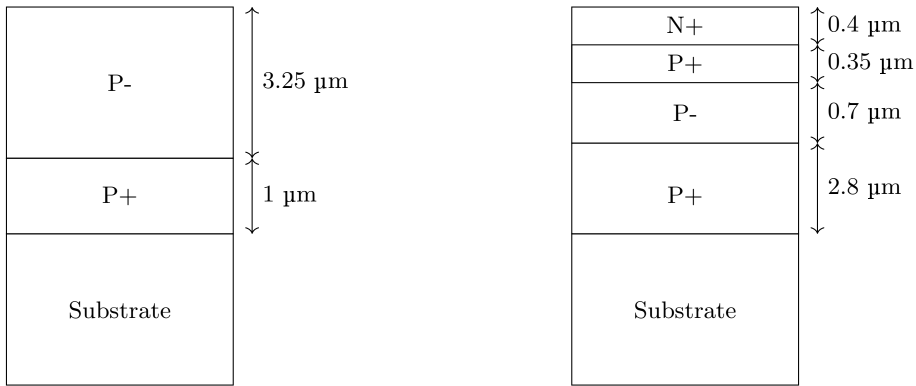

Figure 6.1: Schematic of wafer epitaxial structure (left) and after ion implantation and annealing (right).

using Sentaurus TCAD, to predict the expected doping profile due to the implantation, annealing, and oxidation steps. The achieved sample structure is analysed in section 6.3.

After photolithography to pattern, the Silicon sample was dry etched, using an ICP-RIE system, to form mesa devices. Then SiO_2 is deposited to the thickness of the mesas to act as sidewall passivation and electrical isolation between the bottom of the mesa and the top contact. The SiO_2 is then dry etched using an RIE to open up the device window and the P+ layer. Aluminium contacts and bond pads were then deposited using thermal evaporation to complete the fabrication process.

This chapter describes two sets of mesa diodes, which were fabricated from the same implanted and annealed material. The first set of mesa diodes (sample A), were fabricated externally by iNEX Microtechnology, Newcastle upon Tyne, who used the Sheffield NEW-PIN mask to fabricate circular mesas with radii of 25, 50, 100, and 200 µm. The second set (sample B) were fabricated at Sheffield to produce linear arrays of 128 pixels, where each pixel takes the form of a square mesa 40 µm across with 5 µm radial corners. Though the second set was fabricated into linear arrays, only single pixel performance will be discussed in this chapter. Images of sample A and B can be seen in Figure 6.2.

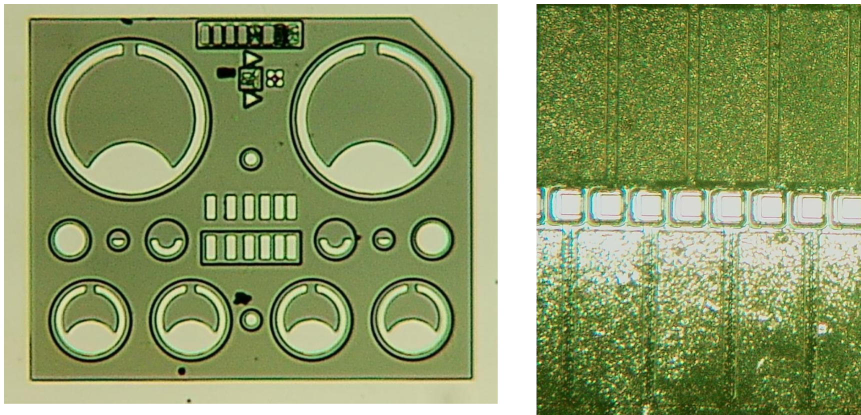

Figure 6.2: Microscope images of sample A (left) and sample B (right)

All the preparation work for sample A ion implantation and annealing processes was performed by myself, with external contractors performing ion implantation and high temperature annealing due to a lack of available equipment in the departmental device fabrication facilities, before the sample was shipped to iNEX Microtechnology for etching and contact deposition. For sample B, all the fabrication apart from ion implantation and annealing was performed by myself. Though this may sound simple, the fabrication facility is a III-V specialist facility that hadn't worked with Si, other than PECVD calibration, for many years, so I had to spend a significant amount of time re-establishing, calibrating, or establishing processes for Si work.

There were some planar Si SPADs planned as part of this work that featured the same main structure and doping profile to the mesa SPADs, to allow for easy comparison between the mesa and planar devices, with the addition of a deeply implanted P+ contact ring (sinker contact) around the active region. This sinker contact would be implanted deep enough to connect the buried P+ epitaxial layer to the surface eliminating the need for dry etching the samples to make a P-type contact. By eliminating the need for dry etching, a lower DCR than presented in Figure 6.10 should have been achievable. A new 10-layer mask set had to be designed for these devices, however due to a number of complications with fabrication and the facility these devices were never completed.

6.2 I-V

I-V measurements were performed upon sample A. The I-V results, shown in Figure 6.3, indicate a high contact resistance (forward bias data), but uniform reverse dark current and avalanche breakdown voltage for all diode sizes (reverse bias data). I-V results for sample B, shown in Figure 6.4, indicating a slightly higher contact resistance compared to the iNEX devices.

The series resistance from the sample B devices, was calculated to be 120 kΩ using Equation (3.5). The fitting to the forward I-V and the change in ideality factor are shown in Figure 6.5.

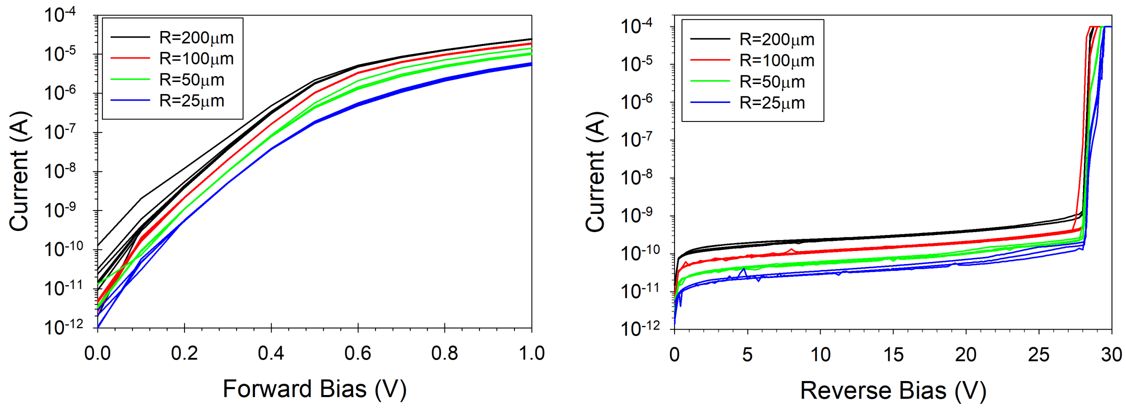

Figure 6.3: I-V results of sample A. Radius of diodes denoted by R in legend. Forward biased (left) and reversed biased (right).

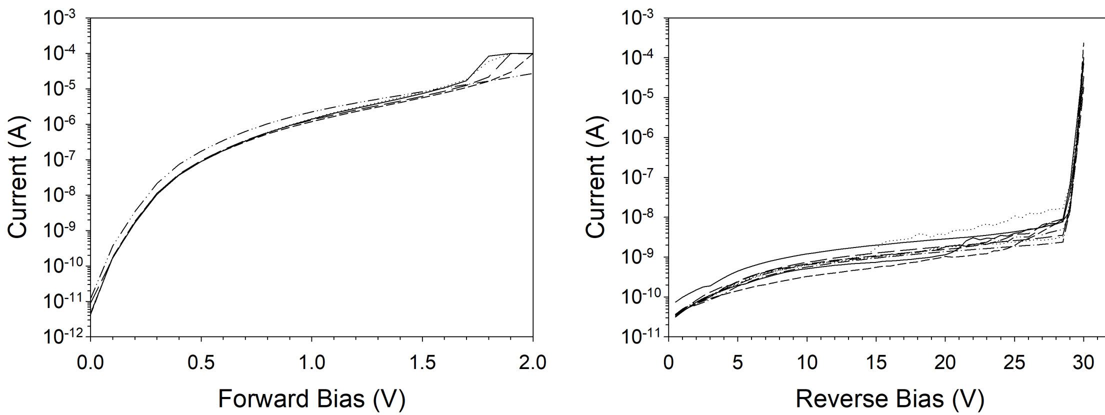

Figure 6.4: I-V results of sample B diodes forward biased (left) and reversed biased (right).

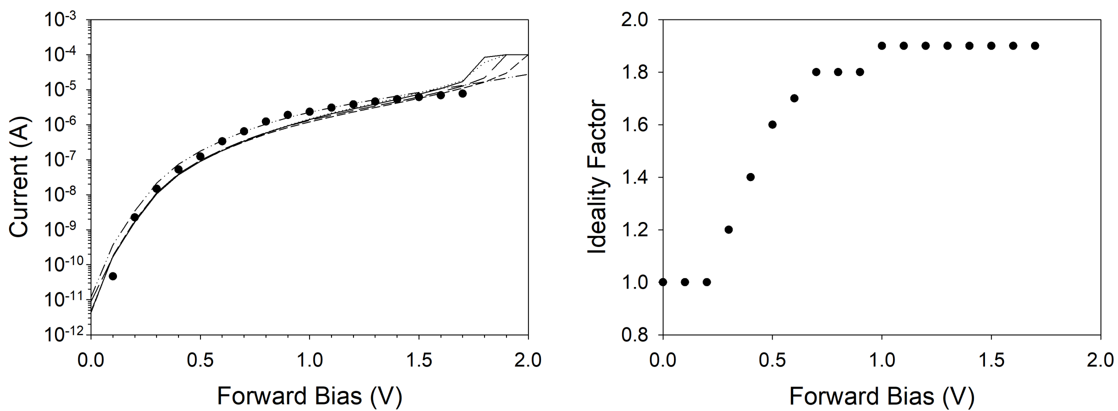

Figure 6.5: (Top) Fit (symbols) used to calculate a series resistance of 120 kΩ from the forward bias of the sample B mesa devices, and the obtained ideality factor (bottom)

6.3 C-V

C-V characteristics of the sample A devices scaled with area, indicating the devices had been etched to the expected size. Fitting to the C-V data of the 200 µm radius devices (from sample A) facilitated extraction of the doping profile. Using the electric field solver, described in appendix B, the electric field for a range of biases was calculated. Using the calculated electric field profiles, the capacitance could be extracted from the depletion region widths. The extracted capacitance values did not agree with the measured values, but a good fit could be obtained as shown in Figure 6.6, by scaling the simulated doping profile to account for incomplete electrical activation of the dopants. To achieve a good C-V fit a scaling factor of 0.35 was used, the difference between the simulated and measured profiles is shown in Figure 6.7.

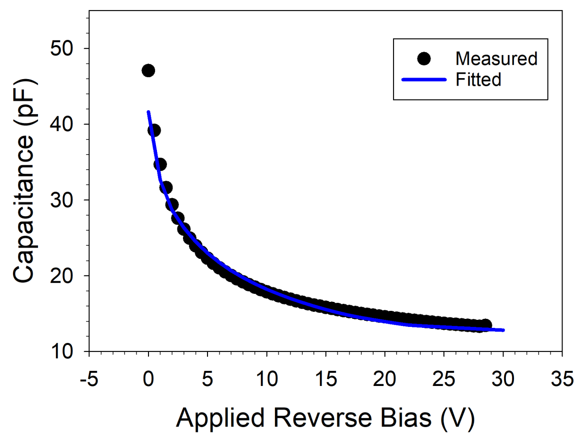

Figure 6.6: C-V data (symbols) and fitted (line) of 200 µm radius mesa from sample A. The fit used the adjusted doping profile shown in Figure 6.7.

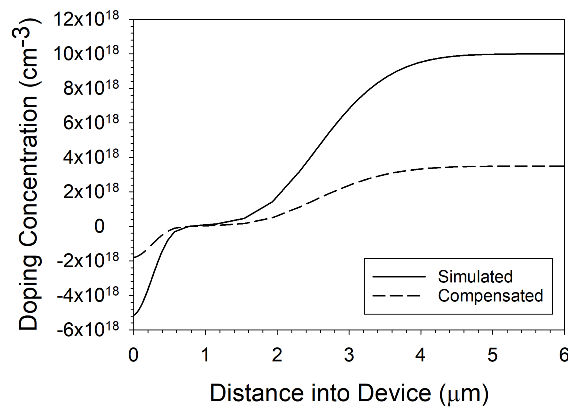

Figure 6.7: Si doping profiles originally simulated in Sentaurus TCAD (solid) and adjusted to fit C-V data (dashed).

6.4 Gain

Avalanche Gain measurements were performed on these devices using the method outlined in section 3.3 with a 633 nm HeNe laser as the illumination source. The sample B diodes were able to achieve uniform repeatable avalanche gain up to $M = 60$, with a maximum avalanche gain of 283 before undergoing avalanche breakdown. Avalanche gain results are shown in Figure 6.8, which also shows agreement to gain simulated using the SMC Simulator (chapter 4) and the doping profile shown in Figure 6.7. It should be noted that there is avalanche gain at reverse bias beyond the breakdown voltage suggested from the I-V data. This "increase" in the breakdown voltage is thought to be due to the series resistance resulting from the aluminium contacts used for this work. The series resistance acts in a similar manner to a simple passive quenching circuit (PQC) where a large quenching resistor is placed in series with the SPAD. As the photocurrent/ avalanche current increases a higher voltage is dropped across the SPAD, and a high series resistance contact, under high current conditions, can drop enough of the applied bias to keep the device below the breakdown voltage.

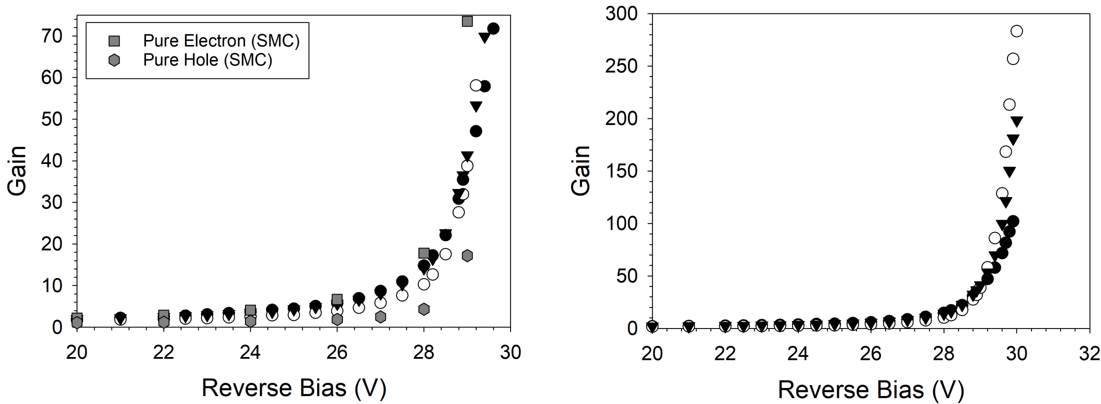

Figure 6.8: Gain data from 3 devices from sample B using 633 nm illumination. (Left) shows uniform gain to $M = 60$ with the addition of SMC simulated M using the simulated doping profile, whilst (Right) shows that the largest recorded M is 283.

6.5 DCR

Dark count measurements were performed on sample B using the Janis ST-500 probe station (section 3.4). The Janis ST-500 uses a combination of liquid nitrogen and electrical heating to maintain the sample stage at a constant temperature. To asses thermal instability, dark count was measured repeatedly, using back-to-back 10 s measurements, with the device reverse biased at 25.1 V with a fixed pulse height of 3 V, pulse duration of 20 ns, and a repetition frequency of 100 kHz. As shown in Figure 6.9, initially the count rate increases sharply, possibly the SPAD heating up due to breakdown, then the dark count rate decreases before reaching a stable level after around 5 minutes of operation.

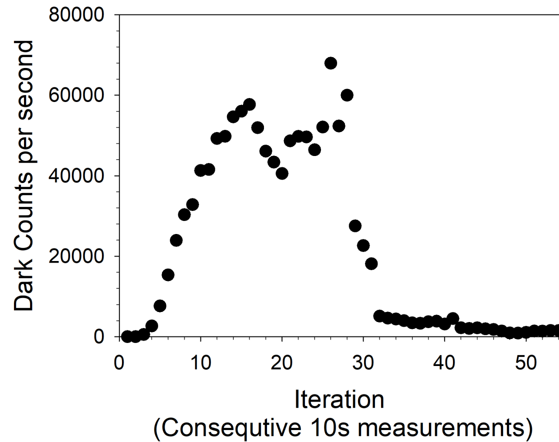

Figure 6.9: Temperature stability of dark count measurements at 25.1 V with an additional 3 V pulse

To counteract the thermal stability issues in subsequent measurements, each SPAD was operated under high DCR conditions for a minimum duration of 5 minutes to allow the device to reach a thermal equilibrium. In addition, for each condition, three consecutive measurements were taken to check for any wide variation.

Parts of sample B have poor quality devices that reached 100% dark counts from 0% within 0.7 V cf. devices which took approx. 2.1 V. The non-uniformity has been attributed to the silicon dry etching process, which created small pockets on the sample that were visibly rougher than other areas (the pockets, showed as surface discolouration, which is the beginning of black silicon formation). For future fabrication, the Si dry etching process needs to be refined to improve the yield of the process or the Si dry etching process needs to be outsourced to a more suitable facility rather than using a device fabrication facility predominately setup for III-V semiconductor fabrication. Increasing the amount of Argon in the Si plasma etch could help as it reduces the formation of black Si. In practise SPADs will usually not be operated to reach breakdown on 100% of the overbias pulses for counting purposes, however 10 devices were tested in this way to assess the robustness. Device failures would indicate a problem in fabrication and/or SPAD design. None of the SPADs tested failed prematurely when being pushed to a high overbias. The area of interest for SPAD operation is sub-40% dark count, so a more detailed measurement was performed on 3 devices, shown in Figure 6.10. There is reasonable uniformity in the dark count data (excluding the yield issues previously discussed).

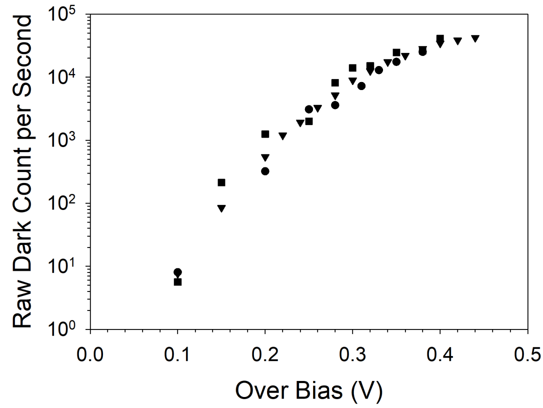

Figure 6.10: 0-40% DCR for 3 diodes from sample B.

6.6 Photon Counting

Photon counting measurements using a 633 nm pulsed laser were performed on the sample B devices cooled to 280 K (section 3.5). Optical pulses from the laser were attenuated to deliver 0.1 photons per pulse. The value of 0.1 photons per pulse was chosen as it shows the extremely low photon level detectable by the device, and that the number of incident photons over a 10 s measurement, using a 100 kHz repetition rate, is 100000 (making it easier to quickly interoperate raw count results). This attenuation was achieved using a 50:50 fibre optic coupler, and two electronically variable optical attenuators (EVOA). A schematic of the light paths can be seen in Figure 6.11. The 50:50 coupler was required to couple a 633 nm He-Ne laser into the same fibre path, while minimising the disruption to the light path caused by breaking/remaking fibre connections. This laser was used to aid aligning the output fibre with the active area of the SPAD as the fast pulses can't be detected with a USB camera.

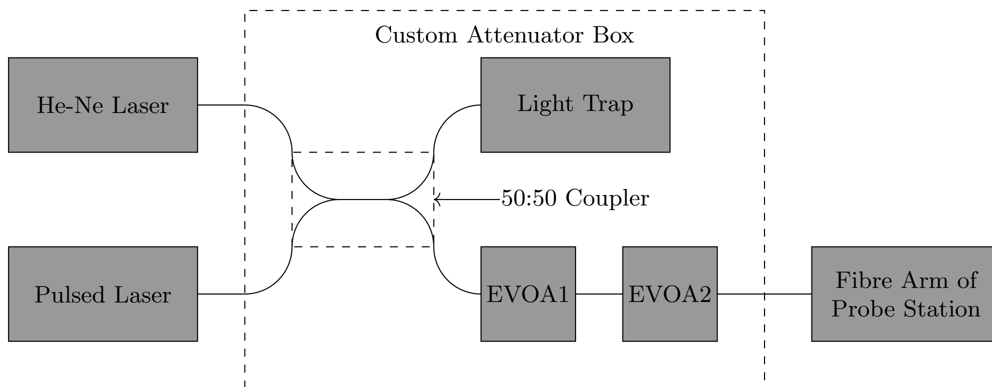

Figure 6.11: Schematic showing optical paths for photon counting setup

To reach an attenuation level corresponding to 0.1 photons per pulse, the setup's attenuation factors must be known accurately. The fibre arm of the ST-500 and the custom attenuation box, were found to have an attenuation of -5.7 dB and -10.9 dB respectively with the EVOAs off. A significant contribution to the attenuation of the custom attenua-

tor box comes from the 50:50 coupler which directs 50% of the incoming light into a light trap. Interface losses from the 4 fibre connections are the next largest contribution to the attenuation of the custom attenuator box. Interface losses could potentially be reduced by remaking all interfaces using an index matching gel at the interface connections. At a 20% tune power, operating at 100 kHz, the 633 nm pulsed laser head emitted an average power of 54 nW (measured with a ThorLabs PM100D and S150C), requiring an additional -56 dB attenuation from the EVOAs to achieve an average of 0.1 photons per pulse.

Measuring 3 devices from sample B at the 0.1 photon per pulse level, for 10 second count durations, taking three measurements per voltage, with a 3 V pulse, and 100 kHz repetition rate resulted in the average counts per second shown in Figure 6.12 at 280 K. Counts were consistent between the devices for the over-bias used.

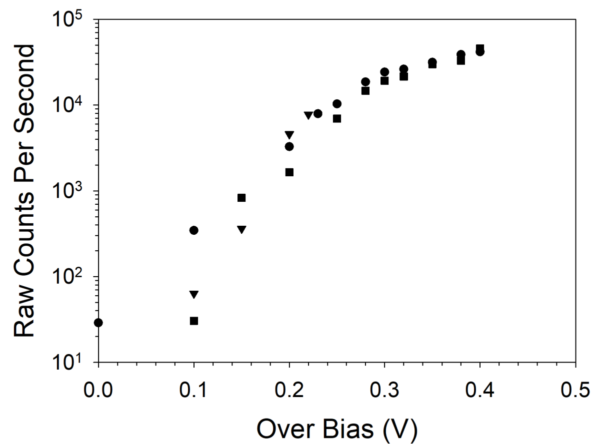

Figure 6.12: Photon and Dark counts per second from 3 devices from sample B, when operated at 100 kHz, with an incoming optical power of 0.1 photons per pulse at 280 K.

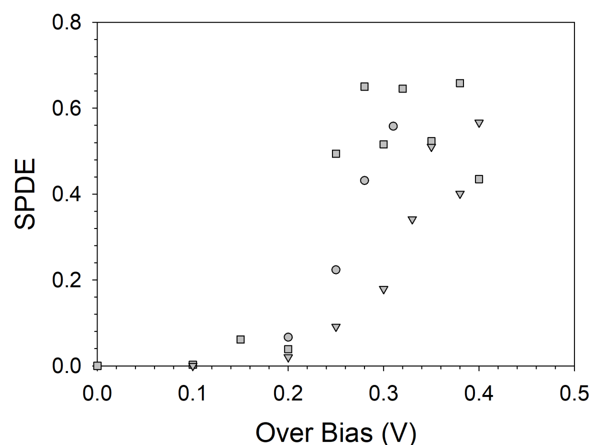

Figure 6.13: SPDE from 3 devices from sample B, when operated at 100 kHz, with 0.1 photons per pulse for photon counting at 280 K.

Single photon detection efficiency (SPDE) was calculated using the data in Figure 6.12 using Equation (3.9). This resulted in a maximum SPDE of 69% at 633 nm, despite the relatively high dark count rate. Calculated SPDE against over-bias is shown in Figure 6.13.

There are three factors that may affect the consistency of SPDE between devices. The first factor is the alignment of the laser spot onto the centre of the device and the second is the uniformity of the mesa etch. The mesa etch surface showed clear signs of variation over the sample, which is thought to be due to the configuration of the ICP-RIE system. The third factor is that for each measured device the breakdown voltage is calculated as the final voltage which achieved 0% DCR, this variation is thought to be due to the etch surface.

6.7 Chapter Summary

In this chapter, a square mesa SPAD has been presented. The mesa SPAD achieved an SPDE of 69% at 633 nm despite a high dark count rate, and limited usable overbias range of 0.4 V. This shows that mesa SPADs offer the alternative to more complex planar SPAD designs.

Chapter 7

Conclusion & Future Work

7.1 Conclusion

A simple monte carlo simulator simulator for SPADs has been developed. The simulator was implemented for Si SPADs using an existing Si parameter set. The model was extensively validated against experimental Si APD characteristics, $M(V)$ and $F(M)$, along with reported drift velocities, diffusion coefficients, and impact ionisation coefficients. The SPAD characteristics were validated against simpler models using effective ionisation coefficients (extracted from the Si simple monte carlo model).

The Si SPAD model was then used to assess the SPAD characteristics of p-on-n and n-on-p designs. The n-on-p design was found to offer a better timing performance for a given breakdown probability, however the p-on-n design achieved a greater breakdown probability for a given bias.

Using the simple monte carlo simulator, a new temperature-dependent simple monte carlo parameter set has been presented for InP APDs. This parameter set was validated for the temperature dependent modelling of InP between 150 and 290 K for electric fields at 400-800 kV.cm^{-1}. The validation data included avalanche gain and excess noise factor data from 7 InP APDs (with intrinsic region thicknesses of 0.125 - 2.5 µm). Also, effective ionisation coefficients have been extracted for the use with simpler models at 150, 200, 250, and 290 K.

Finally, a 40 µm $\times$ 40 µm Si mesa SPAD has been demonstrated. This Si mesa SPAD shows a breakdown voltage $<$ 30 V and despite its high dark count rate demonstrates a

single photon detection efficiency of 69% for 633 nm photons when operated at 280 K.

Analysis of the Si mesa SPAD device characteristics was performed using the simple monte carlo simulator. It was found that the experimentally measured gain has compared favourably to the simulated gain.

7.2 Future Work

As it stands the SMC Simulator, presented in chapter 4, is able to achieve a good agreement with experimental avalanche gain results for the Si Mesa SPADs, shown in chapter 6, but due to setup issues other comparisons are not possible. To enable future comparison of timing jitter, breakdown probability, and SPDE the Mesa SPADs need to be packaged. Packaging them will allow them to be used in the black box setup, described in section 3.4, rather than the Janis probe station. By doing this, it is hoped the devices can be impedance matched which would reduce the pulse distortion currently seen on the Janis setup. To aid the impedance matching process the contact series resistance needs to be reduced. This reduction could be achieved by either 1) trying a different contact deposition method, such as sputter coating, as it is believed that some of the tungsten filament wire used in the thermal evaporation process has alloyed with the Al or 2) if altering the deposition method is unsuccessful, increasing the implant doses to increase the dopant concentration at the wafer surface.

Though the SPDE of the Mesa SPADs was respectable, the DCR from them was quite high. Before packaging them it would be worth refining the dry etching process or outsourcing the process to a fabrication facility better equipped for fabricating Si. Completing the fabrication for the originally planned planar devices would allow for a comparison between the DCR from the mesa and planar structures.

Extending the SMC simulator to include the simulation of a ballast resistor to passively quench the SPAD would enable the optimised ballast resistor selection to turn the Mesa or Planar SPADs into a complete passively quenched detector module. Further characterisation of Si mesa SPADs including measuring more pixels at 633 nm and using a second wavelength would improve information upon the sample yield and consistency of the Si mesa SPADs. The Si mesa arrays could be fabricated upon SOI, rather than on a conductive substrate to enable complete electrical isolation between the diodes in the array.

For the InP work, it would be interesting to characterise a set of InP pin diodes as SPADs. The SMC simulator used for the temperature dependent InP work has SPAD functionality, however due to a lack of InP only SPAD results with a known doping profile this functionality has never been able to be benchmarked against any experimental results. Not only would this benchmarking allow for the validation of anther part of the SMC Simulator, it would potentially allow studies to compare the separate influences of InP and $In_{0.53}Ga_{0.47}As$regions in $In_{0.53}Ga_{0.47}As$/InP SAM-SPADs. Better understanding these regions contributions separately could allow for improvements in $In_{0.53}Ga_{0.47}As$/InP SPAD design.

References

[1] A. Einstein, "Über einen die erzeugung und verwandlung des lichtes betreffenden heuristischen gesichtspunkt," *Ann. Phys.*, vol. 322, no. 6, pp. 132–148, 1905.

[2] J. Wojtanowski, M. Zygmunt, M. Kaszczuk, Z. Mierczyk, and M. Muzal, "Comparison of 905 nm and 1550 nm semiconductor laser rangefinders' performance deterioration due to adverse environmental conditions," *Opto-Electronics Rev.*, vol. 22, no. 3, Jan. 2014.

[3] C. Godard, O. M. Aodha, and G. J. Brostow, "Unsupervised monocular depth estimation with left-right consistency," in *CVPR*, 2017.

[4] Y. Wang, W. Chao, D. Garg, B. Hariharan, M. Campbell, and K. Q. Weinberger, "Pseudo-lidar from visual depth estimation:bridging the gap in 3d object detection for autonomous driving," in *CVPR*, 2019.

[5] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.

[6] M. Menze and A. Geiger, "Object scene flow for autonomous vehicles," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

[7] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, "Quantum cryptography," *Rev. Mod. Phys.*, vol. 74, pp. 145–195, Mar. 2002.

[8] W. K. Wootters and W. H. Zurek, "A single quantum cannot be cloned," *Nature*, vol. 299, no. 5886, pp. 802–803, Oct. 1982.

[9] W. P. Grice, P. G. Evans, B. Lawrie, M. Legré, P. Lougovski, W. Ray, B. P. Williams, B. Qi, and A. M. Smith, "Two-party secret key distribution via a modified quantum secret sharing protocol," *Opt. Express*, vol. 23, no. 6, p. 7300, Mar. 2015.

[10] M. Hillery, V. Bužek, and A. Berthiaume, "Quantum secret sharing," *Phys. Rev. A*, vol. 59, no. 3, pp. 1829–1834, Mar. 1999.

[11] M. Minder, M. Pittaluga, G. L. Roberts, M. Lucamarini, J. F. Dynes, Z. L. Yuan, and A. J. Shields, "Experimental quantum key distribution beyond the repeaterless secret key capacity," *Nat. Photon.*, vol. 13, no. 5, pp. 334–338, Mar. 2019.

[12] M. Lucamarini, Z. L. Yuan, J. F. Dynes, and A. J. Shields, "Overcoming the rate–distance limit of quantum key distribution without quantum repeaters," *Nature*, vol. 557, no. 7705, pp. 400–403, May 2018.

[13] D. Boas, D. Brooks, E. Miller, C. DiMarzio, M. Kilmer, R. Gaudette, and Q. Zhang, "Imaging the body with diffuse optical tomography," *IEEE Signal Process. Mag.*, vol. 18, no. 6, pp. 57–75, 2001.

[14] W. Cong, X. Intes, and G. Wang, "Optical tomographic imaging for breast cancer detection," *J. Biomed. Opt.*, vol. 22, no. 09, p. 1, Sep. 2017.

[15] M. Kacprzak, A. Liebert, P. Sawosz, R. Maniewski, W. Staszkiewicz, A. Gabrusiewicz, and G. Madycki, "Application of a time-resolved optical brain imager for monitoring cerebral oxygenation during carotid surgery," *J. Biomed. Opt.*, vol. 17, no. 1, pp. 1 – 11, 2012.

[16] P. Matousek and N. Stone, "Recent advances in the development of raman spectroscopy for deep non-invasive medical diagnosis," *J. Biophotonics*, vol. 6, no. 1, pp. 7–19, Nov. 2013.

[17] D. V. Martyshkin, R. C. Ahuja, A. Kudriavtsev, and S. B. Mirov, "Effective suppression of fluorescence light in raman measurements using ultrafast time gated charge coupled device camera," *Rev. Sci. Instrum.*, vol. 75, no. 3, pp. 630–635, Mar. 2004.

[18] C. V. Raman and K. S. Krishnan, "A new type of secondary radiation," *Nature*, vol. 121, no. 3048, pp. 501–502, Mar. 1928.

[19] R. Y. Chiao, C. H. Townes, and B. P. Stoicheff, "Stimulated brillouin scattering and coherent generation of intense hypersonic waves," *Phys. Rev. Lett.*, vol. 12, no. 21, pp. 592–595, May 1964.

[20] G. S. Bumbrah and R. M. Sharma, "Raman spectroscopy – basic principle, instrumentation and selected applications for the characterization of drugs of abuse," *Egypt. J. Forensic Sci.*, vol. 6, no. 3, pp. 209–215, Sep. 2016.

[21] M. J. Foster, J. Storey, and J. Lapington, "Raman spectroscopy," WO Patent WO2018/015 309A1, Jan., 2018.

[22] E. Abdollahi, G. Taucher-Scholz, and B. Jakob, "Application of fluorescence lifetime imaging microscopy of DNA binding dyes to assess radiation-induced chromatin compaction changes," *Int. J. Mol. Sci*, vol. 19, no. 8, p. 2399, Aug. 2018.

[23] PiqoQuant. (2017) FLIMbee. [Online]. Available: https://www.picoquant.com/images/uploads/downloads/flimbee_brochure.pdf [Accessed: 2019-11-11]

[24] D. Renker, "Geiger-mode avalanche photodiodes, history, properties and problems," *Nuc. Instrum. Methods Phys. Res. A*, vol. 567, no. 1, pp. 48–56, Nov. 2006.

[25] Hamamatsu. (2013, Feb.) Photomultiplier tubes R2083, R3377. [Online]. Available: https://www.hamamatsu.com/resources/pdf/etd/R2083_R3377_TPMH1227E.pdf [Accessed: 2019-02-06]

[26] Hamamatsu. (2014, Aug.) Photomultiplier tubes R877, R877-01. [Online]. Available: https://www.hamamatsu.com/resources/pdf/etd/R877_TPMH1112E.pdf [Accessed: 2019-02-06]

[27] Hamamatsu. (2018, Oct.) High speed compact hpd (hybrid photo detector) SERIES R10467Useries/ R11322U-40/ H13223-40. [Online]. Available: https://www.hamamatsu.com/resources/pdf/etd/HPD_TPMH1361E.pdf [Accessed: 2019-02-06]

[28] Hamamatsu. (2016, Jan.) MCP (microchannel plate) and MCP assembly. [Online]. Available: https://www.hamamatsu.com/resources/pdf/etd/MCP_TMCP0002E.pdf [Accessed: 2019-02-06]

[29] F. Acerbi, M. Anti, A. Tosi, and F. Zappa, "Design criteria for InGaAs/InP single-photon avalanche diode," *IEEE Photon. J.*, vol. 5, no. 2, pp. 6 800 209–6 800 209, Apr. 2013.

[30] X. Zhou, L. L. G. Pinel, S. J. Dimler, S. Zhang, J. S. Ng, and C. H. Tan, "Thin $Al_{1-x}Ga_xAs_{0.56}Sb_{0.44}$ diodes with low excess noise," *IEEE J. Sel. Top. Quantum Electron.*, vol. 24, no. 2, pp. 1–5, Mar. 2018.

[31] Hamamatsu. (2016, Aug.) MPPC (multi-pixel photon counter) S13360 series. [Online]. Available: https://www.hamamatsu.com/resources/pdf/ssd/s13360_series_kapd1052e.pdf [Accessed: 2019-04-13]

[32] Ketek. (2019) Sipm – silicon photomultiplier pm1125-wb-b0. [Online]. Available: https://www.ketek.net/wp-content/uploads/KETEK-PM1125-WB-B0-Datasheet.pdf [Accessed: 2019-08-19]

[33] Ketek. (2019) Sipm – silicon photomultiplier pm3315-wb-b0. [Online]. Available: https://www.ketek.net/wp-content/uploads/2018/12/KETEK-PM3315-WB-B0-Datasheet.pdf [Accessed: 2019-08-19]

[34] A. D. Semenov, G. N. Gol'tsman, and A. A. Korneev, "Quantum detection by current carrying superconducting film," *Physica C Supercond.*, vol. 351, no. 4, pp. 349–356, Apr. 2001.

[35] O. Kahl, S. Ferrari, V. Kovalyuk, G. N. Goltsman, A. Korneev, and W. H. P. Pernice, "Waveguide integrated superconducting single-photon detectors with high internal quantum efficiency at telecom wavelengths," *Sci. Rep.*, vol. 5, no. 1, Jun. 2015.

[36] A. C. Farrell, X. Meng, D. Ren, H. Kim, P. Senanayake, N. Y. Hsieh, Z. Rong, T.-Y. Chang, K. M. Azizur-Rahman, and D. L. Huffaker, "InGaAs–GaAs nanowire avalanche photodiodes toward single-photon detection in free-running mode," *Nano Lett.*, vol. 19, no. 1, pp. 582–590, Dec. 2018.

[37] K. G. McKay, "Avalanche breakdown in silicon," *Phys. Rev.*, vol. 94, no. 4, pp. 877–884, May 1954.

[38] A. G. Chynoweth and K. G. McKay, "Photon emission from avalanche breakdown in silicon," *Phys. Rev.*, vol. 102, no. 2, pp. 369–376, Apr. 1956.

[39] R. L. Batdorf, A. G. Chynoweth, G. C. Dacey, and P. W. Foy, "Uniform silicon p-n junctions. i. broad area breakdown," *J. Appl. Phys.*, vol. 31, no. 7, pp. 1153–1160, Jul. 1960.

[40] K. Johnson, "Photodiode signal enhancement effect at avalanche breakdown voltage," in *1964 IEEE International Solid-State Circuits Conference. Digest of Technical Papers.* Institute of Electrical and Electronics Engineers, 1964.

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

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

[43] Y. Okuto and C. R. Crowell, "Ionization coefficients in semiconductors: A nonlocalized property," *Phys. Rev. B*, vol. 10, no. 10, pp. 4284–4296, Nov. 1974.

[44] S. Cova, M. Ghioni, A. Lacaita, C. Samori, and F. Zappa, "Avalanche photodiodes and quenching circuits for single-photon detection," *Appl. Opt.*, vol. 35, no. 12, p. 1956, Apr. 1996.

[45] C. Scarcella, G. Boso, A. Ruggeri, and A. Tosi, "InGaAs/InP single-photon detector gated at 1.3 ghz with 1.5% afterpulsing," *IEEE J. Sel. Top. Quantum Electron.*, vol. 21, no. 13, pp. 17–22, May 2015.

[46] S. Cova, A. Longoni, and G. Ripamonti, "Active-quenching and gating circuits for single-photon avalanche diodes (spads)," *IEEE Trans. Nucl. Sci.*, vol. 29, no. 1, pp. 599–601, 1982.

[47] S. Cova, A. Longoni, A. Andreoni, and R. Cubeddu, *IEEE J. Quantum Electron.*

[48] D. Bronzi, F. Villa, S. Tisa, A. Tosi, and F. Zappa, "SPAD figures of merit for photon-counting, photon-timing, and imaging applications: A review," *IEEE Sens. J.*, vol. 16, no. 1, pp. 3–12, Jan. 2016.

[49] F. Zappa, S. Tisa, A. Tosi, and S. Cova, "Principles and features of single-photon avalanche diode arrays," *Sens. Actuators A*, vol. 140, no. 1, pp. 103–112, Oct. 2007.

[50] M. A. Itzler, r. Ben-Michael, C. F. Hsu, K. Slomkowski, A. Tosi, S. Cova, F. Zappa, and R. Ispasoiu, "Single photon avalanche diodes (SPADs) for 1.5 µm photon counting applications," *J. Mod. Opt.*, vol. 54, no. 2-3, pp. 283–304, Jan. 2007.

[51] G. Hurkx, D. Klaassen, and M. Knuvers, "A new recombination model for device simulation including tunneling," *IEEE Trans. Electron Devices*, vol. 39, no. 2, pp. 331–338, 1992.

[52] W. Kindt, "Geiger mode avalanche photodiode arrays," Ph.D. dissertation, Delft, 1999.

[53] G. Hurkx, "On the modelling of tunnelling currents in reverse-biased p-n junctions," *Solid State Electron.*, vol. 32, no. 8, pp. 665–668, Aug. 1989.

[54] S. Yanikgonul, V. Leong, J. R. Ong, C. E. Png, and L. Krivitsky, "Simulation of silicon waveguide single-photon avalanche detectors for integrated quantum photonics," *IEEE J. Sel. Top. Quantum Electron.*, vol. 26, no. 2, pp. 1–8, 2020.

[55] I. Rech, A. Ingargiola, R. Spinelli, I. Labanca, S. Marangoni, M. Ghioni, and S. Cova, "Optical crosstalk in single photon avalanche diode arrays: a new complete model," *Opt. Express*, vol. 16, no. 12, p. 8381, May 2008.

[56] S. Cova, A. Lacaita, and G. Ripamonti, "Trapping phenomena in avalanche photodiodes on nanosecond scale," *IEEE Electron Device Lett.*, vol. 12, no. 12, pp. 685–687, Dec. 1991.

[57] M. Stipčević, "Active quenching circuit for single-photon detection with geiger mode avalanche photodiodes," *Appl. Opt.*, vol. 48, no. 9, p. 1705, Mar. 2009.

[58] M. A. Itzler, X. Jiang, and M. Entwistle, "Power law temporal dependence of InGaAs/InP SPAD afterpulsing," *J. Mod. Opt.*, vol. 59, no. 17, pp. 1472–1480, Oct. 2012.

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

[60] A. W. Ziarkash, S. K. Joshi, M. Stipčević, and R. Ursin, "Comparative study of afterpulsing behavior and models in single photon counting avalanche photo diode detectors," *Sci. Rep.*, vol. 8, no. 1, Mar. 2018.

[61] F. Christnacher, G. S. Buller, R. Tobin, A. Halimi, A. McCarthy, and M. Laurenzis, "Depth imaging through obscurants using time-correlated single-photon counting," in *Advanced Photon Counting Techniques XII*, M. A. Itzler and J. C. Campbell, Eds. SPIE, May 2018.

[62] A. Spinelli and A. Lacaita, "Physics and numerical simulation of single photon avalanche diodes," *IEEE Trans. Electron Devices*, vol. 44, no. 11, pp. 1931–1943, 1997.

[63] F. Acerbi, M. Cazzanelli, A. Ferri, A. Gola, L. Pavesi, N. Zorzi, and C. Piemonte, "High detection efficiency and time resolution integrated-passive-quenched single-photon avalanche diodes," *IEEE J. Sel. Top. Quantum Electron.*, vol. 20, no. 6, pp. 268–275, Nov. 2014.

[64] J. D. Petticrew, S. J. Dimler, X. Zhou, A. P. Morrison, C. H. Tan, and J. S. Ng, "Avalanche breakdown timing statistics for silicon single photon avalanche diodes," *IEEE J. Sel. Top. Quantum Electron.*, vol. 24, no. 2, pp. 1–6, Mar. 2018.

[65] I. Rech, D. Resnati, A. Gulinatti, M. Ghioni, and S. Cova, "Self-suppression of reset induced triggering in picosecond SPAD timing circuits," *Rev. Sci. Instrum.*, vol. 78, no. 8, p. 086112, Aug. 2007.

[66] J. Jackson, A. Morrison, D. Phelan, and A. Mathewson, "A novel silicon geiger-mode avalanche photodiode," in *Digest. International Electron Devices Meeting,*. IEEE, 2002.

[67] H. J. Geipel and W. K. Tice, "Critical microstructure for ion-implantation gettering effects in silicon," *Appl. Phys. Lett.*, vol. 30, no. 7, pp. 325–327, Apr. 1977.

[68] J. S. Kang and D. K. Schroder, "Gettering in silicon," *J. Appl. Phys.*, vol. 65, no. 8, pp. 2974–2985, Apr. 1989.

[69] S. Cova, A. Longoni, and A. Andreoni, "Towards picosecond resolution with single-photon avalanche diodes," *Rev. Sci. Instrum.*, vol. 52, no. 3, pp. 408–412, Mar. 1981.

[70] R. H. Haitz, "Mechanisms contributing to the noise pulse rate of avalanche diodes," *J. Appl. Phys.*, vol. 36, no. 10, pp. 3123–3131, Oct. 1965.

[71] S. Cova, G. Ripamonti, and A. Lacaita, "Avalanche semiconductor detector for single optical photons with a time resolution of 60 ps," *Nuc. Instrum. Methods Phys. Res. A*, vol. 253, no. 3, pp. 482–487, Jan. 1987.

[72] A. Lacaita, M. Ghioni, and S. Cova, "Double epitaxy improves single-photon avalanche diode performance," *Electron. Lett.*, vol. 25, no. 13, p. 841, 1989.

[73] M. Ghioni, S. Cova, A. Lacaita, and G. Ripamonti, "New silicon epitaxial avalanche diode for single-photon timing at room temperature," *Electron. Lett.*, vol. 24, no. 24, p. 1476, 1988.

[74] C. Cammi, F. Panzeri, A. Gulinatti, I. Rech, and M. Ghioni, "Custom single-photon avalanche diode with integrated front-end for parallel photon timing applications," *Rev. Sci. Instrum.*, vol. 83, no. 3, p. 033104, Mar. 2012.

[75] F. Ceccarelli, G. Acconcia, A. Gulinatti, M. Ghioni, and I. Rech, "83-ps timing jitter with a red-enhanced SPAD and a fully integrated front end circuit," *IEEE Photonics Technol. Lett.*, vol. 30, no. 19, pp. 1727–1730, Oct. 2018.

[76] F. Ceccarelli, A. Gulinatti, I. Labanca, M. Ghioni, and I. Rech, "Red-enhanced photon detection module featuring a 32x1 single-photon avalanche diode array," *IEEE Photonics Technol. Lett.*, vol. 30, no. 6, pp. 557–560, Mar. 2018.

[77] F. Nolet, S. Parent, N. Roy, M.-O. Mercier, S. Charlebois, R. Fontaine, and J.-F. Pratte, "Quenching circuit and SPAD integrated in CMOS 65 nm with 7.8 ps FWHM single photon timing resolution," *Instruments*, vol. 2, no. 4, p. 19, Sep. 2018.

[78] D. P. Palubiak and M. J. Deen, "CMOS SPADs: Design issues and research challenges for detectors, circuits, and arrays," *IEEE J. Sel. Top. Quantum Electron.*, vol. 20, no. 6, pp. 409–426, Nov. 2014.

[79] Synopsys. Avalanche photodiodes for industrial & analytical applications. [Online]. Available: https://www.pacer.co.uk/Assets/User/1626-Analytical_APDs.pdf [Accessed: 2019-08-14]

[80] S. Lindner, S. Pellegrini, Y. Henrion, B. Rae, M. Wolf, and E. Charbon, "A high-PDE, backside-illuminated SPAD in 65/40-nm 3d IC CMOS pixel with cascoded passive quenching and active recharge," *IEEE Electron Device Lett.*, vol. 38, no. 11, pp. 1547–1550, Nov. 2017.

[81] E. Charbon, C. Bruschini, and M.-J. Lee, "3d-stacked CMOS SPAD image sensors: Technology and applications," in *2018 25th IEEE International Conference on Electronics, Circuits and Systems (ICECS).* IEEE, Dec. 2018.

[82] M.-J. Lee, A. R. Ximenes, P. Padmanabhan, T.-J. Wang, K.-C. Huang, Y. Yamashita, D.-N. Yaung, and E. Charbon, "High-performance back-illuminated three-dimensional stacked single-photon avalanche diode implemented in 45-nm CMOS technology," *IEEE J. Sel. Top. Quantum Electron.*, vol. 24, no. 6, pp. 1–9, Nov. 2018.

[83] R. K. Henderson, N. Johnston, F. M. D. Rocca, H. Chen, D. D.-U. Li, G. Hungerford, R. Hirsch, D. Mcloskey, P. Yip, and D. J. S. Birch, "A 192 x 128 time correlated SPAD image sensor in 40-nm CMOS technology," *IEEE J. Solid-State Circuits*, vol. 54, no. 7, pp. 1907–1916, Jul. 2019.

[84] E. Charbon, H.-J. Yoon, and Y. Maruyama, "A geiger mode APD fabricated in standard 65nm CMOS technology," in *2013 IEEE International Electron Devices Meeting.* IEEE, Dec. 2013.

[85] M. Sanzaro, F. Signorelli, P. Gattari, A. Tosi, and F. Zappa, "0.16 µm BCD silicon photomultipliers with sharp timing response and reduced correlated noise," *Sensors*, vol. 18, no. 11, p. 3763, Nov. 2018.

[86] C. Niclass, M. Soga, H. Matsubara, S. Kato, and M. Kagami, "A 100-m range 10-frame/s 340 x 96-pixel time-of-flight depth sensor in 0.18-µm CMOS," *IEEE J. Solid-State Circuits*, vol. 48, no. 2, pp. 559–572, Feb. 2013.

[87] C. Niclass, M. Soga, H. Matsubara, M. Ogawa, and M. Kagami, "A 0.18-µm CMOS SoC for a 100-m-range 10-frame/s 200 x 96-pixel time-of-flight depth sensor," *IEEE J. Solid-State Circuits*, vol. 49, no. 1, pp. 315–330, Jan. 2014.

[88] I. Takai, H. Matsubara, M. Soga, M. Ohta, M. Ogawa, and T. Yamashita, "Single-photon avalanche diode with enhanced NIR-sensitivity for automotive LIDAR systems," *Sensors*, vol. 16, no. 4, p. 459, Mar. 2016.

[89] Y. Zou, F. Villa, D. Bronzi, S. Tisa, A. Tosi, and F. Zappa, "Planar CMOS analog SiPMs: design, modeling, and characterization," *J. Mod. Opt.*, vol. 62, no. 20, pp. 1693–1702, Jun. 2015.

[90] F. Villa, R. Lussana, D. Bronzi, S. Tisa, A. Tosi, F. Zappa, A. D. Mora, D. Contini, D. Durini, S. Weyers, and W. Brockherde, "CMOS imager with 1024 SPADs and TDCs for single-photon timing and 3-d time-of-flight," *IEEE J. Sel. Top. Quantum Electron.*, vol. 20, no. 6, pp. 364–373, Nov. 2014.

[91] F. Villa, D. Bronzi, Y. Zou, C. Scarcella, G. Boso, S. Tisa, A. Tosi, F. Zappa, D. Durini, S. Weyers, U. Paschen, and W. Brockherde, "CMOS SPADs with up to 500 µm meter and 55% detection efficiency at 420 nm," *J. Mod. Opt.*, vol. 61, no. 2, pp. 102–115, Jan. 2014.

[92] H. Ruokamo, L. W. Hallman, and J. Kostamovaara, "An 80 x 25 pixel CMOS single-photon sensor with flexible on-chip time gating of 40 subarrays for solid-state 3-d range imaging," *IEEE J. Solid-State Circuits*, vol. 54, no. 2, pp. 501–510, Feb. 2019.

[93] D. Stoppa, L. Pancheri, M. Scandiuzzo, L. Gonzo, G.-F. D. Betta, and A. Simoni, "A CMOS 3-d imager based on single photon avalanche diode," *IEEE Trans. Circuits Syst. I*, vol. 54, no. 1, pp. 4–12, Jan. 2007.

[94] C. Niclass, A. Rochas, P.-A. Besse, and E. Charbon, "Design and characterization of a CMOS 3-d image sensor based on single photon avalanche diodes," *IEEE J. Solid-State Circuits*, vol. 40, no. 9, pp. 1847–1854, Sep. 2005.

[95] J. Jackson, J. Donnelly, B. O'Neill, A.-M. Kelleher, G. Healy, A. Morrison, and A. Mathewson, "Integrated bulk/SOI APD sensor: bulk substrate inspection with geiger-mode avalanche photodiodes," *Electronics Letters*, vol. 39, no. 9, p. 735, 2003.

[96] BS EN 60825-1, "Safety of laser products. part 1: Equipment classification and requirements," British Standards Institute, London, UK, Standard, 2014.

[97] J. Zhang, M. A. Itzler, H. Zbinden, and J.-W. Pan, "Advances in InGaAs/InP single-photon detector systems for quantum communication," *Light Sci. Appl.*, vol. 4, no. 5, pp. e286–e286, May 2015.

[98] L. J. J. Tan, D. S. G. Ong, J. S. Ng, C. H. Tan, S. K. Jones, Y. Qian, and J. P. R. David, "Temperature dependence of avalanche breakdown in InP and InAlAs," *IEEE J. Quantum Electron.*, vol. 46, no. 8, pp. 1153–1157, Aug. 2010.

[99] G. Karve, S. Wang, F. Ma, X. Li, J. C. Campbell, R. G. Ispasoiu, D. S. Bethune, W. P. Risk, G. S. Kinsey, J. C. Boisvert, T. D. Isshiki, and R. Sudharsanan, "Origin of dark counts in $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ avalanche photodiodes operated in geiger mode," *Appl. Phys. Lett.*, vol. 86, no. 6, p. 063505, 2005.

[100] R. McIntyre, "Multiplication noise in uniform avalanche diodes," *IEEE Trans. Electron Devices*, vol. ED-13, no. 1, pp. 164–168, Jan. 1966.

[101] G. J. Rees and J. P. R. David, "Nonlocal impact ionization and avalanche multiplication," *J. Phys. D*, vol. 43, no. 24, p. 243001, Jun. 2010.

[102] M. Hayat, B. Saleh, and M. Teich, "Effect of dead space on gain and noise of double-carrier-multiplication avalanche photodiodes," *IEEE Trans. Electron Devices*, vol. 39, no. 3, pp. 546–552, Mar. 1992.

[103] D. S. Ong, K. F. Li, G. J. Rees, J. P. R. David, and P. N. Robson, "A simple model to determine multiplication and noise in avalanche photodiodes," *J. Appl. Phys.*, vol. 83, no. 6, pp. 3426–3428, Mar. 1998.

[104] T. Kunikiyo, M. Takenaka, Y. Kamakura, M. Yamaji, H. Mizuno, M. Morifuji, K. Taniguchi, and C. Hamaguchi, "A monte carlo simulation of anisotropic electron transport in silicon including full band structure and anisotropic impact-ionization model," *J. Appl. Phys.*, vol. 75, no. 1, pp. 297–312, Jan. 1994.

[105] J. Požela and A. Reklaitis, "Electron transport properties in GaAs at high electric fields," *Solid State Electron.*, vol. 23, no. 9, pp. 927–933, Sep. 1980.

[106] J. C. Phillips, "Energy-band interpolation scheme based on a pseudopotential," *Phys. Rev.*, vol. 112, no. 3, pp. 685–695, Nov. 1958.

[107] D. Dolgos, H. Meier, A. Schenk, and B. Witzigmann, "Full-band monte carlo simulation of high-energy carrier transport in single photon avalanche diodes: Computation of breakdown probability, time to avalanche breakdown, and jitter," *J. Appl. Phys.*, vol. 110, no. 8, p. 084507, Oct. 2011.

[108] E. Pop, R. W. Dutton, and K. E. Goodson, "Analytic band monte carlo model for electron transport in si including acoustic and optical phonon dispersion," *J. Appl. Phys.*, vol. 96, no. 9, pp. 4998–5005, Nov. 2004.

[109] H. Arabshahi, M. R. Khalvati, and M. R. Rokn-Abadi, "Temperature and doping dependencies of electron mobility in InAs, AlAs and AlGaAs at high electric field application," *Braz. J. Phys.*, vol. 38, no. 3a, Sep. 2008.

[110] H. Arabshahi and F. Taghavi, "Calculation of high field electron transport properties in GaSb and GaAS using a monte carlo method," *Res. J. Appl. Sci.*, vol. 6, no. 3, pp. 213–217, Mar. 2011.

[111] S. Chan and K. Choo, "Monte carlo modeling of carrier transport and impact ionization in GaSb," in *2011 2nd International Conference on Photonics.* IEEE, Oct. 2011.

[112] L. Messias and E. M. Jr., "Hole transport characteristics in pure and doped GaSb," *Braz. J. Phys.*, vol. 32, no. 2a, pp. 402–404, Jun. 2002.

[113] F. M. A. El-Ela, "Monte carlo simulation of electron transport in AlGaAs." American Institute of Physics, 2005.

[114] I. C. Sandall, J. S. Ng, S. Xie, P. J. Ker, and C. H. Tan, "Temperature dependence of impact ionization in InAs," *Opt. Express*, vol. 21, no. 7, p. 8630, Apr. 2013.

[115] S. Plimmer, J. David, D. Ong, and K. Li, "A simple model for avalanche multiplication including deadspace effects," *IEEE Trans. Electron Devices*, vol. 46, no. 4, pp. 769–775, Apr. 1999.

[116] C. H. Tan, R. Ghin, J. P. R. David, G. J. Rees, and M. Hopkinson, "The effect of dead space on gain and excess noise in $In_{0.48}Ga_{0.52}P$ $P^{+}IN^{+}$ diodes," *Semicond. Sci. Tech.*, vol. 18, no. 8, pp. 803–806, Jul. 2003.

[117] S. C. L. T. Mun, C. H. Tan, Y. L. Goh, A. R. J. Marshall, and J. P. R. David, "Modeling of avalanche multiplication and excess noise factor in $In_{0.52}Al_{0.48}As$ avalanche photodiodes using a simple monte carlo model," *J. Appl. Phys.*, vol. 104, no. 1, p. 013114, Jul. 2008.

[118] J. D. Petticrew, S. J. Dimler, C. H. Tan, and J. S. Ng, "Modeling temperature dependent avalanche characteristics of InP," *J. Light. Technol.*, vol. PP, pp. 1–1, Oct. 2019.

[119] C. Groves, C. N. Harrison, J. P. R. David, and G. J. Rees, "Temperature dependence of breakdown voltage in $Al_{x}Ga_{1-x}As$," *J. Appl. Phys.*, vol. 96, no. 9, pp. 5017–5019, Nov. 2004.

[120] X. Zhou, J. S. Ng, and C. H. Tan, "A simple monte carlo model for prediction of avalanche multiplication process in silicon," *J. Instrum.*, vol. 7, no. 08, p. 08006, Aug. 2012.

[121] L. V. Keldysh, "Kinetic theory of impact ionization in semiconductors," *Sov. Phys. JETP*, vol. 37, no. 3, p. 509, Mar. 1959.

[122] B. K. Ng, "Impact ionization in wide band gap semiconductors: $Al_{x}Ga_{1-x}As$ and 4H-SiC," Ph.D. dissertation, University of Sheffield, 2002.

[123] 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 J. Sel. Top. Quantum Electron.*, vol. 13, no. 4, pp. 919–925, 2007.

[124] S. J. Dimler, "A capacitive quenching characterisation system for single photon avalanche diodes and avalanche photodiodes," Ph.D. dissertation, University of Sheffield, 2015.

[125] B. Levine, C. Bethea, and J. Campbell, "Near room temperature 1.3 µm single photon counting with a InGaAs avalanche photodiode," *Electron. Lett.*, vol. 20, no. 14, p. 596, 1984.

[126] B. F. Levine and C. G. Bethea, "Single photon detection at 1.3 µm using a gated avalanche photodiode," *Appl. Phys. Lett.*, vol. 44, no. 5, pp. 553–555, Mar. 1984.

[127] B. F. Levine and C. G. Bethea, "10-MHz single photon counting at 1.3 µm," *Appl. Phys. Lett.*, vol. 44, no. 6, pp. 581–582, Mar. 1984.

[128] B. F. Levine and C. G. Bethea, "Error rate measurement for single photon detection at 1.3 µm," *Applied Physics Letters*, vol. 44, no. 7, pp. 649–650, Apr. 1984.

[129] B. Levine and C. Bethea, "Detection of single 1.3 µm photons at 45 mbit/s," *Electron. Lett.*, vol. 20, no. 6, p. 269, 1984.

[130] Y. Kang, H. X. Lu, Y.-H. Lo, D. S. Bethune, and W. P. Risk, "Dark count probability and quantum efficiency of avalanche photodiodes for single-photon detection," *Appl. Phys. Lett.*, vol. 83, no. 14, pp. 2955–2957, Oct. 2003.

[131] W. Kern and D. Puotinen, "Cleaning solutions based on hydrogen peroxide for use in silicon semiconductor technology," *RCA Rev.*, vol. 31, pp. 187–205, 1970.

[132] R. Ridley, T. Grebs, J. Trost, R. Webb, M. Schuler, R. Longenberger, T. Fenstemacher, and M. Caravaggio, "Advanced aqueous wafer cleaning in power semiconductor device manufacturing," in *IEEE/SEMI 1998 IEEE/SEMI Advanced Semiconductor Manufacturing Conference and Workshop (Cat. No.98CH36168).* IEEE, 1998.

[133] G. W. Gale, B. K. Kirkpatrick, and F. W. Kern, "Surface preparation," in *Handbook of Semiconductor Manufacturing Technology*, 2nd ed., R. Doering and Y. Nishi, Eds. Boca Raton, FL: CRC Press, 2008, ch. 5.

[134] Dow. (2014, Mar.) Megaposit^{TM} spr^{TM}220 series i-line photoresists. [Online]. Available: http://microchem.com/PDFs_Dow/SPR%20220%20DATA%20SHEET%20R%26H.pdf [Accessed: 2019-08-01]

[135] Rohm and Haas Electronic Materials. (2007, May) Megaposit^{TM} spr^{TM}350 series photoresist. [Online]. Available: http://micromaterialstech.com/wp-content/dow_electronic_materials/datasheets/SPR350_Photoresist.pdf [Accessed: 2019-08-01]

[136] R. G. Narechania, "Polyimide adhesion," *J. Electrochem. Soc.*, vol. 132, no. 11, p. 2700, 1985.

[137] J. P. Biersack and L. G. Haggmark, "A monte carlo computer program for the transport of energetic ions in amorphous targets," *Nucl. Instrum. Methods*, vol. 174, no. 1-2, pp. 257–269, Aug. 1980.

[138] Synopsys. (2012) Sentaurus tcad. [Online]. Available: https://www.synopsys.com/content/dam/synopsys/silicon/datasheets/sentaurus_ds.pdf [Accessed: 2019-08-01]

[139] N. Gherardi, S. Martin, and F. Massines, "A new approach to SiO_2 deposit using a N_2-SiH_4-N_2O glow dielectric barrier-controlled discharge at atmospheric pressure," *J. Phys. D*, vol. 33, no. 19, pp. L104–L108, Sep. 2000.

[140] S. A. Rosli, A. A. Aziz, and H. A. Hamid, "Characteristics of RIE SF6/O2/Ar plasmas on n-silicon etching," in *2006 IEEE International Conference on Semiconductor Electronics.* IEEE, Nov. 2006.

[141] DOW, "Microposit™ remover 1165 for post-etch applications," no. 889-00026,Rev.0, Feb. 2014. [Online]. Available: http://microchem.com/products/images/uploads/Remover-1165-DataSheet-RH.pdf [Accessed: 2019-08-05]

[142] J. D. Petticrew, S. J. Dimler, and J. S. Ng, "Simple monte carlo simulator for modelling linear mode and geiger mode avalanche photodiodes in C++," *J. Open Res. Softw.*, vol. 6, 2018.

[143] C. H. Tan, J. S. Ng, G. J. Rees, and J. P. R. David, “Statistics of avalanche current buildup time in single-photon avalanche diodes,” *IEEE J. Sel. Top. Quantum Electron.*, vol. 13, no. 4, pp. 906–910, 2007.

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

[145] P. J. Hambleton, S. A. Plimmer, and G. J. Rees, “Limitations of the saturated drift velocity approximation for time domain modelling,” *Semicond. Sci. Tech.*, vol. 17, no. 2, pp. 124–128, Jan. 2002.

[146] J. Y. Tang and K. Hess, “Impact ionization of electrons in silicon (steady state),” *J. Appl. Phys.*, vol. 54, no. 9, pp. 5139–5144, Sep. 1983.

[147] M. V. Fischetti and S. E. Laux, “Monte carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects,” *Phys. Rev. B*, vol. 38, no. 14, pp. 9721–9745, Nov. 1988.

[148] S. Ramo, “Currents induced by electron motion,” *Proc. IRE*, vol. 27, no. 9, pp. 584–585, Sep. 1939.

[149] S. Cova, M. Ghioni, and F. Zappa, “Circuit for high precision detection of the time of arrival of photons falling on single photon avalanche diodes,” US Patent 6384663 B2, 2002.

[150] C. Canali, G. Ottaviani, and A. A. Quaranta, “Drift velocity of electrons and holes and associated anisotropic effects in silicon,” *J. Phys. Chem. Solids*, vol. 32, no. 8, pp. 1707–1720, Jan. 1971.

[151] A. Prior, “The field-dependence of carrier mobility in silicon and germanium,” *J. Phys. Chem. Solids*, vol. 12, no. 2, pp. 175–180, Jan. 1960.

[152] C. Norris and J. Gibbons, “Measurement of high-field carrier drift velocities in silicon by a time-of-flight technique,” *IEEE Trans. Electron Devices*, vol. 14, no. 1, pp. 38–43, Jan. 1967.

[153] C. Canali, C. Jacoboni, G. Ottaviani, and A. Alberigi-Quaranta, “High-field diffusion of electrons in silicon,” *Appl. Phys. Lett.*, vol. 27, no. 5, pp. 278–280, Sep. 1975.

[154] F. Nava, C. Canali, L. Reggiani, D. Gasquet, J. C. Vaissiere, and J. P. Nougier, “On the diffusivity of holes in silicon,” *J. Appl. Phys.*, vol. 50, no. 2, pp. 922–924, Feb. 1979.

[155] D. Massey, J. David, and G. Rees, “Temperature dependence of impact ionization in submicrometer silicon devices,” *IEEE Trans. Electron Devices*, vol. 53, no. 9, pp. 2328–2334, Sep. 2006.

[156] D. Massey, J. David, and G. Rees, “Temperature dependence of avalanche multiplication in submicron silicon devices,” in *Proceedings of 35th European Solid-State Device Research Conference, 2005. ESSDERC 2005.* IEEE.

[157] R. V. Overstraeten and H. D. Man, “Measurement of the ionization rates in diffused silicon p-n junctions,” *Solid State Electron.*, vol. 13, no. 5, pp. 583–608, May 1970.

[158] C. H. Tan, “Measurements of excess avalanche noise in sub-micron Si and $Al_{0.6}Ga_{0.4}As$ avalanche photodiodes,” Ph.D. dissertation, University of Sheffield, 2001.

[159] P. Hambleton, B. Ng, S. Plimmer, J. David, and G. Rees, “The effects of nonlocal impact ionization on the speed of avalanche photodiodes,” *IEEE Trans. Electron Devices*, vol. 50, no. 2, pp. 347–351, Feb. 2003.

[160] E. A. G. Webster, L. A. Grant, and R. K. Henderson, “A high-performance single-photon avalanche diode in 130-nm CMOS imaging technology,” *IEEE Electron Device Lett.*, vol. 33, no. 11, pp. 1589–1591, Nov. 2012.

[161] L. Commandar, B. Fröhlich, J. Dynes, A. Sharpe, M. Lucamarini, Z. Yuan, R. Penty, and A. Shields, “Gigahertz-gated InGaAs/InP single-photon detector with detection efficiency exceeding 55% at 1550nm,” *J. Appl. Phys.*, vol. 117, no. 8, p. 083109, Feb. 2015.

[162] J. S. Ng, C. H. Tan, G. J. Rees, and J. P. R. David, “Effects of dead space on breakdown probability in geiger mode avalanche photodiode,” *J. Mod. Opt.*, vol. 54, no. 2-3, pp. 353–360, Jan. 2007.

[163] S. C. L. T. Mun, C. H. Tan, S. J. Dimler, L. J. J. Tan, J. S. Ng, Y. L. Goh, and J. P. R. David, “A theoretical comparison of the breakdown behavior of $In_{0.52}Al_{0.48}As$ and InP near-infrared single-photon avalanche photodiodes,” *IEEE J. Quantum Electron.*, vol. 45, no. 5, pp. 566–571, May 2009.

[164] Y. Takanashi and Y. Horikoshi, “Temperature dependence of ionization coefficients for InP and 1.3 µm InGaAsP avalanche photodiodes,” *Jpn. J. Appl. Phys.*, vol. 20, no. 10, pp. 1907–1913, Oct. 1981.

[165] F. Osaka, Y. Kishi, M. Kobayashi, and T. Mikawa, “Low-temperature impact ionization rates in (111) oriented InP,” *Appl. Phys. Lett.*, vol. 47, no. 8, pp. 865–866, Oct. 1985.

[166] C. Tan, K. Li, S. Plimmer, J. David, G. Rees, J. Clark, and C. Button, “Improved excess noise and temperature dependence of multiplication characteristics in thin InP avalanching regions,” in *Conf. Proc. Indium Phosph.* IEEE, 1999.

[167] L. W. Cook, G. E. Bulman, and G. E. Stillman, “Electron and hole impact ionization coefficients in InP determined by photomultiplication measurements,” *Appl. Phys. Lett.*, vol. 40, no. 7, pp. 589–591, Apr. 1982.

[168] M. Saleh, M. Hayat, P. Sotirelis, A. Holmes, J. Campbell, B. Saleh, and M. Teich, “Impact-ionization and noise characteristics of thin III-V avalanche photodiodes,” *IEEE Trans. Electron Devices*, vol. 48, no. 12, pp. 2722–2731, 2001.

[169] L. J. J. Tan, J. S. Ng, C. H. Tan, and J. P. R. David, “Avalanche noise characteristics in submicron InP diodes,” *IEEE J. Quantum Electron.*, vol. 44, no. 4, pp. 378–382, Apr. 2008.

[170] S. McKenzie and M. G. Burt, “A test of the lucky-drift theory of the impact ionisation coefficient using monte carlo simulation,” *J. Phys. C*, vol. 19, no. 12, pp. 1959–1973, Apr. 1986.

[171] D. Lockwood, G. Yu, and N. Rowell, “Optical phonon frequencies and damping in AlAs, GaP, GaAs, InP, InAs and InSb studied by oblique incidence infrared spectroscopy,” *Solid State Commum.*, vol. 136, no. 7, pp. 404–409, Nov. 2005.

[172] L. Kleinman and J. C. Phillips, “Crystal potential and energy bands of semiconductors. I. self-consistent calculations for diamond,” *Phys. Rev.*, vol. 116, no. 4, pp. 880–884, Nov. 1959.

[173] D. J. Chadi and M. L. Cohen, “Special points in the brillouin zone,” *Phys. Rev. B*, vol. 8, no. 12, pp. 5747–5753, Dec. 1973.

[174] T. H. Windhorn, L. W. Cook, M. A. Haase, and G. E. Stillman, “Electron transport in InP at high electric fields,” *Appl. Phys. Lett.*, vol. 42, no. 8, pp. 725–727, Apr. 1983.

[175] K. Brennan and K. Hess, “Theory of high-field transport of holes in GaAs and InP,” *Phys. Rev. B*, vol. 29, no. 10, pp. 5581–5590, May 1984.

[176] L. J. J. Tan, "Telecommunication wavelength InP based avalanche photodiodes," Ph.D. dissertation, University of Sheffield, 2008.

[177] J. S. Ng, C. H. Tan, J. P. R. David, and G. J. Rees, "A general method for estimating the duration of avalanche multiplication," *Semicond. Sci. Tech.*, vol. 17, no. 10, pp. 1067–1071, Sep. 2002.

Appendix A

The Simple Monte Carlo Simulator

The source code for the Simple Monte Carlo Simulator was released under an Apache 2.0 licence as part of this work. The source code (implemented in C++) and documentation for the Simple Monte Carlo Simulator can be found on GitHub (jdpetticrew/Simple-Monte-Carlo-Simulator) and on the University of Sheffield's Online Research Data (DOI: 10.15131/shef.data.5683939). The simulator has the capabilities to calculate the following device characteristics for any given SMC material parameter set:

- Avalanche gain
- Excess noise factors
- Breakdown probability
- Mean time to breakdown
- Timing jitter

alongside the electric field dependent material properties:

- Electron and Hole Drift Velocity
- Electron and Hole Impact Ionisation Coefficients.

Currently the simulator can be used for Si [120], InGaP [116], InP [118] and GaAs [115]. The various characteristics can be obtained using the three distinct operating modes.

The three modes, which have been implemented as separate functions, are namely device properties, drift velocity, and impact ionisation (ii_coef).

A.1 Architecture

The simulator has been implemented with 5 classes, to achieve the three operating modes. The class hierarchy of these functions can be seen in Figure A.1, whilst a brief description of each class can be seen in Table A.1. The 1D Gaussian Histogram fitting class source code was released separately to this work under an Apache 2.0 Licence and can be found on GitHub (jdpetticrew/Gaussian-Histogram-Fitter).

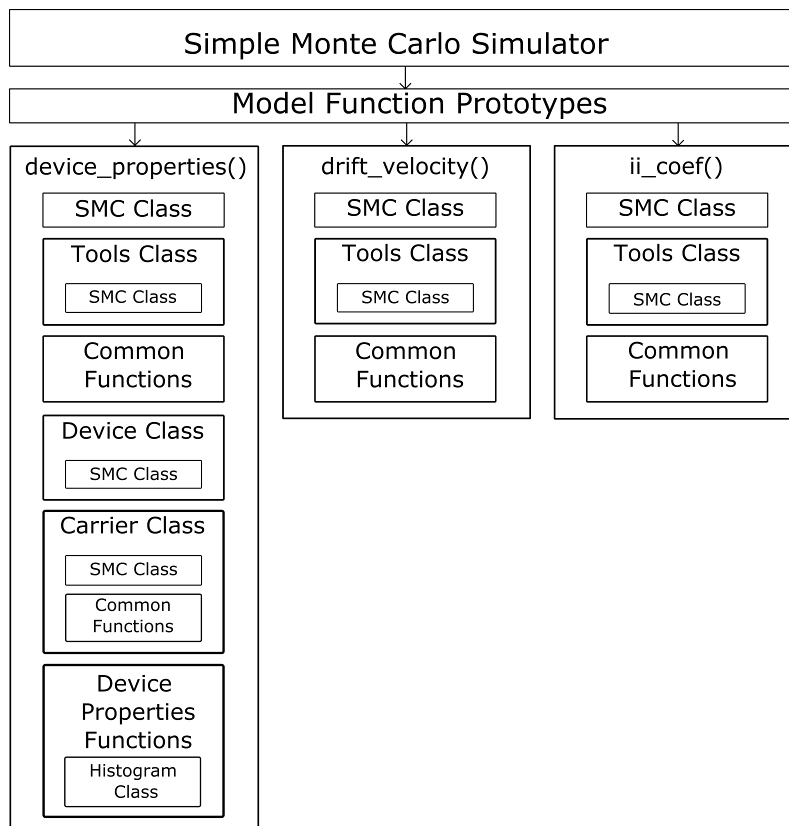

Figure A.1: Schematic of class dependencies of the Simple Monte Carlo Simulator

Table A.1: Simple Monte Carlo Simulator Class Descriptions

Class	Description
SMC	Contains the material parameter sets
Tools	Calculates and stores the electron and hole interaction probabilities and scattering rates
Device	Uses the supplied doping profile to calculate the electric field profile for any given voltage (see appendix B)
Carrier	Stores all the tracked variables for each electron and hole in the simulation. It is also used to pick the carriers random scattering directions.
Histogram	Calculates the mean and standard deviation of the data sets and can fit data sets to a 1D Gaussian.

Appendix B

Electric Field Solver

Electric field profiles govern the operating conditions of all APD and SPAD devices. It is possible to use the Poisson equation Equation (B.1) to calculate a 1-dimensional electric field profile for any device for any applied bias when the doping profile comprising of doping concentrations and region thicknesses is known.

$$\frac{dE}{dx} = \frac{qXN}{\epsilon} \tag{B.1}$$

B.1 Calculating Simple Electric Fields

B.1.1 2 Regions

The simplest doping profile resulting in an electric field is a P-N diode, i.e. Figure B.1. From Figure B.1 it is possible to see that

$$E = \frac{qN_1X_1}{\epsilon_1} = \frac{qN_2X_2}{\epsilon_2}.$$

The total voltage, V_t, can be calculated as the area under the electric field profile,

$$V_t = \frac{EX_1}{2} + \frac{EX_2}{2}.$$

On the condition that the P and N regions are made of the same material, $\epsilon_1 = \epsilon_2$, the equation for V_t can be rearranged to

$$V_t = \frac{qN_2^2X_2}{2\epsilon_2N_1}$$

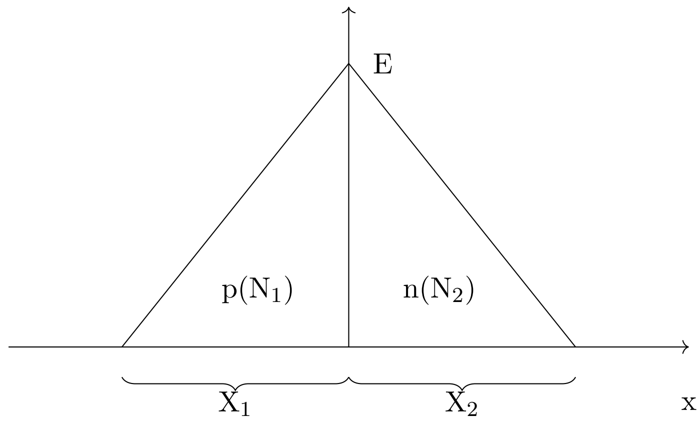

Figure B.1: Schematic of PN diode

which leads to the applied voltage dependent solutions

$$X_2 = \frac{2V_t\epsilon_2 N_1}{qN_2^2} \quad , \quad X_1 = \frac{2V_t\epsilon_1}{qN_2} \quad \text{and,} \quad E = \frac{2V_t N_1}{qN_2}$$

B.1.2 3 Regions

The 2 region case will be true until one of the layers has become fully depleted, which is when the 3 region case takes over. An example would be a P-I-N diode where, if the i-region was p-type as in Figure B.2, the i-region would have to be fully depleted before the p-region would begin to deplete. An example of the 3 region case is given below, where solutions need to be found for the p-region and n-region depletion thicknesses as the i-region is fully depleted.

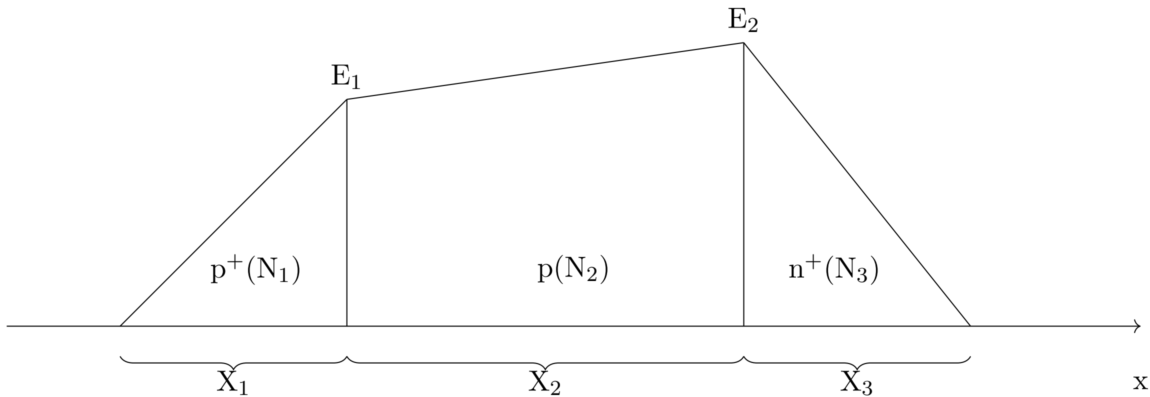

Figure B.2: Schematic of PIN diode

$$E_1 = \frac{qN_1X_1}{\epsilon_1}, E_2 = E_1 + \frac{qN_2X_2}{\epsilon_2}, E_2 = \frac{qN_3X_3}{\epsilon_3}$$

Assuming the regions are all made of the same material so $\epsilon_1 = \epsilon_2 = \epsilon_3 = \epsilon$, rearanging leads to,

$$X_3 = \frac{N_1 X_1 + N_2 X_2}{N_3}.$$

The total voltage is given by,

$$V_t = \frac{1}{2}\left(E_1 X_1 + E_2 X_3 + (E_1 + E_2) X_2\right)$$

which becomes the following when substituting in the equations for E_1, E_2, and X_3,

$$V_t = \frac{q}{2\epsilon}\left(N_1\left(1 + \frac{N1}{N3}\right)X_1^2 + 2N_1\left(X_2 + \frac{N_2 X_2}{N_3}\right)X_1 + \left(\frac{N_2^2 X_2^2}{N_3} + N_2 X_2^2\right)\right).$$

This solution of V_t leads itself to a quadratic equation solution for X_1 using the quadratic formula.

$$X_1 = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},$$

where,

$$a = N_1\left(1 + \frac{N_1}{N_3}\right),$$

$$b = 2N_1\left(X_2 + \frac{N_2 X_2}{N_3}\right),$$

$$c = \frac{N_2^2 X_2^2}{N_3} + N_2 X_2^2 - \frac{2V_t\epsilon}{q}.$$

B.2 Infinite Electric Field Solver

For devices with many doping layers, or even a graded doping profile, this simple 1D solution can become quite complicated, having to determine how many regions have been depleted to then work out which case to apply then calculating all of the region widths. This simply isn't a practical solution for many region devices, it is quicker to write a generic infinite region solver than to write a special case for each number of regions.

This approach for a simple infinite electric field solver is limited by the assumption that the device contains only a single P-N junction and that all P-type regions on one side of the junction and that all N-type regions are on the other side of the junction. This assumption avoids having to deal with multiple PN junctions simultaneously.

This method can be broken down into three key steps outlined in the flow chart shown in Figure B.3.

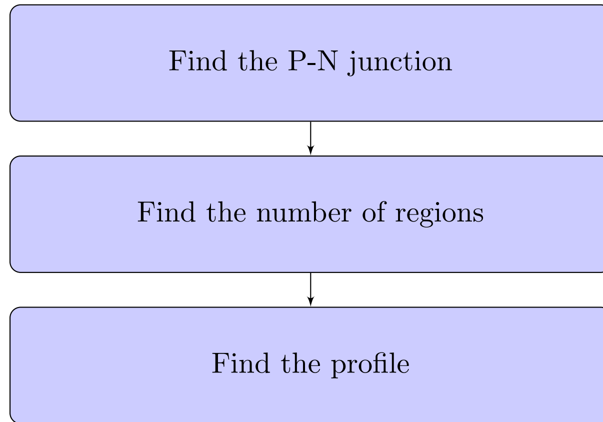

Figure B.3: Flow chart outlining simplified process for the infinite electric field solver

B.2.1 PN junction finder

The first step, to find the P-N junction, can be achieved through a comparison method. The method deployed was to iteratively multiply the inputted doping concentration of neighbouring regions (where positive concentrations represent P-type and negative concentrations represent N-type), if the product was positive then the dopants are of the same type, whilst if the product is negative the the PN junction has been found as the dopant concentration has swapped.

B.2.2 Find the number of regions

Now the PN junction has been found, the second step is to identify the number of regions depleted by the applied bias. Firstly the bias required to completely deplete the first P region, V_p, is calculated.

For the first region, the running total of the electric field E_t is given by,

$$E_t = \frac{qN_1X_1}{\epsilon_1}$$

and the running total applied bias V_{pt} is given by

$$V_{pt} = \frac{E_tX_1}{2}.$$

This is continued for subsequent regions using

$$E_t = E_{tprevious} + \frac{qN_iX_i}{\epsilon_i} \quad \text{and} \quad V_{pt} = V_{ptprevious} + min(E_t, E_{tprevious})X_i + Xi\frac{|E_t - E_{tprevious}|}{2},$$

until E_t becomes negative, at that point the region does not need to be fully depleted and the final value to be added to V_{pt} is

$$V_{pt} = \frac{E_{tprevious}X_i}{2}.$$

The value of V_{pt} can then be compared to the applied bias required for the simulation. If V_{pt} > required bias then the number of regions required has identified, otherwise the p-region needs to be "stepped back" to the previous p-region and repeat the calculation until this condition has been met, potentially "stepping back multiple p-regions" depending on the doping structure of the device.

B.2.3 Finding the profile

Finally the actual electric field profile for the requested voltage can be found, via an iterative method similar to the previous step. The voltage required to fully deplete the p-region is already known, V_2, and the voltage required to deplete the previous p-region, V_1. The condition

$$V_1 < \text{required voltage} < V_2$$

is known. Next V_3, the voltage required to deplete from half-way through the p-region must be calculated. Then if $V_3 < required\ voltage$ the voltage required to deplete from half-way between the V_2 and V_3 positions is calculated otherwise the position halfway between the V_1 and V_3 points is calculated. This half-way cutting of the positions is repeated quickly to converge on the starting position, and thus the electric field for the given applied bias.

Appendix C

Obtaining Diffusion constant from $S(x,t)$

To begin, consider the particle current flux, j, in an electric field,

$$j = S\nu - D\frac{dS}{dx}, \tag{C.1}$$

where S is the number of carriers. The first term, $S\nu$, represents the contribution from the drift current, whilst the second contribution , $-D\frac{dS}{dx}$, is from the diffusion current. The continuity equation,

$$\frac{\delta j}{\delta x} = -\frac{\delta S}{\delta t} \tag{C.2}$$

also needs to be considered, which shows the current changes with position. Combining Equations (C.1) and (C.2) yields,

$$\nu\frac{\delta S}{\delta x} - D\frac{\delta^2 S}{\delta x^2} = -\frac{\delta S}{\delta t}. \tag{C.3}$$

Changing the frame of reference of Equation (C.3) from the initial injection position of the carriers, $x = 0$, to be moving with the carriers so that $y = x - \nu t$, Equation (C.3) becomes,

$$D\frac{\delta^2 S}{\delta y^2} = \frac{\delta S}{\delta t}. \tag{C.4}$$

Assuming the solution to Equation (C.4) follows a normal gaussian, $S(y,t)$ can be repre-

sented as,

$$S(y,t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-y^2}{2\sigma^2}\right), \tag{C.5}$$

where σ takes the form,

$$\sigma = AD^B t^C, \tag{C.6}$$

where A, B, and C are unknown constants that need to be derived using Equations (C.4) and (C.5). To start Equation (C.5) is differentiated with regards to y, resulting in,

$$\frac{\partial S}{\partial y} = -\frac{y}{\sigma^2} S, \text{ and }, \frac{\partial^2 S}{\partial y^2} = \frac{S}{\sigma^2}\left(\frac{y^2}{\sigma^2} - 1\right), \tag{C.7}$$

and with respect to t, resulting in,

$$\frac{\partial S}{\partial t} = \frac{SC}{t}\left(\frac{y^2}{\sigma^2} - 1\right). \tag{C.8}$$

Combining Equations (C.4), (C.7) and (C.8) results in,

$$\frac{SD}{\sigma^2}\left(\frac{y^2}{\sigma^2} - 1\right) = \frac{SC}{t}\left(\frac{y^2}{\sigma^2} - 1\right), \tag{C.9}$$

which can be simplified to,

$$\frac{D}{A^2 D^{2B} t^{2C}} = \frac{C}{t}. \tag{C.10}$$

To cancel the D and t terms in Equation (C.10), both B and C must equal 0.5, meaning that A takes the value of $\sqrt{2}$. Substituting these values back into Equation (C.5) results in the solution,

$$S(y,t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{y^2}{4Dt}\right), \tag{C.11}$$

which when converted back into the original reference frame is,

$$S(x,t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{(x-vt)^2}{4Dt}\right). \tag{C.12}$$

Appendix D

Fitting InP PDFs

Two of the probability density function fits for the InP temperature dependent SMC (chapter 5), were shown in Figure 5.6. This appendices shows the complete fitting graphs, where probability densities were generated in 50 kV.cm^{-1} intervals from 400 kV.cm^{-1} to 800 kV.cm^{-1} for electrons and holes at 150, 200, 250 and 290 K. All the probability densities shown in this chapter were fitted using a hard deadspace approximation [177], using Equation (2.18) for electrons and Equation (2.19) for holes. From this fitting process we were able to obtain equations for $\alpha^*(E)$ and $\beta^*(E)$ at each temperature, these final equations can be seen in Table 5.3.

An example of the fitted probability densities plotted on a linear scale can be seen in Figure D.1 as it is a better visual for the effect of deadspace, however as our deadspace values were previously set by the use of previously published impact ionization threshold energies the remaining Figures D.2 to D.9 are all plotted with a lograithmic y-axis as that shows a better indication of the α^* or β^* fit.

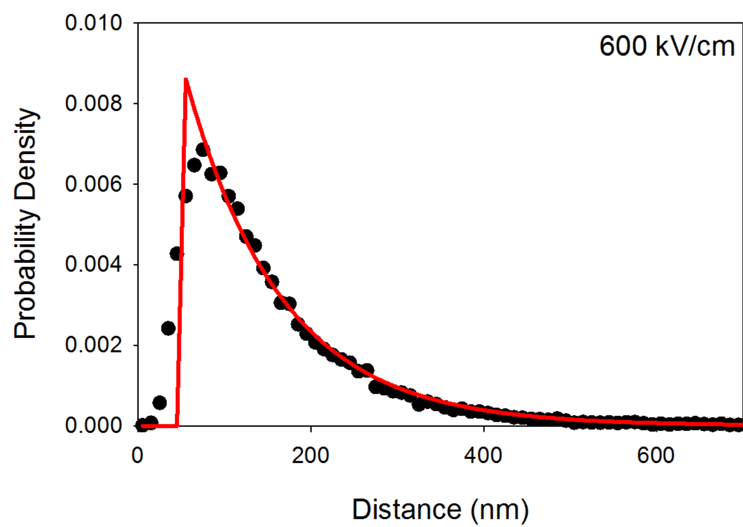

Figure D.1: Example of probability density plot fitted using a linear y-axis (290 K holes at 600 kV.cm^{-1})

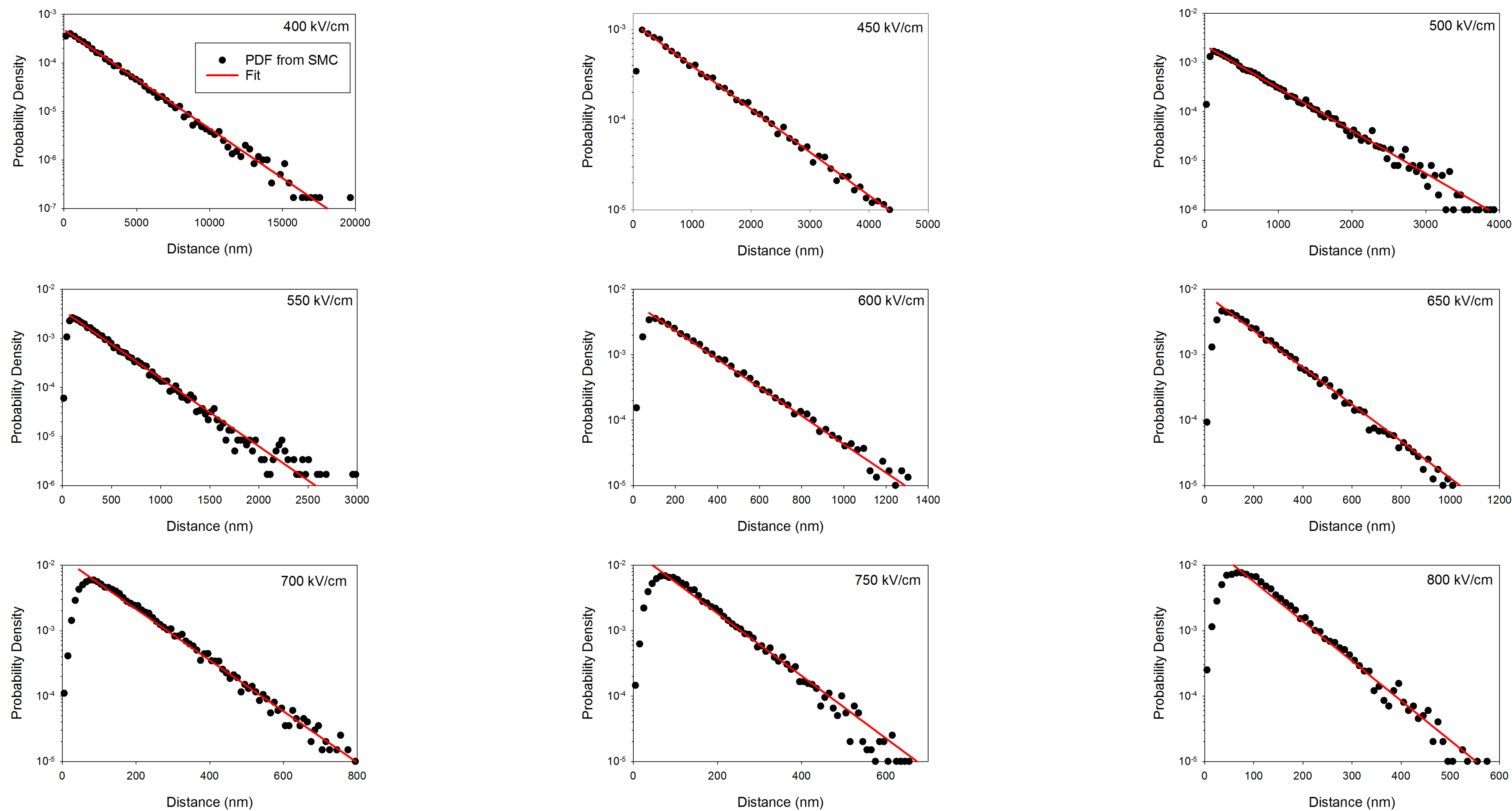

Figure D.2: Fitted probability densities generated from the SMC for electrons at 290 K, fitted using the hard deadspace PDF equation.

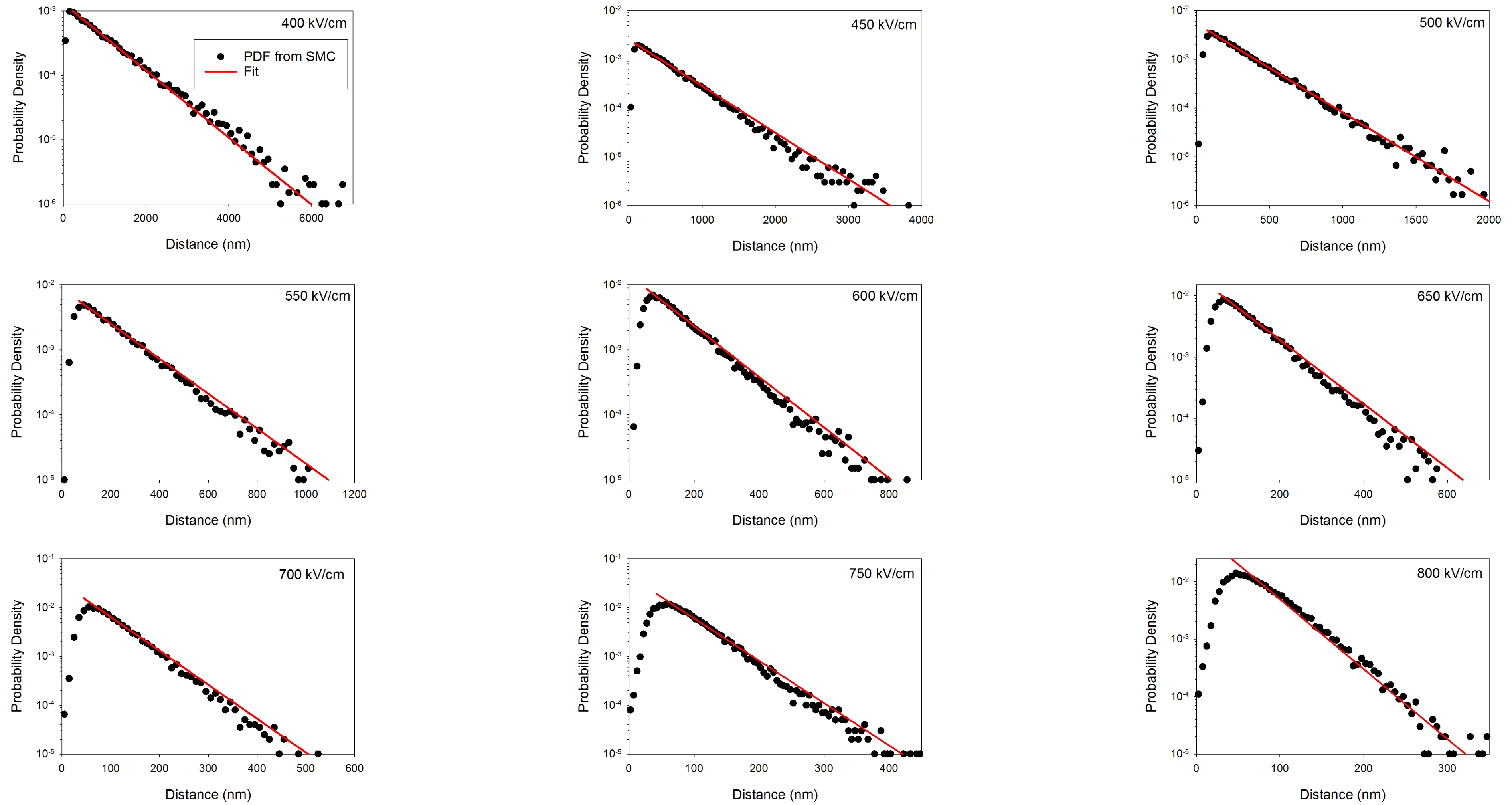

Figure D.3: Fitted probability densities generated from the SMC for holes at 290 K, fitted using the hard deadspace PDF equation.

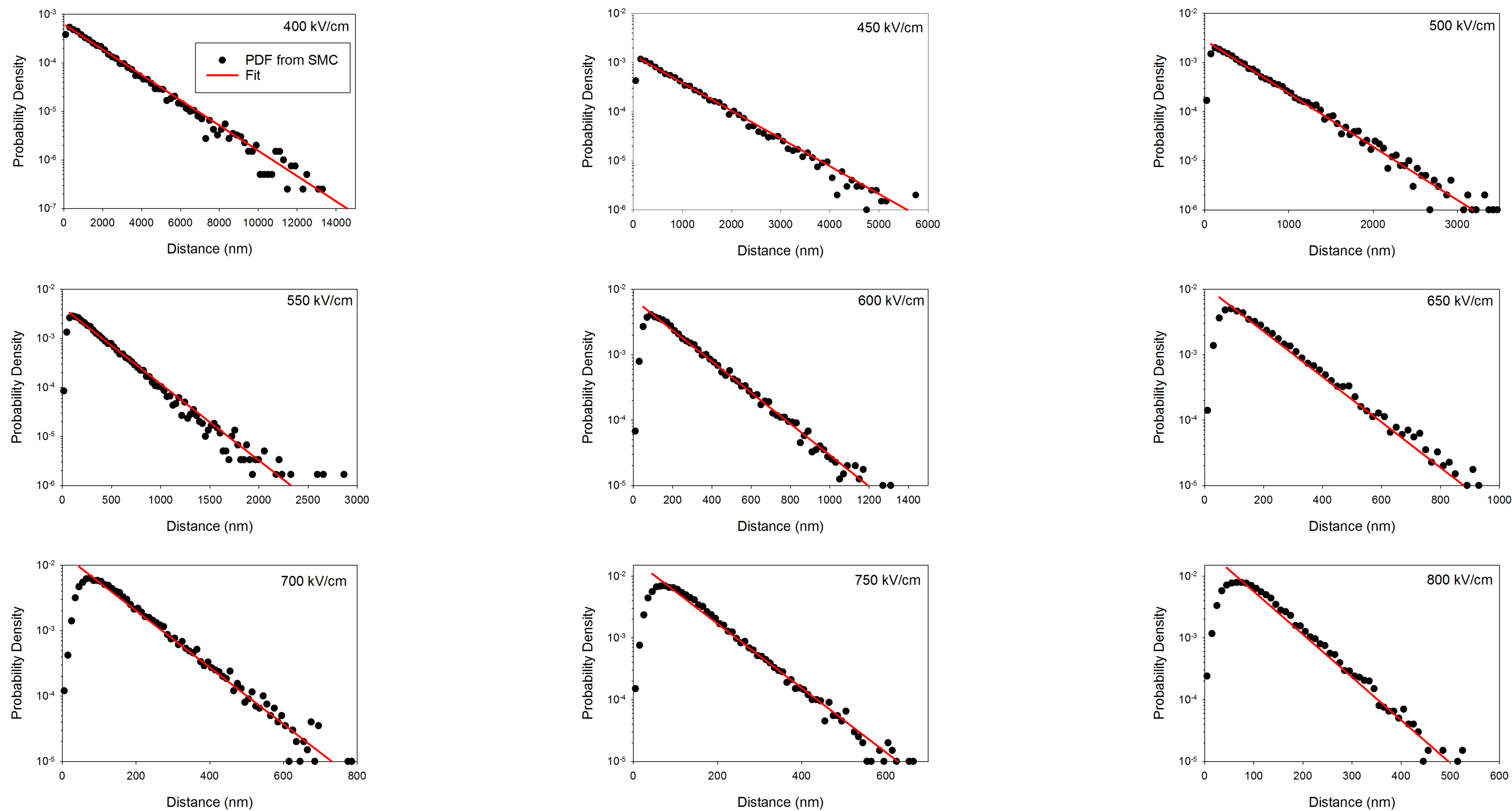

Figure D.4: Fitted probability densities generated from the SMC for electrons at 250 K, fitted using the hard deadspace PDF equation.

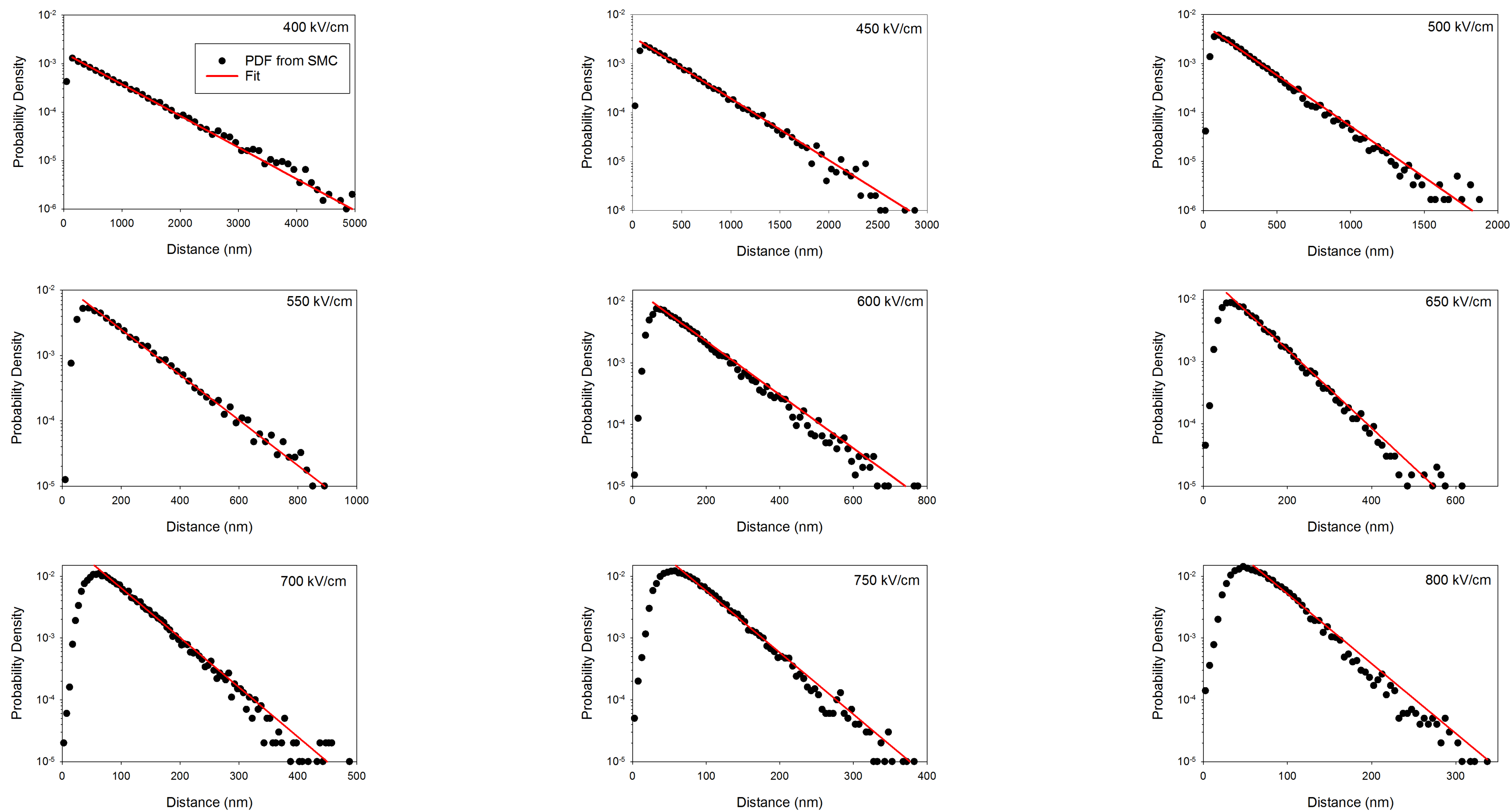

Figure D.5: Fitted probability densities generated from the SMC for holes at 250 K, fitted using the hard deadspace PDF equation.

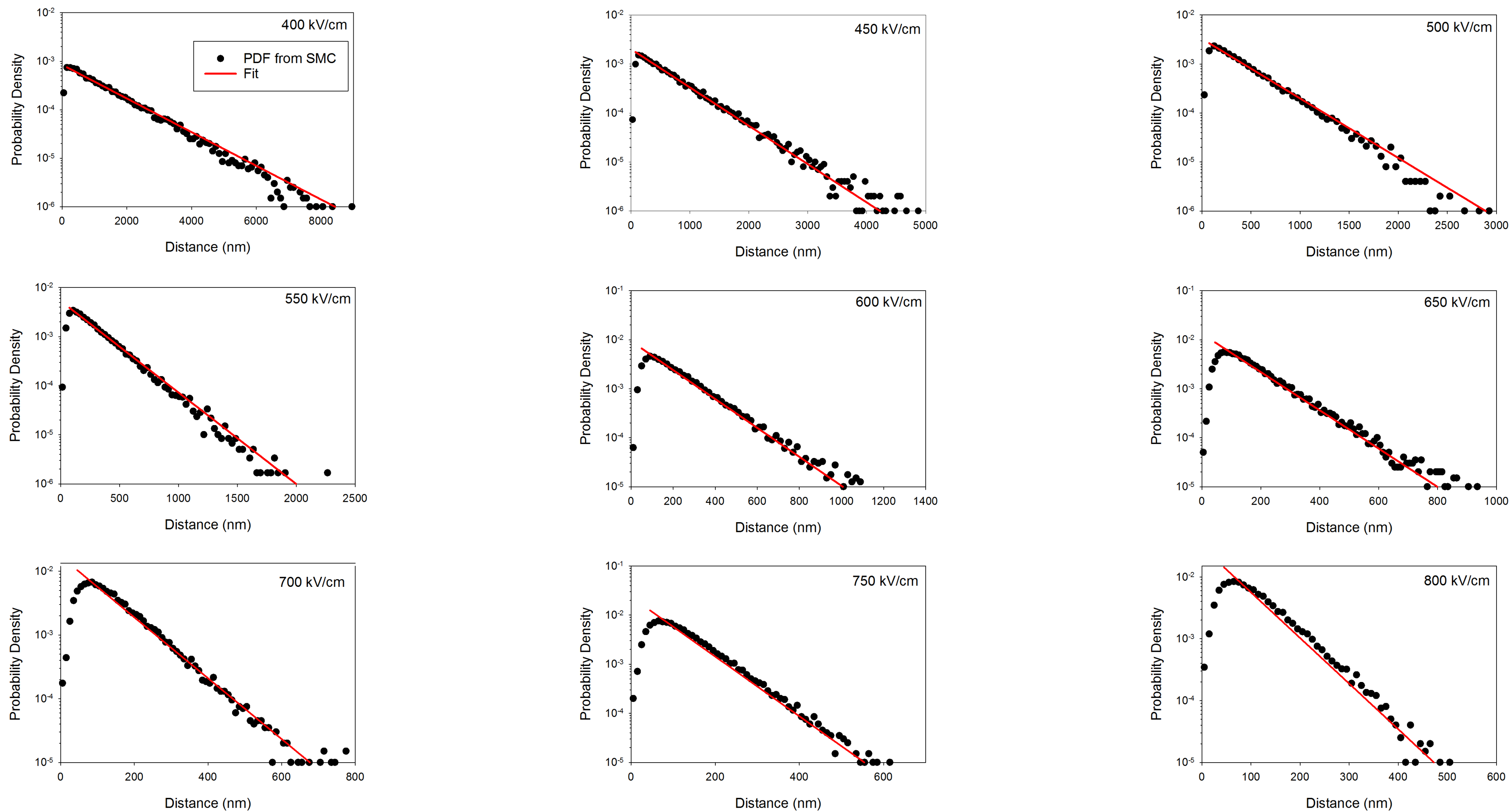

Figure D.6: Fitted probability densities generated from the SMC for electrons at 200 K, fitted using the hard deadspace PDF equation.

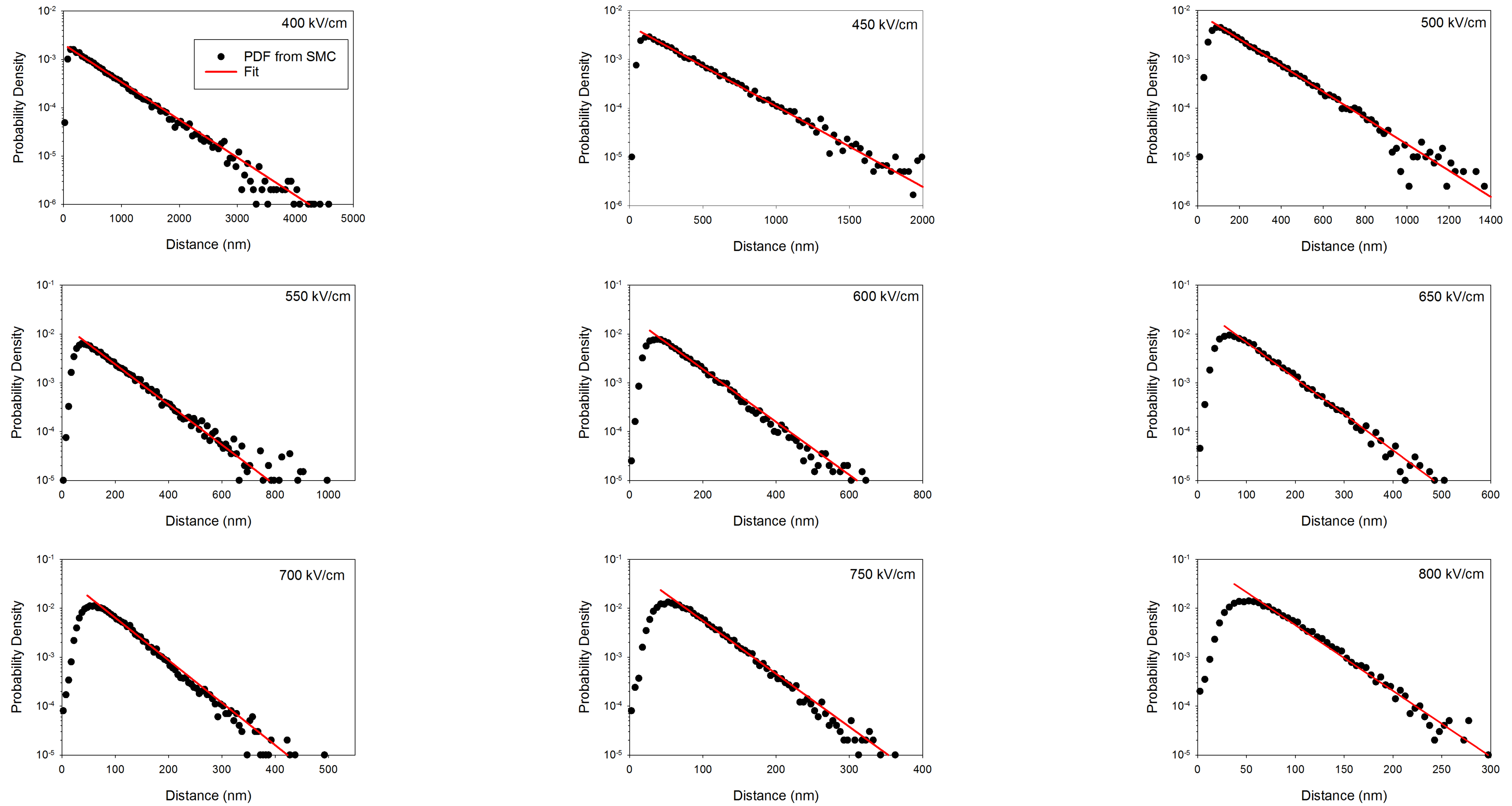

Figure D.7: Fitted probability densities generated from the SMC for holes at 200 K, fitted using the hard deadspace PDF equation.

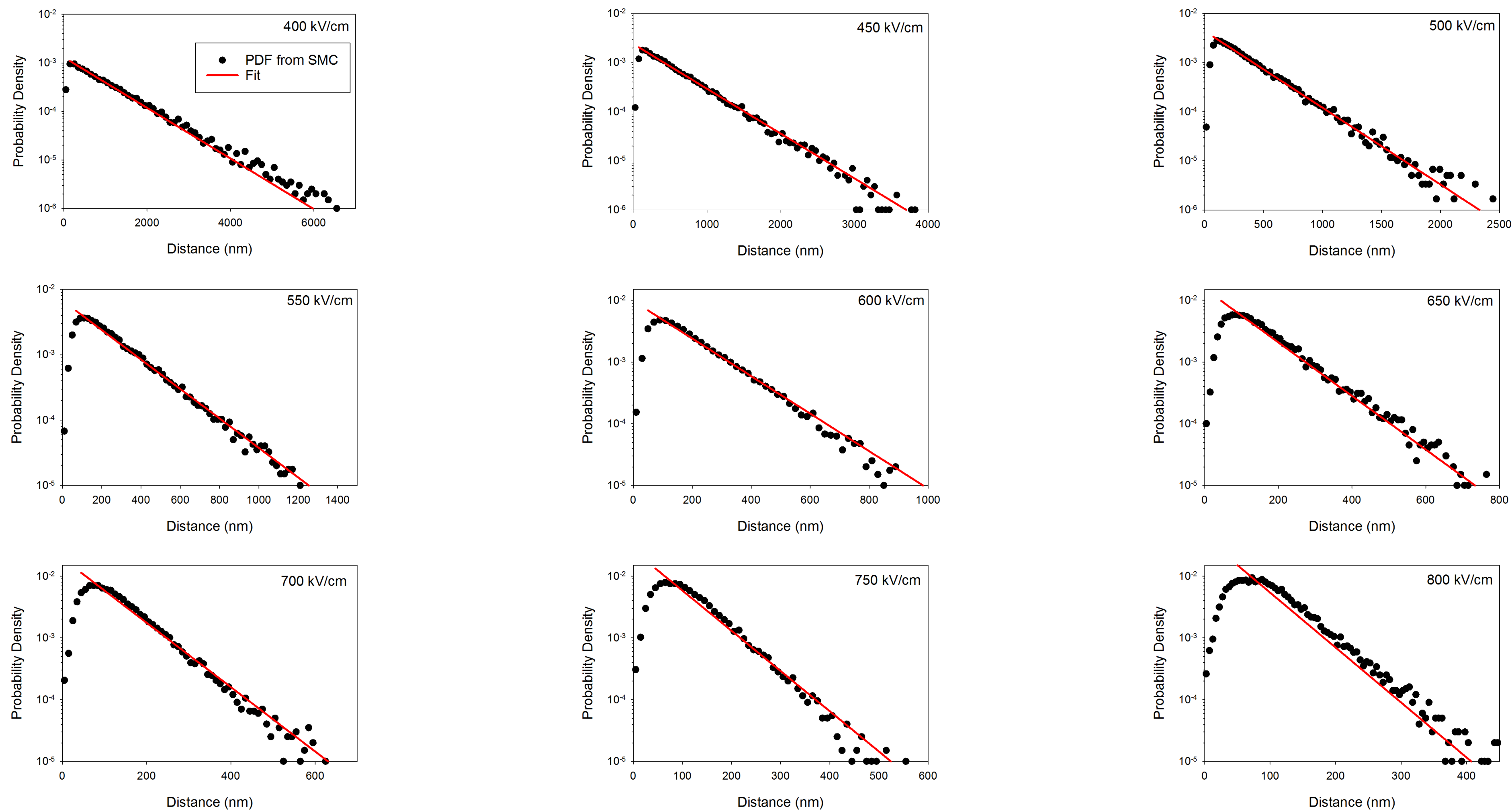

Figure D.8: Fitted probability densities generated from the SMC for electrons at 150 K, fitted using the hard deadspace PDF equation.

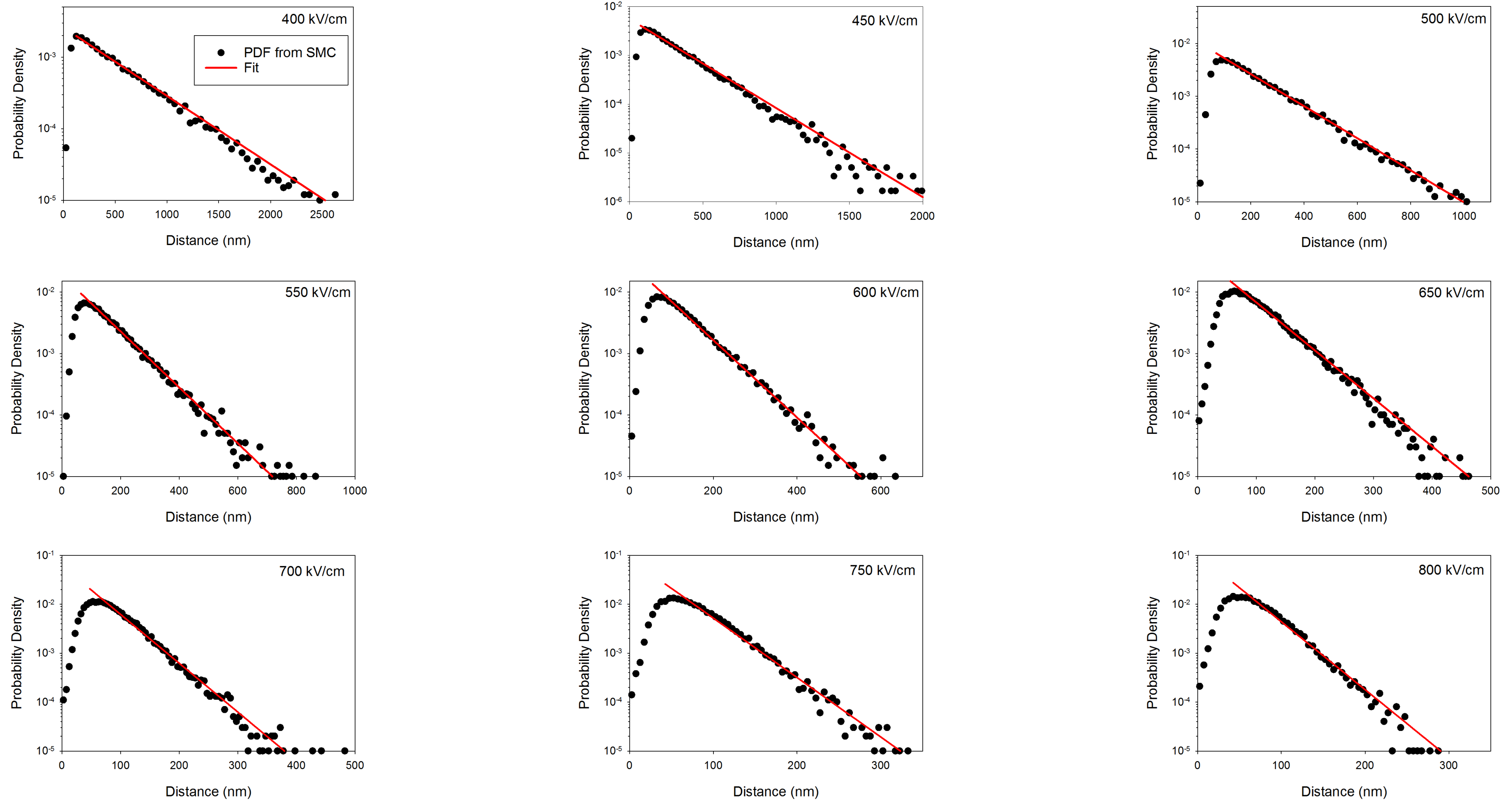

Figure D.9: Fitted probability densities generated from the SMC for holes at 150 K, fitted using the hard deadspace PDF equation.

Appendix E

Mesa Fabrication Process

This chapter contains the outline of the Mesa SPAD fabrication process. This process has been split into 2 sections, the section on how to form the implanted layers (appendix E.1), and the section on sample fabrication (appendix E.2).

E.1 Implantation

1. RCA Clean Wafer
2. Deposit screening oxide via PECVD (30 nm)
3. Send sample for Boron implantation
4. Clean wafer
5. Anneal and Oxidation
6. Strip oxide in 40% HF
7. Clean wafer
8. Deposit screening oxide via PECVD (30 nm)
9. Send sample for Phosphorus implantation
10. Clean wafer
11. Anneal and Oxidation

E.2 Sample Fabrication

1. Cleave sample
2. Strip Oxide in 10% HF ($\sim$ 3 minutes)
3. Clean Sample
4. Bake sample at 100°C for at least 1 minute.
5. Mount sample in open top spinner using green tacky paper.
6. Cover the sample in HMDS, wait 30 seconds, then spin (4000 rpm for 30 s)
7. Deposit SPR220, spin (2000 rpm for 30 s)
8. Remove sample from spinner, remove tacky paper.
9. Bake sample for min. 90 s at 100°C
10. Align to mask plate, and expose sample using UV300.
11. Develop sample for 90 s, using MF26A developer.
12. Etch mesa in ICP using Silicon-1
13. Strip resist using EKC830 resist stripper.
14. Clean sample

    If required perform isolation etch to isolate devices and bottom contacts, otherwise skip:

15. Bake sample at 100°C for at least 1 minute.
16. Mount sample in open top spinner using green tacky paper.
17. Cover the sample in HMDS, wait 30 seconds, then spin (4000 rpm for 30 s)
18. Deposit SPR220, spin (4000 rpm for 30 s)
19. Remove sample from spinner, remove tacky paper.
20. Bake sample for min. 90 s at 100°C
21. Align to mask plate, and expose sample using UV300.
22. Develop sample for 60 s, using MF26A developer.
23. Etch mesa in ICP using Silicon-1

24. Strip resist using EKC830 resist stripper.

25. Clean sample

26. Deposit oxide, to same height as mesa via PECVD.

27. Clean sample

28. Bake sample at 100°C for at least 1 minute.

29. Mount sample in open top spinner using green tacky paper.

30. Cover the sample in HMDS, wait 30 seconds, then spin (4000 rpm for 30 s)

31. Deposit SPR220, spin (4000 rpm for 30 s)

32. Remove sample from spinner, remove tacky paper.

33. Bake sample for min. 90 s at 100°C

34. Align to mask plate, and expose sample using UV300.

35. Develop sample for 60 s, using MF26A developer.

36. Etch Oxide in ICP using oxide-1. Etch rate $\sim$ 20 nm.min^{-1}

37. Strip resist using EKC830 resist stripper.

38. Clean sample

    For contact deposition:

39. Bake sample at 100°C for at least 1 minute.

40. Mount sample in open top spinner using green tacky paper.

41. Cover the sample in HMDS, wait 30 seconds, then spin (4000 rpm for 30 s)

42. Deposit SPR220, spin (4000 rpm for 30 s)

43. Remove sample from spinner, remove tacky paper.

44. Bake sample for min. 90 s at 100°C

45. Align to mask plate, and expose sample using UV300.

46. Develop sample for 60 s, using MF26A developer.

    If contacts are to be placed on SiOx:

47. Place sample in RIE

48. Rougher surface with 2 min Oxide etch to improve contact adhesion
49. Degrease Al wire in lengths of 20 cm
50. Load Al wire into 4 fired coils, then load coils into thermal evaporator. Coils to be placed at 12 cm height in parallel.
51. Coat inside of evaporator bell with Bell Shine
52. Ash sample for 3 minutes in barrel asher
53. Place sample in 19:1 DI water: ammonia solution for 30 s
54. Blow dry sample with N2, load sample into evaporator.
55. Deposit 800 nm of Al with evaporator.
56. Lift-off using DMSO.