Characterisation of GaAsBi based heterojunction and MQW photodiodes

Shiyuan Gao

University of Sheffield

A thesis submitted for the degree of Doctor of Philosophy

School of Electrical and Electronic Engineering

25/06/2025

Declaration of Authorship

I, Shiyuan Gao, declare that this thesis titled, "Characterisation of GaAsBi based heterojunction and MQW photodiodes" and the work presented in it are my own. I confirm that:

- This work was completed primarily during my PhD period at the University of Sheffield.
- Any part of this thesis previously submitted for a degree or other qualification at any institution is clearly indicated.
- Published work consulted is properly attributed.
- All quotations from other works are cited, and apart from those, this thesis represents my original work.
- I have acknowledged all major sources of assistance.
- Where the research involved collaboration, I have clarified my individual contributions and those of others.

Signed: Shiyuan Gao

Date: 25/06/2025

Abstract

This dissertation focuses on incorporating dilute bismuth (Bi) into a GaAs substrate to enable avalanche photodiodes (APDs) operating in the near-infrared range while reducing excess noise at high gain.

GaAsBi epitaxial layers grown by molecular beam epitaxy (MBE) incorporated small amounts of Bi, tuning the bandgap and enhancing spin–orbit splitting. XRD (with consistency checks against the optical response) determined Bi composition and layer thickness, while SIMS provided a qualitative and semi-quantitative indicator of dopant profiles, axial layer placement, and relative Bi incorporation trends.

By fabricating p–i–n and n–i–p heterojunction devices and multiple quantum well (MQW) structures, this work systematically investigated relationships among lattice strain, optical response, dark current, and avalanche gain–noise performance. Moderate Bi content effectively suppresses hole-initiated impact ionization at high reverse biases, reducing the β/α ratio and lowering the APD excess noise. Multi-period MQWs, maintaining controllable strain, achieve noise reduction similar to thick single GaAsBi layers while exhibiting stronger sensitivity around 1 μm.

Excessive Bi content or MQW period number leads to strain accumulation and defects, increasing dark current and highlighting the need for further optimization of epitaxial growth and strain compensation. The dissertation also explores potential applications and limitations, and proposes future directions including temperature-dependent characterization, strain management, advanced multi-region designs, and large-scale integration.

In summary, GaAsBi shows promise for high-gain, low-noise near-infrared detection, and with continued refinement of epitaxial processes and device architectures, could enable high-speed optical communications, lidar, and highly sensitive sensing systems.

List of publications

Journal papers:

1. X. Tao, X. Jin, **S. Gao**, X. Yi, Y. Liu, T. B. O. Rockett, N. J. Bailey, F. Harun, N. A. Adham, C. H. Tan, R. D. Richards and J. P. R. David, "Engineering of Impact Ionization Characteristics in GaAs/GaAsBi Multiple Quantum-Well Avalanche Photodiodes," *ACS Photonics*, vol. 11, no. 11, pp. 4846–4853, Nov. 2024, doi: 10.1021/acsphotonics.4c01343.
2. **S. Gao**, X. Tao, X. Jin, Y. Liu, N. J. Bailey, C. H. Tan, R. D. Richards, J. P. R. David, "Characterisation of GaAs/GaAsBi heterostructure photodiodes on GaAs substrate" (in preparation).
3. X. Tao, X. Jin, **S. Gao**, Y. Chen, Q. Tian, X. Yi, Y. Liu, T. B. O. Rockett, N. A. Adham, C. H. Tan, R. D. Richards and J. P. R. David, "Impact Ionization Characteristics in GaAs/GaAsBi Multiple Quantum-Well Avalanche Photodiodes up to 120 wells" (in preparation).
4. X. Tao, **S. Gao**, J. Bork, X. Jin, Y. Chen, Q. Tian, X. Yi, Y. Liu, R. D. Richards, J. Zide and J. P. R. David, "Impact Ionization Characteristics in $InAlAsBi_{0.02}$ Avalanche Photodiodes on InP substrate" (in preparation).

Oral presentations:

1. **S. Gao**, X. Tao, X. Jin, Y. Liu, N. J. Bailey, C. H. Tan, J. P. R. David and R. D. Richards, "Characterisation of GaAs/GaAsBi heterostructure photodiodes," presented at the *Semiconductor and Integrated Opto-Electronics Conference (SIOE 2024)*, Cardiff, U.K., Apr. 2024.

2. **S. Gao**, X. Tao, X. Jin, Y. Liu, N. J. Bailey, C. H. Tan, J. P. R. David and R. D. Richards, "Characterisation of GaAs/GaAsBi heterojunction photodiodes," presented at *UK Semiconductors 2024*, Sheffield, U.K., Jul. 2024.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my primary supervisor, Professor John P. R. David. His profound insights and unwavering support have been pivotal to this research, offering not only academic guidance but also a significant boost to my professional development.

I also wish to extend my thanks to my co-supervisors, Dr. Robert D. Richards, Professor Jo S. Ng, and Professor Chee H. Tan. Their diverse expertise and practical feedback have continually broadened my perspective, helping me refine both my research methodologies and critical thinking skills.

I am deeply grateful to Dr. Nicholas J. Bailey, Dr. Thomas B. O. Rockett, and Dr. Nada Adham for their constructive suggestions at crucial moments, as well as their enthusiastic discussions on intricate details—both of which have been instrumental in ensuring steady progress in this project.

My special thanks go to Dr. Yuchen Liu and Dr. Xiao Jin for their invaluable guidance during the early stages of my research. Their generosity in sharing knowledge and encouragement to explore innovative ideas laid a solid foundation for my research design and methodology.

I would also like to acknowledge my fellow doctoral companions—Dr. Xiaofeng Tao, Dr. Yifan Liu, Dr. Matthew Carr, Qingyu Tian, and Douglas Crackett,—who stood alongside me throughout my PhD journey. Their camaraderie and support, whether through academic discussions or tackling daily research challenges, have made this phase of my studies both productive and fulfilling.

I am equally appreciative of Dr. Yuting Ji, Dr. Tarick Blain, and Dr. Jonathan Petticrew for the invaluable assistance they provided in the lab. Their professional expertise and commitment to best practices were essential to the smooth execution of my experimental work.

Furthermore, I am thankful for the contributions of Dr. Nicholas J. Bailey and Dr. Thomas B. O. Rockett to the molecular beam epitaxy (MBE) growth and device fabrication. The advanced techniques and resources they introduced have substantially expanded the scope and depth of my research.

I also consider it a privilege to be part of the Impact Ionization Group at the University of Sheffield. The team's enthusiasm and collaborative spirit continuously inspired my research ideas and motivated me to strive further along my academic path.

Lastly, I would like to express my deepest gratitude to my parents. Their unwavering financial and emotional support, coupled with their steadfast belief in my abilities, has served as the bedrock of my academic achievements and remained a constant source of motivation throughout this journey.

Contents

Chapter 1 Introduction 1

1.1 Basic introduction of Group-III–V Semiconductors 1

1.2 Photodetectors 2

1.3 APD 4

1.4 Competing Materials in Telecommunication Technology 6

1 .5 Research Motivation 9

1 .5.1 Key Advantages of GaAsBi 10

1 .5.2 Application Prospects 10

1 .5 .3 Principal Scientific Challenges 10

1 .6 Thesis Organisation 11

Chapter 2: Theoretical Background 16

2.1 PN junction 16

2.2 Absorption of light 17

2.3 Diode equation 20

2.4 Introduction of GaAs and GaAsBi 22

2.4.1 introduction of GaAs 22

2.4.2 Introduction of GaAsBi 23

2.4.3 Valence-Band Anti-Crossing (VBAC) Model 24

2.4.4 Growth of GaAsBi 26

2.5 Avalanche multiplication and Impact ionization coefficient 28

2.6 Random Path Length (RPL) model 32

2.7 Excess noise 34

Chapter3: Molecular Beam Epitaxy (MBE) Growth Techniques 39

3.1 Producing and Sustaining a Vacuum 39

3.2 Molecular Sources and Generating a Molecular Beam 42

3.3 Cell Temperature Control and Outgassing....46

3.4 Automation....47

3.5 III-V Growth Overview....47

3.5.1 Limitations of low-temperature III–V MBE growth....48

3.6 Reflection High-Energy Electron Diffraction (RHEED)....49

3.6.1 RHEED Measurements for Substrate Heater Calibration....50

3.6.2 RHEED Measurements for Finding Growth Rates....52

3.6.3 RHEED Measurements for As2:Ga 1:1 Calibration....53

3.7 Heater Calibration Process....55

3.8 Growth Rate Calibration....60

3.9 GaAs Growth....61

3.10 GaAsBi Growth....62

3.10.1 Review of GaAsBi Growth Investigation....62

3.10.2 Summary of GaAsBi Growth Conditions....64

Chapter 4: Characterization Techniques....71

4.1 X-ray Diffraction (XRD)....71

4.1.1 Bragg's Law....71

4.1.2 Design of Diffractometer....72

4.2 Nomarski Microscopy....73

4.3 Photoluminescence (PL)....75

4.3.1 Radiative Recombination....75

4.3.2 Non-radiative Recombination....76

4.3.3 Experimental Setup....78

4.3.4 Monochromator....80

4.4 Current-Voltage (IV) Measurement....81

4.5 Capacitance-Voltage (CV) Measurement....84

4.6 Capacitance-Voltage (CV) Measurement Insights..88
4.7 Photomultiplication..89
4.8 Photocurrent...93
Chapter 5: Characterisation of GaAs/GaAsBi Heterostructures..................................98
5.1 Introduction..98
5.2 Layer details...100
5.3 SIMS of heterostructures 102
5.4 XRD of heterostructures 105
5.5 Electrical characterisation of heterostructures 107
5.5.1 IV measurements 107
5.5.2 CV measurements 117
5.6 Photo spectrum of heterostructures...123
5.7 Photomultiplication characterisation of heterostructures..............................126
5.8 Discussion...133
5.8.1 GaAsBi Epitaxy and Material Quality..133
5.8.2 Electrical Characteristics and Depletion Region Control 134
5.8.3 Optical Response and Bandgap Tuning 134
5.8.4 Avalanche Multiplication and Hole Ionization Suppression.......................135
5.8.5 Significance, Limitations, and Future Work 135
Chapter 6: Characterisation of GaAs/GaAsBi MQW photodiodes 141
6.1 Introduction..141
6.2 Layer details...142
6.3 XRD Characterization of MQW Diodes..145
6.4 Electrical characterisation of MQW Diodes 148
6.4.1 IV measurements 148
6.4.2 CV measurements 158

6.5 Photo spectrum of MQW Diodes....165
6.6 Photomultiplication characterisation of MQW Diodes....168
6.6 Photomultiplication Characterisation of MQW Diodes....174
6.6.1 Setup Summary....174
6.6.2 Photomultiplication Performance of Different MQW Periods....175
6.6.3 Influence of Average Period Bi Content on Impact Ionization....176
6.7 Discussion....177
6.7.1 Quantum Well Design and Hole Suppression....177
6.7.2 Strain Management and Device Uniformity....177
6.7.3 Overall Performance and Application Prospects....178
6.7.4 Conclusion....178
Chapter 7Conclusion and Future Work....183
7.1 Summary of Research....183
7.2 Future Work....184

List of Figures

Figure 2.1 Band diagrams of a p–n junction at equilibrium (a), forward bias(b), and reverse bias(c). 16

Figure 2.5: Lu model. Process 1 illustrates how Bi–Ga bonds are formed, increasing the Bi content at the growing surface. Process 3 shows how these bonds are replaced by stronger As–Ga bonds, reducing the Bi content. Adapted from [7].......... 27

Figure 2.6: Lewis model. Process 1 shows Bi bonding to an exposed Ga site, increasing Bi surface coverage and reducing Ga surface coverage. Process 2 shows a Ga atom bonding to a Bi site as part of the growth process, decreasing Bi surface coverage. Process 3 shows a Bi atom being thermally ejected to the surface, further reducing Bi surface coverage. Adapted from [6].......... 27

Figure 2.7: Schematic illustration of impact ionisation and avalanche multiplication in a reverse-biased p-i-n junction. (a) Branching of an electron-initiated avalanche within a multiplication region of width w (coordinate x). (b) Reverse-bias band diagram indicating carrier drift: electrons drift towards the n-contact (opposite the field), whereas holes drift towards the p-contact (with the field), motivating the definitions of Me and Mh. (c) Energy-band picture of an impact-ionisation event showing the band gap Eg and the threshold energy Eth.......... 29

Figure 3.1: A diagram illustrating the general positioning of pumps within the MBE-STM system. Gate valves used to isolate chambers and pumps in the system are represented as solid gray rectangles. The scanning tunneling microscope (STM) connected to the growth chamber was not utilized in this work and is therefore not detailed.......... 41

A key requirement for an MBE system is the high stability of the source beams. When source material of sufficient purity is available, the generation of the elemental beam from the bulk material is typically achieved through a thermal process. Heating the material within the vacuum raises its vapor pressure, resulting in a significant flux of evaporated or sublimated material, depending on whether the source is a liquid or a solid at the given temperature. The traditional device used to hold the bulk source material and regulate the beam flux through radiative heating is a Knudsen effusion

cell. 43

Figure 3.3: A generalized cross-section of a Knudsen effusion cell. 44

Figure 3.4: RHEED System Diagram. The diffraction pattern (periodic green rods) arises from elastic diffraction collisions, while the green background haze is caused by inelastic collisions and electron scattering within the reactor. 49

Figure 3.5: Simplified representation of the surface profile during the growth of a single monolayer. 52

Figure 3.6: Comparison of Bi incorporation for various $As2$ and $As4$:Ga atomic flux ratios. Reprinted from *Journal of Crystal Growth*, 390, R. D. Richards et al., *Molecular beam epitaxy growth of GaAsBi using* $As2$ *and* $As4$, 120–124, Copyright (2014) [10]. 53

Figure 3.7: Dynamic reconstructions of GaAs during growth. Adapted from [11]. 54

Figure 3.8: Oxide-removed "3D" RHEED pattern. 56

Figure 3.9: (2×4) RHEED pattern. Annotations indicate the integer and fractional order diffractions. 56

Figure 3.10: Hazy As cap. The incident electron beam is absorbed and scattered by the amorphous As surface, preventing the formation of diffraction peaks. 57

Figure 3.11: $c(4 \times 4)$ Reconstruction after As cap removal. Annotations indicate the integer and fractional order diffractions. Image captured along the $[110]$ azimuth at 14.5 keV. 58

Figure 3.12: Plot of heater calibration points with an exponential trend-line. 59

Figure 3.13: Growth Rate Measurements 61

Figure 3.13: Bake-out preparation of MBE system 66

Figure 4.1: Bragg's Law in X-ray Diffraction 72

Figure 4.2 Schematic diagram of XRD measurement. Note: The monochromator and analyzer are omitted in this diagram for clarity. 73

Figure 4.3 Basic Components in a Nomarski Microscope 74

Figure 4.4: The Processes Involved in Photoluminescence (PL) a) An incoming photon excites an electron from the valence band to the conduction band, leaving a hole in the valence band. b) The electron in the conduction band thermalizes to the lower energy edge of the conduction band. c) During the recombination of the electron with the hole, a photon is emitted with the same energy as the material's bandgap. 75

Figure 4.5 Shockley Read-Hall Recombination via Trap State in the Band Gap. 77

Figure 4.6: Schematic illustration of CHCC (Conduction band to Heavy hole valence band, Conduction band to higher energy state of Conduction band) and CHSH (Conduction band to Heavy hole valence band, Spin orbit split-off band to Heavy hole valence band) Auger recombination processes. 78

Figure 4.7: The Photoluminescence Setup Used in This Work 79

Figure 4.8: Schematic Diagram of a Horiba iHR320 Monochromator 80

Figure 4.9: Schematic Diagram of the I-V Measurement Setup..........................82

Figure 4.10: a) Equivalent Circuit for Series Model - Showcases the series resistance Rs. b) Equivalent Circuit for Parallel Model - Illustrates the parallel resistance Rp and the reactance x. c) Measurement Guidelines for Capacitance - Provides best practices for capacitance measurement using a standard LCR meter. 85

Figure 4.11: Common Errors in Capacitance Measurement This figure illustrates the typical errors encountered in capacitance measurement. The ideal result is represented by the black line, while the offset error and gain error are depicted by the red line and blue line, respectively, as adapted from reference [9]. 87

Figure 4.12: Schematic Diagram of the Photomultiplication Setup This diagram illustrates the setup used to measure photocurrent versus reverse bias. Different wavelength lasers can be utilized depending on the requirements of the experiment. 90

Figure 4.13: Baseline Correction and Gain Calculation for 3100nm GaAsBi p-i-n MQWs Diode..92

Figure 4.14: Schematic diagram of the photocurrent measurement setup...........94

Figure 5.1 Schematic of the n-i-p and p-i-n heterostructure devices structure..100

Figure 5.2 SIMS results of STK heterostructures 102

Figure 5.3 X-ray diffraction results of STK heterostructures 105

Figure 5.4 (a) Reverse dark current, (b) Forward dark current, (c) Reverse dark current density, (d) Forward dark current density of STK2D 109

Figure 5.5 (a) Reverse dark current, (b) Forward dark current, (c) Reverse dark current density, (d) Forward dark current density of STK2J 110

Figure 5.6 (a) Reverse dark current, (b) Forward dark current, (c) Reverse dark current density, (d) Forward dark current density of STK62 111

Figure 5.7 (a) Reverse dark current, (b) Forward dark current, (c) Reverse dark current density, (d) Forward dark current density of STK64 112

Figure 5.8 a) Reverse dark current densities and (b) Forward dark current densities of heterostructure devices 115

Figure 5.9 (a) CV of STK2D; (b) Capacitance per area of STK2D 117

Figure 5.10 (a) CV of STK2J; (b) Capacitance per area of STK2J 118

Figure 5.11 (a) CV of STK62; (b) Capacitance per area of STK62 119

Figure 5.12 (a) CV of STK64; (b) Capacitance per area of STK64 120

Figure 5.13 (a)-(d) Measured CV of heterostructure devices, black dots are raw data and red lines are fitting from Poisson's equation 121

Figure 5.14 Doping profile of heterostructure devices 122

Figure 5.15 (a)-(d) Bias dependent photocurrent of heterostructure devices; (e)Absorption coefficient of heterostructure devices (@0V) 124

Figure 5.16 (a) M_e (Blue dots), M_{530} (Green dots), M_{625} (Red dots), M_{780} (Pink dots), M_{mix} (Cyan dots) of STK2D with at different voltage; (b) M_e -1 (Blue dots), M_{530} -1 (Green dots), M_{625} -1 (Red dots), M_{780} -1 (Pink dots), M_{mix} -1 (Cyan dots) of STK2D with at different voltage; (c) M_e (Blue dots), M_{530} (Green dots), M_{625} (Red dots), M_{780} (Pink dots), M_{mix} (Cyan dots) of STK2D with at different electric field; (d) M_e -1 (Blue dots), M_{530} -1 (Green dots), M_{625} -1 (Red dots), M_{780} -1 (Pink dots), M_{mix} -1 (Cyan dots) of STK2D with at different electric field 127

Figure 5.17 (a) M_h (Blue dots), M_{530} (Green dots), M_{625} (Red dots), M_{780} (Pink dots), M_{mix} (Cyan dots) of STK62 with at different voltage; (b) M_h -1 (Blue

dots), M_{530} -1 (Green dots), M_{625} -1 (Red dots), M_{780} -1 (Pink dots), M_{mix} -1 (Cyan dots) of STK62 with at different voltage; (c) M_h (Blue dots), M_{530} (Green dots), M_{625} (Red dots), M_{780} (Pink dots), M_{mix} (Cyan dots) of STK62 with at different electric field; (d) M_h -1 (Blue dots), M_{530} -1 (Green dots), M_{625} -1 (Red dots), M_{780} -1 (Pink dots), M_{mix} -1 (Cyan dots) of STK62 with at different electric field 128

Figure 5.18 M_e -1 (dots and triangles) verse electric field of heterostructures with RPL simulation (solid lines) 129

Figure 5.19 M_h -1 (dots) verse electric field of heterostructures with RPL simulation (solid lines) 129

Figure 6.1 Schematic diagram of cross-section of the MQW p-i-n device structures 142

Figure 6.2 X-ray diffraction results of MQW diodes 145

Figure 6.3: (a) Reverse dark current, (b) forward dark current, (c) reverse dark current density, and (d) forward dark current density of the 2QW photodiodes measured for multiple nominal mesa diameters (D = 50, 100, 200, and 400 µm). In this and the following diameter-dependent plots in Chapter 6, traces with the same colour correspond to the same nominal diameter, while solid and dashed curves represent measurements from different nominally identical devices of that same diameter, thereby visualising device-to-device spread at fixed geometry. 148

Figure 6.4 (a) Reverse dark current, (b) forward dark current, (c) reverse dark current density, (d) forward dark current density of 5QW 149

Figure 6.5 (a) Reverse dark current, (b) forward dark current, (c) reverse dark current density, (d) forward dark current density of 15QW 150

Figure 6.6 (a) Reverse dark current, (b) forward dark current, (c) reverse dark current density, (d) forward dark current density of 30QW 151

Figure 6.7 (a) Reverse dark current, (b) forward dark current, (c) reverse dark current density, (d) forward dark current density of 40QW 152

Figure 6.8 (a) Reverse dark current, (b) forward dark current, (c) reverse dark current density, (d) forward dark current density of 80QW 153

Figure 6.9 (a) Reverse dark current, (b) forward dark current, (c) reverse dark current density, (d) forward dark current density of 120QW 154

Figure 6.10 (a) Reverse dark current densities and (b) Forward dark current densities of the MQW samples 155

Figure 6.13 (a) CV data of 15QW (b) Capacitance per area of 15QW............. 159

Figure 6.17 (a) CV data of 120QW (b) Capacitance per area of 120QW......... 160

Figure 6.18 Doping profile of MQW samples 161

Figure 6.19 (a)-(g) Bias dependent photocurrent intensity of MQW devices; (h) Photocurrent intensity of MQW devices (@0V) 165

Figure 6.20 (a) M_e (blue dots), M_{530} (green dots), M_{625} (red dots), M_{780} (cyan dots), M_{850} (purple dots), M_{mix} (dark green dots) of 5QW at different voltage; (b) M_e - 1(blue dots), M_{530} - 1(green dots) , M_{625} - 1(red dots), M_{780} - 1(cyan dots), M_{mix} - 1(dark green dots) of 5QW; (c) M_e (blue dots), M_{mix} (dark green dots) of 5QW with RPL fitting (solid lines) at different voltage; (d) M_e - 1(blue dots), M_{mix} - 1(dark green dots) of 5QW with RPL fitting (solid lines) at different voltage. 169

Figure 6.21 (a) M_e (blue dots), M_{530} (green dots), M_{625} (red dots), M_{780} (cyan dots), M_{850} (purple dots), M_{mix} (dark green dots) of 40QW at different voltage; (b) M_e - 1(blue dots), M_{530} - 1(green dots) , M_{625} - 1(red dots), M_{780} - 1(cyan dots), M_{850} - 1(purple dots), M_{mix} - 1(dark green dots) of 40QW; (c) M_e (blue dots), M_{mix} (dark green dots) of 40QW with RPL fitting (solid lines) at different voltage; (d) M_e - 1(blue dots), M_{mix} - 1(dark green dots) of 40QW with RPL fitting (solid lines) at different voltage. 170

Figure 6.22 (a) M_e (blue dots), M_{mix} (dark green dots) of 15QW with RPL fitting (solid lines) at different voltage; (b) M_e - 1(blue dots), M_{mix} - 1(dark green dots) of 15QW with RPL fitting (solid lines) at different voltage. 171

Figure 6.23 (a) M_e (blue dots), M_{mix} (dark green dots) of 30QW with RPL fitting (solid lines) at different voltage; (b) M_e - 1(blue dots), M_{mix} - 1(dark green dots) of 30QW with RPL fitting (solid lines) at different voltage. 171

Figure 6.24 (a) M_e (blue dots), M_{mix} (dark green dots) of 80QW with RPL fitting (solid lines) at different voltage; (b) M_e - 1(blue dots), M_{mix} - 1(dark green dots) of 80QW with RPL fitting (solid lines) at different voltage. 172

Figure 6.25 α of GaAs (black line), β of GaAs (black dash line), α of GaAsBi with 0.7% Bi(red line), β of GaAsBi with 1.7% Bi (red dash line), α/ β ratio of GaAs (black dots), α/ β ratio of MQW(red dots), vs inverse electric field

.... 173

List of Tables

Table 1.1 Principal group-III, -IV and -V elements and those most commonly employed in III–V alloys. ...2

Table 1.2: Qualitative comparison of competing APD material platforms for the 1.3-1.55~μm telecommunication window. ...8

Table 5.1 Details of the heterostructure devices ...100

Table 5.2 Dark current density and ideality factor of heterostructure devices ..116

Table 5.3 heterostructure CV model parameters compared with nominal thickness...122

Table 6.1 Design of MQW samples...143

Table 6.2 XRD fitting results of MQW diodes ...146

Table 6.3 Dark current density and ideality factor of each MQW sample...155

Table 6.4 CV model parameters compared with design thickness ...161

Table 6.5 Average period Bi concentration of MQW diodes by XRD...173

Chapter 1 Introduction

This chapter begins with a concise survey of group-III–V compound semiconductors and highlights the unusual physical properties of dilute-bismide gallium arsenide (GaAsBi) that have spurred intense research interest in recent years. It then assesses the current and prospective technological applications of GaAsBi. The closing section states the motivation for this study and outlines the structure of the dissertation.

1.1 Basic introduction of Group-III–V Semiconductors

Compounds formed by pairing elements from column III (B, Al, Ga, In, Tl) with those from column V (N, P, As, Sb, Bi) of the periodic table are collectively known as III–V semiconductors; GaAs is the prototypical example (Table 1.1). Owing to their predominantly direct band gaps and excellent thermal stability, III–V materials underpin a wide range of optoelectronic and high-speed electronic devices, including photodetectors and light-emitting diodes. Their band gaps span from the deep-ultraviolet to the far-infrared, giving this material family a unique versatility across the electromagnetic spectrum.

During the past three decades, researchers have shown that isovalent incorporation of small amounts of size- or electronegativity-mismatched atoms—most notably nitrogen (N) or bismuth (Bi)—provides a powerful means of tailoring the electronic structure of III–V hosts. In dilute alloys such as $GaAs_{1-x}N_x$ and $GaAs_{1-\gamma}Bi_\gamma$, even minute substitutional fractions produce an anomalously large band-gap bowing, leading to a pronounced reduction of the fundamental gap [1, 2]. This phenomenon opens a viable route to long-wavelength light sources and detectors on the well-established GaAs platform, thereby motivating the detailed investigation presented in this thesis.

Table 1.1 Principal group-III, -IV and -V elements and those most commonly employed in III–V alloys.

III	IV	V
B	C	N
Al	Si	P
Ga	Ge	*As*
In	Sn	*Sb*
Tl	Pb	Bi

1.2 Photodetectors

A photodetector is an essential electronic component that senses incoming light and converts it into an electrical output, enabling a wide spectrum of scientific and technological functions. Its operation relies on the photoelectric effect: photons striking the detector's surface liberate charge carriers—electrons and/or holes—which are subsequently collected to produce a measurable signal. Under linear operating conditions, the electrical power generated scales with the optical intensity, permitting very precise quantification of light levels.

Multiple categories of photodetectors are engineered for particular roles. Photodiodes (PDs) are among the most prevalent, especially in fibre-optic receivers and consumer electronics, where they deliver rapid response times and high sensitivity. Avalanche photodiodes (APDs) include an internal gain mechanism that multiplies weak signals, rendering them well-suited to low-illumination applications such as LIDAR and long-

haul telecommunications; an expanded discussion of PDs and APDs is provided in the following section. Phototransistors, structurally analogous to photodiodes, furnish additional current gain, making them advantageous for very low-light sensing. Photomultiplier tubes (PMTs) are vacuum devices capable of single-photon detection by amplifying the initial electron cascade on the order of 10^5–10^8, and are extensively employed in nuclear research, astronomy, and medical imaging [3]. Finally, solid-state array detectors—charge-coupled devices (CCDs) and complementary metal-oxide–semiconductor (CMOS) sensors—capture light on millions of miniature photodiodes to generate high-resolution image data for cameras, diagnostic instruments, and related equipment.

The effectiveness of a photodetector is characterised by several key parameters. Responsivity quantifies the efficiency with which optical power is translated into electrical current, while quantum efficiency expresses the proportion of collected carriers relative to incident photons. Dark current—current observed in the absence of light—constitutes a noise source that should be minimised. The noise-equivalent power (NEP) defines the weakest optical signal discernible above the noise floor; smaller NEP values indicate superior sensitivity. Response time or bandwidth dictates how swiftly the device reacts to fluctuations in illumination, a decisive factor for high-speed systems. Dynamic range denotes the span of optical intensities accommodated without saturation, and linearity guarantees that output remains proportional to input across that span.

Photodetectors find widespread application across numerous domains. In optical communications, they form the terminal element of fiber networks, translating modulated light into electrical data for high-capacity links. In imaging, they underlie digital cameras as well as medical devices such as microscopes and telescopes, enabling the capture of high-resolution images. Environmental and industrial monitoring leverage photodetectors for air-quality sensing, gas detection, motion control, and automated lighting. In LIDAR, photodetectors measure time-of-flight reflections for autonomous vehicles and topographic surveys. Within medical diagnostics, they are integral to equipment such as CT scanners, X-ray systems, and fluorescence microscopes, facilitating light-based analyses for disease detection.

Despite their versatility, photodetectors exhibit certain constraints. Some variants cannot directly probe non-radiative recombination in materials, potentially overlooking specific energy-loss pathways. Thermal effects pose another challenge; elevated or extreme temperatures can degrade performance metrics such as dark current and responsivity. Moreover, each detector type is optimised for a particular spectral band, and its efficiency diminishes outside that range unless additional engineering measures are adopted.

In summary, photodetectors are indispensable for translating optical information into electrical form, supporting a vast array of applications in telecommunications, imaging, environmental surveillance, and research. Their accuracy, sensitivity, and adaptability across diverse wavelength regimes render them foundational to modern technology.

1.3 APD

Avalanche photodiodes (APDs) are a distinct subset of semiconductor photodetectors whose defining attribute is an internal carrier-multiplication process that dramatically elevates detection sensitivity relative to conventional p–n or p–i–n photodiodes. When a sufficiently high reverse bias is applied—typically just below the breakdown threshold—the electric field within the depletion region becomes intense enough to accelerate primary photo-generated electrons and holes to energies capable of initiating impact-ionisation events. Each such event liberates additional electron–hole pairs, precipitating a cascade that yields an avalanche gain factor M that can readily exceed two orders of magnitude. In linear-mode operation, the mean avalanche multiplication (gain) is commonly defined as

$$M \equiv \frac{I_{\mathrm{APD}}}{I_{\mathrm{ph}}},$$

i.e., the ratio of the measured photocurrent in the APD to the unity-gain photocurrent generated by the same absorbed optical power.

This intrinsic gain permits the direct conversion of extremely weak optical powers into measurable currents without the immediate need for external, low-noise amplification stages, thereby enhancing signal-to-noise performance in photon-starved applications

such as long-haul optical communication, lidar time-of-flight ranging, quantum-optical experiments, and single-molecule fluorescence detection.

A critical consequence of the stochastic nature of impact ionisation is that the internal gain is accompanied by additional noise beyond the Poisson shot noise of a unity-gain photodiode. This behaviour is commonly quantified by the excess-noise factor $F(M)$, which captures the variance of the multiplication process and therefore increases with gain. For a broad class of uniform-field devices, the classic McIntyre model expresses this trend approximately as

$$F(M) = kM + (1-k)\left(2 - \frac{1}{M}\right), k = \frac{\beta}{\alpha},$$

where α and β are the electron and hole impact-ionisation coefficients, respectively, and k indicates how strongly holes participate in the avalanche chain. Increasing M thus boosts the signal amplitude but also elevates the noise through $F(M)$, imposing a fundamental trade-off between gain and noise performance.

In this thesis, the term signal fidelity refers to the extent to which the detector preserves the incident optical waveform with minimal added noise and distortion. In practical receiver terms, fidelity is therefore linked to metrics such as output signal-to-noise ratio (SNR) and bit-error rate (BER) for communications, or timing jitter / range precision in time-of-flight sensing. Consequently, APD designs aim to maximise M while constraining $F(M)$ (and, in practice, while balancing the gain--bandwidth trade-off set by avalanche build-up time). [6]

Temperature represents another pivotal consideration: elevated operating temperatures raise dark-current levels, shift breakdown voltages, and accentuate avalanche initiation probability, collectively degrading sensitivity and linearity. To mitigate these effects, APDs often incorporate thermoelectric coolers or active bias-compensation circuits that stabilise gain across a broad thermal envelope. Reliability concerns stem from phenomena such as micro-plasma breakdown, surface-state leakage, and cumulative avalanche-induced damage; modern fabrication techniques address these challenges through carefully engineered junction-termination structures, passivation layers that shield the multiplication region from surface states, and robust packaging strategies that minimise mechanical and electrostatic stress. [7-12]

At the system level, APDs are frequently integrated with low-noise trans-impedance amplifiers in hybrid or monolithic photoreceiver modules, yielding compact assemblies

with minimal parasitic capacitance and thus superior frequency response. Advances in integrated photonics further enable the co-location of APDs with on-chip waveguides and complementary metal-oxide-semiconductor circuitry, creating highly integrated optoelectronic transceivers that occupy negligible footprint and exhibit reduced power consumption. The confluence of internal gain, high temporal resolution, and continued improvements in noise management and integration technologies underscores the pivotal role of avalanche photodiodes in contemporary and next-generation optical sensing and communication systems, where the imperative to detect extremely low photon fluxes must be reconciled with stringent speed, reliability, and size constraints. [13-17]

1.4 Competing Materials in Telecommunication Technology

The relentless demand for high-performance avalanche photodiodes in the 1.3–1.55 µm telecommunication windows has driven sustained investigation into absorber–multiplication material systems whose intrinsic electronic properties can sustain both large gain–bandwidth products and minimal excess noise. Silicon, by far the most mature and cost-effective semiconductor, offers exceptionally low intrinsic noise and benefits from the vast infrastructure of complementary metal–oxide–semiconductor processing. [18-20] Nevertheless, its indirect 1.12 eV bandgap confines spectral responsivity to wavelengths shorter than approximately 1.0 µm; therefore, silicon APDs cannot be deployed in long-haul fibre networks without the added complexity of wavelength conversion, a penalty that is difficult to reconcile with strict power and latency budgets.

Indium-gallium-arsenide(InGaAs) [21, 22] absorbers paired with indium-phosphide or indium-aluminium-arsenide(InAlAs) [23] multiplication layers currently underpin most commercial telecommunication receivers because their direct bandgaps accurately target both the O- and C-bands. Yet the narrow bandgap that confers this spectral reach also induces substantial thermally generated dark current and elevates after-pulsing probabilities, forcing the use of thermoelectric coolers or complex gating schemes in photon-starved applications such as coherent deep-space links and single-photon lidar. The economic burden is further amplified by small-diameter InP substrates, which limit

wafer throughput and impede seamless co-integration with the silicon electronics that dominate post-detection signal processing.

After-pulsing refers to spurious avalanche events triggered after a genuine detection event, most prominently in high-gain and gated (Geiger-mode) operation. Physically, a fraction of carriers generated during an avalanche can become captured by deep traps in the multiplication region; their delayed release in a subsequent gate (or shortly after bias recovery) can re-initiate impact ionisation, thereby producing false counts and excess noise. The after-pulsing probability generally increases with avalanche charge and decreases with longer hold-off time and reduced temperature, motivating mitigation strategies such as rapid quenching, dead-time control, and thermoelectric cooling. [33]

Germanium grown hetero-epitaxially on silicon has emerged as an attractive compromise, extending responsivity beyond 1.6 μm while retaining compatibility with large-area, low-cost Si wafers. Nonetheless, the 4 % lattice mismatch between Ge and Si introduces a high density of threading dislocations as well as interface trap states. [20] These crystalline imperfections act as generation–recombination centres that elevate dark current by several orders of magnitude relative to III–V counterparts, thereby eroding the very sensitivity gains sought in weak-signal regimes. Notwithstanding these materials challenges, Ge/Si APDs have also been demonstrated with gain--bandwidth products comparable to mature III-V platforms when the multiplication region is carefully engineered, highlighting both the promise of Si-compatible integration and the central importance of defect and interface control.[34]

Table 1.2: Qualitative comparison of competing APD material platforms for the 1.3-1.55~μm telecommunication window.

Platform	**Spectral reach**	**Integration / cost**	**Key strengths**	**Key limitations (APD-relevant)**
Si APD	$< 1.0~\mu$m	CMOS-native; low cost	Very low excess noise (typically $k \ll 1$); mature processing	Not directly usable at 1.3-1.55~μm (requires wavelength conversion)
InGaAs absorber + InP/InAlAs multiplication	O- and C-bands	III-V wafers; heterogeneous integration to Si	High responsivity in telecom window; commercially mature receivers	Narrow gap raises dark current; trap-assisted effects can require cooling and (in photon-starved regimes) gating; higher substrate / packaging cost
Ge-on-Si (Ge absorption; Si-compatible platform)	$\gtrsim 1.6~\mu$m possible	Si wafers; CMOS-compatible	Potential monolithic co-integration with Si photonics/electronics; demonstrated high speed in integrated platforms	4% mismatch introduces dislocations and interface traps; dark current and noise strongly depend on defect density and surface/interface passivation
$GaAs_{1-x}Bi_x$ (this work)	Tunable into NIR--SWIR (composition-dependent)	GaAs platform (larger wafers than InP)	Band-structure engineering (e.g., reduced hole ionisation / improved noise prospects) on a scalable III-V substrate	Bi incorporation requires low-temperature epitaxy; strain/defects at higher Bi; technology less mature than InGaAs/InP

Against this backdrop, dilute-bismide $GaAs_{1-x}Bi_x$ alloys constitute a compelling next-generation candidate. Incorporation of only a few percent bismuth induces a pronounced band-anticrossing interaction that simultaneously lowers the fundamental bandgap and enlarges the spin–orbit splitting energy.[24] In practice, this permits the

long-wavelength cut-off to be engineered well beyond 1.3 μm on a lattice-matched GaAs platform while suppressing Auger-assisted carrier recombination, a principal source of excess noise in narrow-gap semiconductors. GaAs substrates are available in large diameters and support high-yield epitaxy, thereby mitigating the cost and scalability limitations inherent to InP technology. Furthermore, the favourable thermal conductivity and mechanical robustness of GaAs simplify monolithic integration of GaAsBi APDs with high-speed electronics and light-emitting devices fabricated within the same process flow. Early optoelectronic characterisation reports corroborate the theoretical advantages: devices exhibit competitive gain, reduced temperature-dependent dark current and a lower excess-noise factor than conventional InGaAs/InP structures operating at equivalent wavelength targets. Consequently, GaAsBi stands poised to overcome the spectral mismatch of silicon, the thermal noise and cost penalties of InGaAs, and the defect-induced dark current challenges of Ge-on-Si, offering a versatile and manufacturable platform for the next generation of low-noise, high-bandwidth telecommunication photoreceivers.

1 .5 Research Motivation

Incorporating trace amounts of bismuth (Bi) into the GaAs crystal lattice produces the dilute bismide alloy GaAsBi, which markedly narrows the band-gap while preserving lattice matching and simultaneously boosts the spin–orbit splitting energy. This band-engineering capability extends the material's emission–absorption window from the conventional GaAs range of 870–900 nm to beyond 1 μm in the near-infrared to short-wavelength-infrared (NIR–SWIR) region [25,26]. Bi incorporation also suppresses Auger recombination and other non-radiative channels, thereby enhancing the efficiency and thermal stability of high-power optoelectronic devices [27]. Moreover, the high volatility of Bi necessitates a reduced epitaxial temperature (350-400~ °C) in order to suppress Bi desorption and achieve stable, uniform incorporation. While lower growth temperatures can kinetically hinder some thermally activated relaxation pathways, they also reduce adatom mobility and can increase the susceptibility to point-defect formation and incorporation of residual background species. Consequently, achieving both low defect densities and low impurity

incorporation in low-temperature MBE critically depends on stringent ultra-high-vacuum conditions and careful optimisation of flux ratios and growth kinetics. [28,35]

1 .5.1 Key Advantages of GaAsBi

Band-gap tunability is GaAsBi's primary strength: with only 1.8 % Bi the LED peak can be driven to 987 nm, while 6 % Bi extends the emission to 1.2 μm [25,26]. The rapid increase in spin–orbit splitting with Bi content weakens valence-band-assisted non-radiative processes and improves carrier stability under strong electric fields [27]. Low-temperature MBE growth, meanwhile, mitigates thermal stress and deep-level defect formation, providing high-quality crystals for subsequent device fabrication [28].

1 .5.2 Application Prospects

Light emitters. GaAsBi offers excellent wavelength tunability and thermal robustness. A 390 nm GaAsBi active-layer Fabry–Pérot laser achieves continuous-wave operation from room temperature down to 240 K with a threshold temperature coefficient only 40 % of that of a GaAs device at the same wavelength [28]. Increasing the Bi content to 5.8 % pushes the lasing wavelength to 1.204 μm while retaining the low thermal drift [29]. Electrically pumped single-quantum-well lasers containing 2.2 % Bi exhibit a room-temperature threshold current density of only 1.56 kA cm^{-2} at 947 nm [27].

Terahertz devices. GaAsBi layers grown at 280 °C show electron lifetimes of ~1 ps and deliver optical-pumped THz emission exceeding 3 THz; subsequent optimisation extended emission to 4.5 THz with high optical-to-THz conversion efficiency [30, 31].

High-efficiency photovoltaics. $GaAs_{0.94}Bi_{0.06}$ provides a direct 1 eV band-gap with only 0.7 % lattice mismatch to GaAs, making it an ideal fourth junction in four-junction solar cells and enabling theoretical efficiencies beyond 47 % [32].

1 .5 .3 Principal Scientific Challenges

The band-edge engineering of GaAsBi heterostructures requires systematic study. This work employs a 200 nm GaAs/200 nm GaAsBi bilayer and compares devices in which the Bi-containing layer is placed either at the top or bottom of the intrinsic region; Bi content is varied from 1 % to 3.8 % to examine the influence of interface barriers on carrier injection, internal field distribution, and breakdown characteristics.

Optimising multi-quantum-well (MQW) parameters is equally important. To balance strong absorption, low noise, and controllable breakdown, each GaAsBi quantum well is fixed at 10 nm, while the number of wells is stepped from 2 to 120. Increasing the period number enlarges the effective absorption volume and can markedly improve utilisation of > 1 μm photons; simultaneously, periodic barriers discretise the high-field region, potentially suppressing hole impact ionisation, increasing the electron–hole ionisation-coefficient ratio (α/β), reducing multiplication noise, redistributing the internal field, delaying avalanche breakdown, and lowering dark current.

Finally, experimental data on α/β in GaAsBi remain scarce. By combining pure-injection multiplication and excess-noise measurements, this study quantifies how Bi content, Bi-layer placement, and MQW period number jointly tailor α/β, providing concrete guidelines for low-noise GaAsBi avalanche photodiodes (APDs).

1 .6 Thesis Organisation

This dissertation comprises seven chapters that progress from theoretical foundations through material growth and device design to performance evaluation and synthesis.

Chapter 1 introduces the background, motivation, and objectives of GaAsBi research.

Chapter 2 reviews the underlying theory: III–V band structures, impact-ionisation mechanisms, α/β and multiplication noise models, and the effects of alloying and strain on band alignment.

Chapter 3 details MBE growth procedures for GaAs and GaAsBi, emphasising high-purity sources, ultra-high-vacuum maintenance, and in-situ RHEED diagnostics, and discusses Bi incorporation kinetics and defect-suppression strategies.

Chapter 4 describes the optical, electrical, and structural characterisation methods—photoluminescence, X-ray diffraction, atomic-force microscopy, pure-injection multiplication, and excess-noise measurements—together with unified data-analysis protocols.

Chapter 5 focuses on 200 nm GaAs/200 nm GaAsBi heterojunction photodiodes. By varying Bi-layer placement and Bi content, the chapter quantifies how interface barriers modify breakdown voltage, dark current, and α/β, clarifying the role of band-edge engineering in noise suppression.

Chapter 6 investigates GaAs/GaAsBi MQW photodiodes. With a fixed 10 nm well thickness and 2–120 periods, it elucidates the cooperative effects of quantum confinement and field redistribution on absorption, avalanche gain, α/β, and breakdown, yielding design rules for low-noise and delayed-breakdown structures.

Chapter 7 synthesises the comparative merits and limitations of heterojunction and MQW devices, summarises optimum ranges for Bi content, Bi-layer placement, and well number, and outlines future directions toward NIR detectors, thermally stable lasers, and broadband THz sources based on GaAsBi.

Reference

[1] K. Oe and H. Okamoto, "New Semiconductor Alloy $GaAs_{1-x}Bi_x$ Grown by Metal-Organic Vapor Phase Epitaxy," *Jpn. J. Appl. Phys.*, vol. 37, Part 2, no. 11A, pp. L1283–L1285, Nov. 1998, doi: 10.1143/JJAP.37.L1283.

[2] M. Weyers, M. Sato, and H. Ando, "Red Shift of Photoluminescence and Absorption in Dilute GaAsN Alloy Layers," *Jpn. J. Appl. Phys.*, vol. 31, Part 2, no. 7A, pp. L853–L855, Jul. 1992, doi: 10.1143/JJAP.31.L853.

[3] "Photomultiplier Tubes: Basics and Application Third Edition," Hamamatsu Photonics K. K.

[4] R. B. Emmons, "Avalanche photodiode frequency response," J. Appl. Phys., vol. 38, no. 9, pp. 3705–3714, 1967.

[5] R.B. Emmons and G. Lucovsky, "The frequency response of avalanching photodiodes,"IEEE Trans. Electron Devices, vol. 13, pp. 297-305, 1966.

[6] R. J. McIntyre, "Multiplication noise in avalanche photodiodes", IEEE Trans. Electron Devices, vol. 13, no. 1, pp. 164–168, Jan. 1966.

[7] M. M. Hayat, B. E. A. Saleh, and M. C. 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.

[8] X.W. Li,X.G. Zheng, S. L.Wang, F. Ma, and J. C. Campbell, "Calculation of gain and noise with dead space for GaAs and AlxGa1-xAs avalanche photodiode," IEEE Trans. Electron Devices, vol. 49, no. 7, pp. 1112–1117, Jul. 2002.

[9] J. C. Campbell, S. Chandrasekhar, W. T. Tsang, G. J. Qua, and B. C. Johnson, "Multiplication noise of wide-bandwidth InP/InGaAsP/InGaAs Avalanche Photodiodes," Journal of Lightwave Technology, vol. 7, no. 3, pp. 473–478, Mar. 1989. doi:10.1109/50.16883.

[10] P. Yuan et al., "Impact ionization characteristics of III-V semiconductors for a wide range of multiplication region thicknesses," IEEE J. Quantum Electron, vol. 36, no. 2, pp. 198–204, Feb. 2000.

[11] M. A. Saleh et al., "Impact-ionization and noise characteristics of thin III–V avalanche photodiodes," IEEE Trans. Electron Devices, vol. 48, no. 12, pp. 2722–2731, Dec. 2001.

[12] P. Yuan et al., "A new look at impact ionization—Part II: Gain and noise in short

avalanche photodiodes," IEEE Trans. Electron Devices, vol. 46, no. 8, pp. 1632–1639, Aug. 1999.
[13] P. Yuan, S. Wang, X. Sun, X. G. Zheng, A. L. Holmes, and J. C. Campbell, "Avalanche photodiodes with an impact-ionization-engineered multiplication region," IEEE Photon. Technol. Lett., vol. 12, no. 10, pp. 1370–1372, Oct. 2000.
[14] O. H. Kwon et al., "Optimal excess noise reduction in thin heterojunction Al0.6Ga0.4As-GaAs avalanche photodiodes," IEEE J. Quantum Electron, vol. 39, no. 10, pp. 1287–1296, Oct. 2003.
[15] M.M. Hayat, O.H. Kwon, S.L. Wang, J.C. Campbell, B.E. A. Saleh, and M.C. Teich, "Boundary effects on multiplication noise in thin heterostructure avalanche photodiodes: Theory and experiment," IEEE Trans. Electron Devices, vol. 49, no. 12, pp. 2114–2123, Dec.2002.
[16] S. Wang et al., "Low-noise impact-ionization-engineered avalanche photodiodes grown on InP substrates," IEEE Photon. Technol. Lett., vol. 14, no. 12, pp. 1722–1724, Dec. 2002.
[17] N. Duan et al., "High-speed and low-noise SACM avalanche photodiodes with an impact-ionization-engineered multiplication region," IEEE Photon. Technol. Lett., vol. 17, no. 8, pp. 1719–1721, Aug. 2005.
[18] D. Cannon, J. Liu, Y. Ishikawa, K. Wada, D. T. Danielson, S. Jongthammanurak, J. Michel, and L. C. Kimerling, "Tensile strained epitaxial Ge films on Si(100) substrates with potential application in L-band telecommunications," Appl. Phys. Lett., 84, 906-908 (2004).
[19] G. Dehlinger and S.J. Koester and J. D. Schaub and J. O. Chu and Q.C. Ouyang and A. Grill, "High-Speed Germanium-on-SOI Lateral PIN photodiodes," IEEE Photon. Technol. Lett. 16, 2547-2549 (2004).
[20] J. Liu, J. Michel, W. Giziewicz, D. Pan, K. Wada, D. Cannon, L.C. Kimerling, J. Chen, F. O. Ilday, F. X. Kartner and J. Yasaitis, "High-performance, tensile-strained Ge p-i-n photodetectors on a Si platform," Appl. Phys. Lett. 87, 103501 (2005).
[21] A.C. Farrell *et al.*, "Plasmonic field confinement for separate absorption-multiplication in InGaAs nanopillar avalanche photodiodes," *Sci. Rep.*, vol. 5, pp. 17580, 2015.

[22] J.S. Ng, C.H. Tan, J.P.R. David, G. Hill and G.J. Rees, "Field dependence of InGaAs impact ionization coefficients," *IEEE Trans. Electron Devices*, vol. 50, no. 4, pp. 901–905, Apr. 2003.

[23] M.E. Woodson, M. Ren, S.J. Maddox, Y. Chen, S. Bank and J. Campbell, "Low-noise AlInAsSb avalanche photodiode," *Appl. Phys. Lett.*, vol. 108, Feb. 2016.

[24] Y. Liu *et al.*, "Valence band engineering of GaAsBi for low-noise avalanche photodiodes," *Nature Communications*, vol. 12, no. 1, Aug. 2021, doi:10.1038/s41467-021-24966-0.

[25] R. Lewis *et al.*, "GaAsBi LEDs emitting at 987 nm with 1.8 % Bi," 2009.

[26] R. Richards *et al.*, "Room-temperature 1.2 µm GaAsBi LEDs with 6 % Bi," 2010.

[27] P. Ludewig *et al.*, "Electrically pumped 2.2 % Bi GaAsBi single-QW laser at 947 nm," 2013.

[28] M. Yoshimoto *et al.*, "Room-temperature GaAs/GaAsBi/GaAs Fabry–Pérot laser at 982.8 nm," *Appl. Phys. Lett.*, 2010.

[29] M. Yoshimoto *et al.*, "GaAsBi laser diode emitting at 1.204 µm with 5.8 % Bi," 2013.

[30] K. Bertulis *et al.*, "GaAsBi photoconductive THz emitters and detectors with sub-picosecond lifetimes," 2006.

[31] K. Bertulis *et al.*, "Enhanced 4.5 THz emission from LT-GaAsBi antennas," 2009.

[32] I. P. Marko and S. J. Sweeney, "GaAsBi for a 1 eV sub-cell in high-efficiency multi-junction photovoltaics," 2017.

[33] J. Liu et al., "Reducing Afterpulsing in InGaAs(P) Single-Photon Detectors with Hybrid Quenching," Sensors (Basel), vol. 20, no. 16, 4384, Aug. 2020, doi:10.3390/s20164384.

[34] Y. Kang et al., "Monolithic germanium/silicon avalanche photodiodes with 340 GHz gain–bandwidth product," Nature Photonics, vol. 3, pp. 59–63, 2009, doi:10.1038/nphoton.2008.247.

[35] D. J. Eaglesham, "Semiconductor molecular-beam epitaxy at low temperatures," Journal of Applied Physics, vol. 77, no. 8, pp. 3597–3617, Apr. 1995, doi:10.1063/1.358597.

Chapter 2: Theoretical Background

2.1 PN junction

Figure 1 schematically depicts the essential physics of a p–n junction by means of conduction- and valence-band diagrams under three operating conditions. Figure 1 (a) shows the junction at thermal equilibrium. The p-type region on the left is rich in holes (open circles), whereas the n-type region on the right contains a high electron concentration (solid circles). Carrier diffusion across the metallurgical interface establishes a space-charge region devoid of mobile carriers; the resulting internal electric field bends the energy bands upward in the p-layer and downward in the n-layer until the Fermi level (dashed line) becomes flat. This built-in potential counterbalances further diffusion and defines the depletion width.

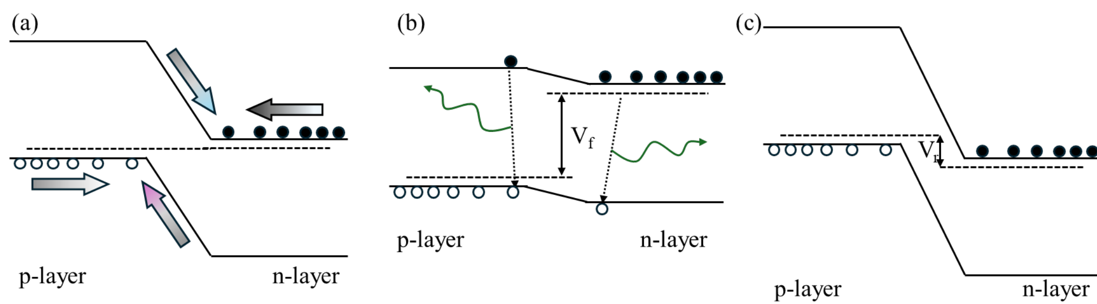


Figure 2.1 Band diagrams of a p–n junction at equilibrium (a), forward bias(b), and reverse bias(c).

Figure 2.1 (b) illustrates forward bias, where an external voltage V_f is applied so that the p-side is raised and the n-side lowered with respect to equilibrium. The applied field partially cancels the built-in field, narrowing the depletion region and reducing the potential barrier by V_f. Electrons injected from the n-layer and holes injected from the p-layer overcome the diminished barrier and enter the opposite quasi-neutral regions as minority carriers, indicated by the curved green arrows recombining within the bulk. The exponential growth of this injection current with V_f underpins the rectifying behaviour captured by the Shockley diode equation.

Figure 2.1 (c) represents reverse bias, in which the external voltage V_r augments the built-in potential and widens the depletion width. The greater band bending elevates the

barrier for majority carriers, effectively suppressing their flow so that only a small leakage current—due to thermally generated minority carriers or surface states—persists. At sufficiently high V_r the electric field becomes large enough to trigger either Zener tunnelling or avalanche multiplication, leading to breakdown.

Together, the three panels underscore how the interplay between diffusion, drift, and recombination in the depletion region governs current–voltage characteristics, capacitance, and breakdown phenomena, making the p–n junction the foundational element of diodes, transistors, solar cells, and myriad optoelectronic devices.

2.2 Absorption of light

The interaction of photons with a semiconductor is governed by the material's band-gap energy E_{g}. When the photon energy $h\nu$ exceeds E_{g}, electrons can be excited from the valence band to the conduction band, creating electron–hole pairs that underpin photodetection and photovoltaic conversion. The longest wavelength capable of initiating this interband transition is termed the cut-off wavelength λ_{c} and is obtained from the energy–wavelength relation

$$\lambda_{\mathrm{c}} = \frac{hc}{E_{\mathrm{g}}} \approx \frac{1.24}{E_{\mathrm{g}}}\ (\mu\mathrm{m}), \tag{2.1}$$

where h is Planck's constant and c is the speed of light. Photons with $\lambda > \lambda_{\mathrm{c}}$ possess insufficient energy and therefore propagate through the crystal without generating carriers.

Once $h\nu \geq E_{\mathrm{g}}$, the rate at which optical power diminishes inside the semiconductor is quantified by the absorption coefficient γ. A large γ signifies that carriers are generated within a shallow depth—an essential attribute for thin-film photodiodes and surface-emitting lasers—whereas a small γ implies that the photon flux penetrates deeply before being absorbed, an advantage for bulk solar cells that seek volumetric generation. The spatial decay of the optical intensity $\varphi(x)$ follows Beer's law,

$$\varphi(x) = \varphi_0 \exp(-\gamma x), \tag{2.2}$$

with φ_0 the incident intensity and x the propagation distance. Although Eq. (2.1) prescribes the long-wavelength limit of absorption, a practical short-wavelength cut-off can also arise: in highly excited semiconductors, rapid carrier recombination, intervalence-band absorption, or surface recombination may curtail the collection of photogenerated carriers at photon energies well above E_g. Consequently, both the intrinsic band structure and extrinsic recombination dynamics must be considered when designing optoelectronic devices for a specific spectral window.

Figure 2.2 compares calculated electronic band structures for GaAs (a) and Si (b) along high-symmetry directions of the first Brillouin zone (L–Λ–Γ–Δ–X–U, K–Σ–Γ). The vertical axis gives energy E in electron-volts, while the horizontal axis tracks crystal momentum k. In GaAs, the conduction-band minimum and valence-band maximum coincide at the Γ point (k = 0), yielding a direct band gap; electrons can therefore be excited vertically in k-space without phonon assistance, leading to a large optical matrix element and an absorption coefficient that exceeds 10^4 cm^{-1} within tens of meV above the gap. By contrast, the lowest conduction band in silicon occurs at the Δ valley near the X point, whereas the valence-band maximum remains at Γ. This indirect gap forces phonon participation to conserve momentum, suppressing the transition probability and reducing the near-edge absorption coefficient by two to three orders of magnitude. The diagrams also reveal the larger density of conduction-band valleys in Si, which affects carrier mobility anisotropy and intervalley scattering, while GaAs exhibits a lighter Γ_6 conduction band that underpins its high electron mobility. Consequently, the direct-gap topology of GaAs favours efficient light emission and detection, whereas Si's indirect character necessitates thicker absorbers and limits radiative recombination efficiency.

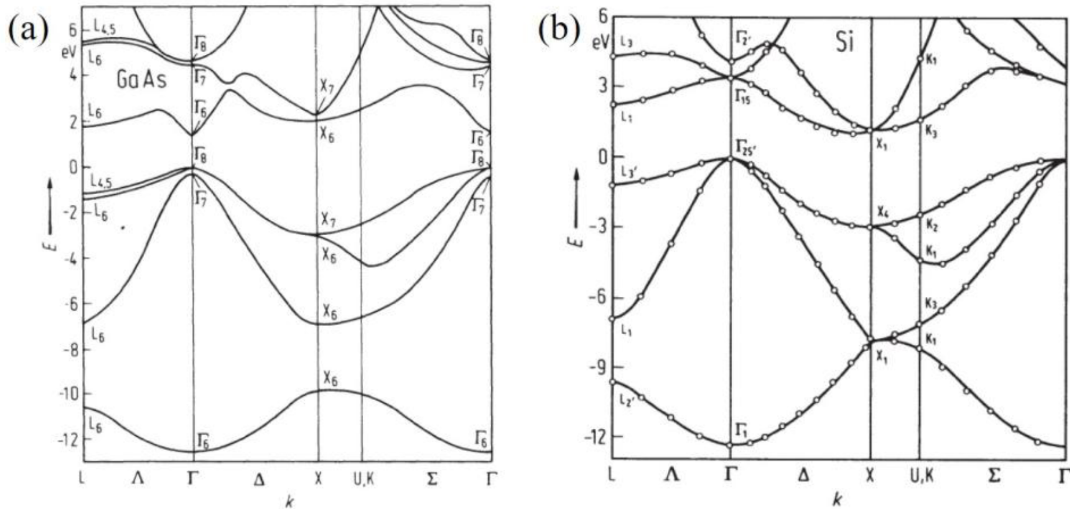


Figure 2.2 Band structure of (a) GaAs (b) Si [1].

In a direct-band-gap semiconductor the conduction-band minimum and valence-band maximum reside at the same crystal momentum (typically the Γ point). When a photon with energy $h\nu \geq E_g$ is absorbed, an electron can be promoted vertically in k-space without violating momentum conservation because the photon's momentum is negligibly small. The transition probability is governed only by the dipole matrix element between the initial and final Bloch states, which is generally large for allowed direct transitions. As a result the optical absorption coefficient $\alpha(h\nu)$ rises steeply at the band edge, following $\alpha \propto (h\nu - E_g)^{1/2}$, and reaches values of 10^4–10^5 cm^{-1} within a few tens of millielectron-volts. Practically, this means that sub-micrometre films of GaAs, InP or GaN can absorb nearly all incident photons whose energies exceed their gap, enabling thin active regions in light-emitting diodes, laser diodes and high-speed photodiodes.

In an indirect-band-gap material the conduction-band minimum occurs at a momentum offset from the valence-band maximum, as seen for silicon in Figure 2.2. Because a photon cannot supply the requisite crystal momentum, absorption must involve a

second quasi-particle—a phonon—to conserve total momentum. This phonon-assisted, second-order process possesses a much smaller matrix element, and the near-edge absorption coefficient is reduced by two to three orders of magnitude, typically 10^2–10^3 cm^{-1}. The spectral dependence becomes $\alpha \propto (h\nu - E_g \pm E_{\mathrm{ph}})^2$, reflecting the need to absorb or emit a phonon of energy E_{ph}. Efficient absorption therefore requires material thicknesses of tens of micrometres, which is why crystalline-silicon solar cells employ wafers hundreds of micrometres thick. Conversely, the low radiative recombination probability inherent to indirect gaps hampers silicon's usefulness in light-emitting applications unless augmented by nanostructuring or defect engineering.

Thus the key distinction is that direct gaps permit one-step, momentum-conserving optical transitions with high absorption and emission efficiency, while indirect gaps rely on phonon mediation, leading to much weaker absorption near the fundamental edge and intrinsically poor light emission.

2.3 Diode equation

The Shockley diode equation provides the canonical description of the current–voltage response of an abrupt p–n junction. Originating from an analytical solution of the coupled continuity and Poisson equations under low-level injection, the relation captures the combined effects of minority-carrier diffusion, recombination, and the thermally activated nature of carrier transport across the space-charge region. In its general form the terminal current I is expressed as

$$I = I_s \left(\mathrm{e}^{\frac{qV}{nkT}} - 1 \right), \tag{2.3}$$

where q is the elementary charge, k is Boltzmann's constant, T is the absolute temperature, V is the externally applied bias, and I_s designates the reverse-saturation current. The parameter n, often called the ideality or emission factor, quantifies

deviations from purely diffusive transport. For an ideal junction limited solely by carrier diffusion across the depletion edge, n approaches unity; when Shockley–Read–Hall recombination within the depletion region becomes appreciable, n climbs toward two. Values exceeding two usually signify additional series resistance or high-level injection effects.

The saturation current I_s is governed by the equilibrium minority-carrier concentrations on either side of the junction and by their respective diffusion lengths. Because these quantities depend exponentially on the band-gap energy, I_s is minuscule for wide-gap materials and rises sharply as the gap narrows. Temperature exerts an equally pronounced influence: an increase of 10 K can augment I_s by an order of magnitude in silicon diodes. At the standard reference temperature of 300 K, the thermal voltage $V_T = kT/q$ attains a value of approximately 25.85 mV; this small potential sets the natural voltage scale for exponential growth in equation (2.3).

When a positive bias is applied to the p-side (forward bias), the external field partially cancels the built-in potential, thereby contracting the depletion layer and lowering the potential barrier. Minority carriers injected across the junction recombine with majority carriers after diffusing into the quasi-neutral regions, and the ensuing current escalates exponentially with V. At moderate forward biases—typically above 0.1 V for silicon—the exponential term greatly exceeds unity and the -1 in Eq. (2.3) may be neglected, giving

$$I \approx I_s\, e^{\frac{qV}{nkT}}. \tag{2.4}$$

If series resistance is negligible, the current density at 0.7 V in a silicon diode can surpass the saturation current by more than nine decades, which explains the device's utility as a rectifying element. Below this "turn-on" voltage the current remains limited by the -1 term, ensuring that negligible conduction occurs when $V = 0$.

Under reverse bias the junction field is reinforced: the depletion region widens, majority

carriers are swept away from the interface, and the exponential in Eq. (2.3) tends toward zero. The net result is a nearly bias-independent leakage current equal to I_s. Only when the reverse voltage approaches the breakdown threshold do additional mechanisms—Zener tunnelling in highly doped devices and impact ionisation in more lightly doped ones—overwhelm the diffusion current, at which point Eq. (2.3) ceases to apply. Prior to breakdown, however, the constancy of I_s with respect to voltage makes the Shockley equation a reliable predictor of leakage, an attribute critical for analogue circuit design and detector noise analysis.

The exponential dependence in Eqs. (2.3) and (2.4) is rooted in the Maxwell–Boltzmann distribution of carrier energies. Forward bias flattens the quasi-Fermi-level separation across the junction, allowing electrons from the n-type side and holes from the p-type side to surmount the diminished barrier; in reverse bias, the enlarged barrier repels these majority carriers while thermally generated minority carriers are swept across, giving rise to I_s. Because I_s scales with the product of the intrinsic carrier concentration n_i^2 and diffusion coefficients, both of which increase with temperature, the Shockley relation intrinsically links thermal effects to observable diode behaviour. Consequently, temperature-compensated circuit topologies and precision instrumentation frequently exploit V_T and $I_s(T)$ to achieve accurate biasing or sensing functions.

Although additional complexities—series resistance, high-level injection, and generation–recombination noise—can necessitate refined models, equation (2.3) remains the cornerstone of diode theory. Its compact analytical form distils the essence of p–n junction transport and continues to inform the design of modern power rectifiers, radio-frequency detectors, and photonic transceivers.

2.4 Introduction of GaAs and GaAsBi

2.4.1 introduction of GaAs

Gallium arsenide (GaAs) is a III-V semiconductor with a direct band gap of 1.424 eV

at room temperature and a lattice constant of 5.6533 Å (at 300 K) [3]. It has a zinc blende crystal structure, as illustrated in Figure 2.3, which is equivalent to two face-centered cubic (FCC) sub-lattices: one composed of arsenic (As) atoms and the other of gallium (Ga) atoms [3]. These sub-lattices are superimposed with a relative displacement of 5.6533 Å / 4 in the (001), (010), and (100) unit directions [3].

In a unit cell diagram, this structure appears as a single FCC lattice with four tetrahedral interstices. Although the unit cell shows more of the light grey group in the structure, the corner and face-centered atoms are shared among multiple cells. Each unit cell actually contains four As atoms and four Ga atoms.

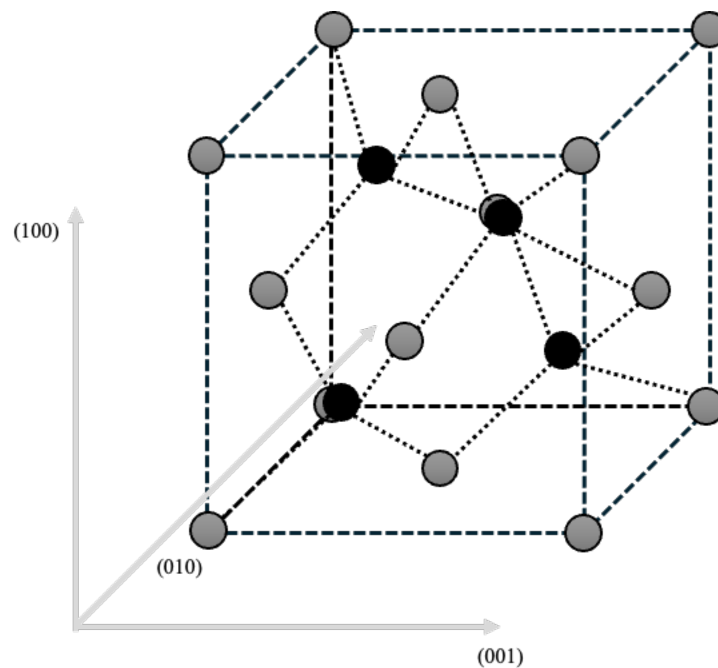


Figure 2.3: Zinc blende unit cell. FCC lattice (light grey) and tetrahedral interstices (black).

2.4.2 Introduction of GaAsBi

Gallium-arsenide bismide (GaAsBi) is produced when a dilute concentration of bismuth substitutes for arsenic on the anion lattice of GaAs. The large atomic-size and

electronegativity mismatch between Bi and As introduces resonant impurity states that interact strongly with the GaAs valence band, driving a pronounced and nearly linear reduction of the fundamental band gap while preserving close lattice matching to the GaAs substrate [3]. By judiciously adjusting the Bi mole fraction, this gap can be engineered to cover the critical 1.3 µm and 1.55 µm wavelength windows that dominate contemporary fibre-optic infrastructure. The same tunability suggests the use of GaAsBi as a low-band-gap junction in multi-junction photovoltaic architectures, where its incorporation has the potential to improve current matching and to raise the overall conversion efficiency. An additional electronic consequence of Bi incorporation is a sizeable downward shift of the spin-orbit split-off band, which may render the spin-orbit splitting energy Δ_{SO} larger than the band gap. In this regime, Auger recombination and inter-valence-band absorption become energetically unfavourable, a feature that is especially advantageous for long-wavelength lasers and other telecommunication-grade optoelectronic devices [4].

2.4.3 Valence-Band Anti-Crossing (VBAC) Model

The valence-band anti-crossing (VBAC) model provides a physically transparent description of the unusually strong bandgap bowing observed in highly mismatched III–V alloys. In GaAsBi, substitutional Bi on the group-V sublattice introduces Bi-derived, p-like localized (resonant) states whose energy lies close to the GaAs valence-band edge. Because these localized Bi states have the same symmetry character as the extended valence-band states near the Brillouin-zone centre, they can couple (hybridise). The coupling produces level repulsion and splits the original valence band into two hybridised branches: an upper E_+ branch and a lower E_- branch (analogous to the formation of hybrid modes in two coupled oscillators or coupled optical waveguides). The upper branch E_+ forms the new valence-band maximum, thereby shifting the valence-band edge upward and reducing the band gap. [19, 20]

Within the simplest two-level VBAC picture, the hybridised energies can be written as

$$E_{\pm}(k) = \frac{E_v^{\mathrm{GaAs}}(k) + E_{Bi}}{2} \pm \frac{1}{2}\sqrt{\left(E_v^{\mathrm{GaAs}}(k) - E_{Bi}\right)^2 + 4xC_{Bi}^2} \tag{2.5}$$

where E_{Bi} is the energy of the Bi-related resonant level (referenced to the GaAs valence-band edge), x is the Bi composition, and C_{Bi} quantifies the coupling strength between the localized Bi states and the extended GaAs valence-band states. Near the zone centre, a parabolic approximation for the host valence band is

$$E_v^{\mathrm{GaAs}}(k) = E_v^{\mathrm{GaAs}}(0) - \frac{\hbar^2 k^2}{2m_h^*} \tag{2.6}$$

so that even small Bi fractions can produce a comparatively large upward shift of the valence-band edge through the $\sqrt{(E_v - E_{Bi})^2 + 4xC_{Bi}^2}$ coupling term. In more complete multiband descriptions, the same anti-crossing physics also modifies the split-off band and can increase Δ_{SO}, which is directly relevant to suppressing hole-initiated impact ionisation and reducing avalanche excess noise in GaAsBi-based APDs. [20]

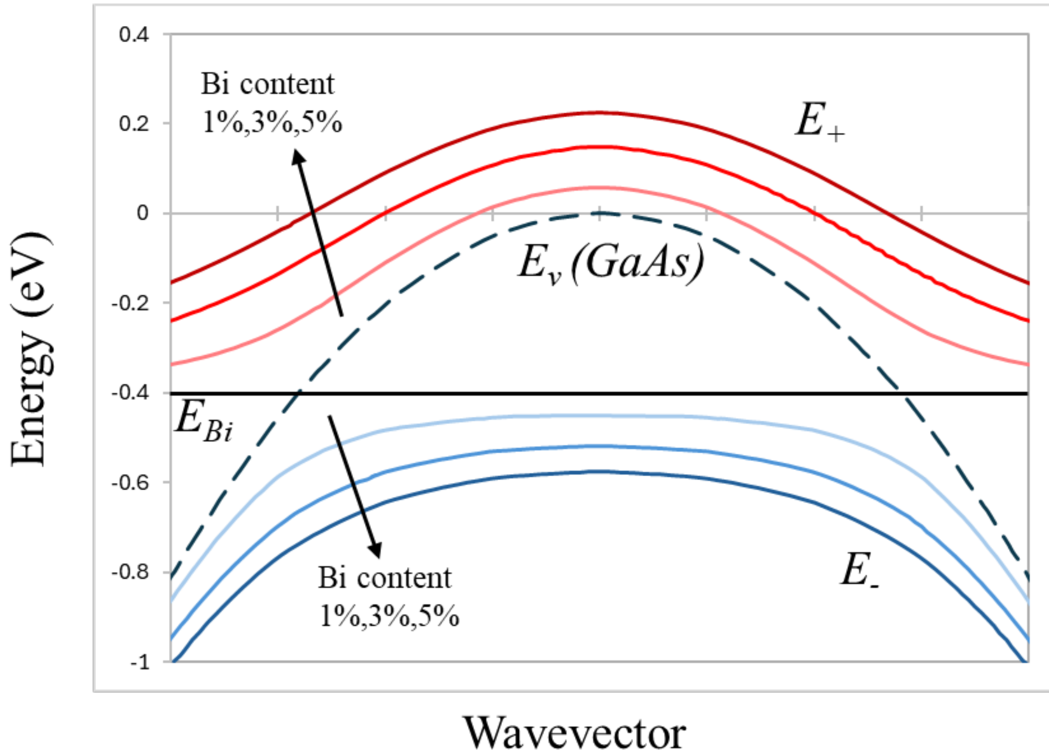


Figure 2.4 Calculated valence-band structure of GaAsBi for Bi concentrations of 1 %, 3 % and 5 % using the VBAC model.

2.4.4 Growth of GaAsBi

The epitaxial growth of GaAsBi is markedly more restrictive than that of GaAs because the Ga–Bi pair exhibits both a large miscibility gap and a comparatively weak bond energy. At the substrate temperatures typically used for GaAs (≈ 580 °C) bismuth desorbs before it can be incorporated, so high-quality GaAsBi layers have been realised by molecular-beam epitaxy only at or below 400 °C [6]–[8]. Significant Bi incorporation also hinges on supplying arsenic in near-stoichiometric proportions, either as As_4 sublimed from elemental arsenic or as As_2 produced by thermal cracking. Experiments show that the optimum arsenic flux differs for these species because of their distinct sticking probabilities; when this is accounted for, both precursors can deliver an effective As:Ga ratio close to unity [9]. On the GaAs(001) surface, however, arsenic uptake is curtailed by the need for two adjacent ad-atoms to form a stable As_2 dimer; a fraction of impinging As_4 therefore desorbs intact, reducing the net incorporation rate [10, 11]. Because Bi competes with As for surface sites, lower arsenic coverage increases the Bi residence time, whereas excess arsenic promotes Bi displacement.

Two kinetic descriptions capture these trends while differing in the number of explicit reaction channels. Lu and co-workers outline three surface processes—Ga bonding to Bi-terminated sites, Ga bonding between two Bi atoms, and As atoms replacing Ga–Bi bonds—and emphasise the competition between Bi incorporation and As-induced bond exchange (Figure 2.5) [7]. Lewis et al. adopt a simplified picture that neglects the energetically unfavourable Ga-to-Bi bonding pathway and describe the evolution of Bi coverage solely in terms of the relative surface densities of Bi and Ga (Figure 2.6) [6]. In this formulation the steady-state Bi content is controlled primarily by the As:Ga flux ratio and the growth temperature, which together govern the rates of Bi desorption and As-driven bond replacement. Although neither model fully reproduces layers grown under extreme arsenic-deficient conditions, both indicate the practical route to higher

Bi fractions: suppress desorption by lowering the temperature and maintain a carefully balanced arsenic flux to avoid Bi displacement.

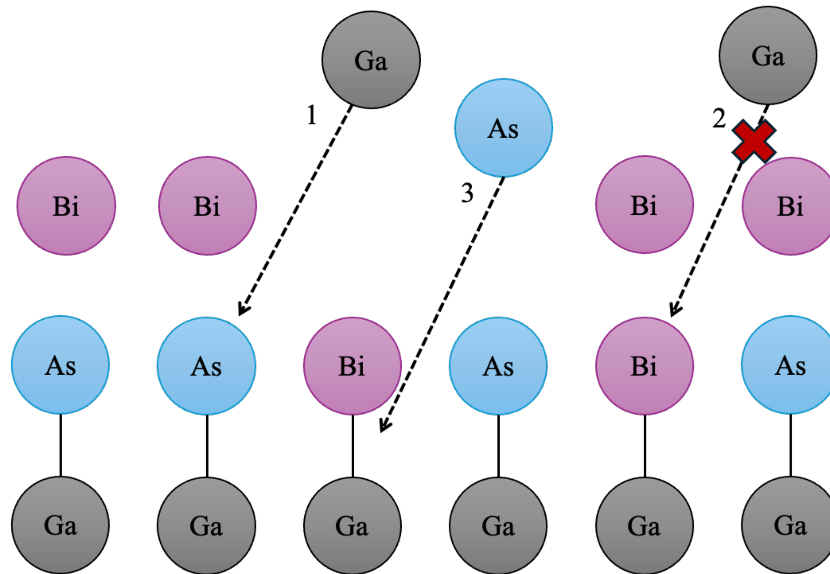


Figure 2.5: Lu model. Process 1 illustrates how Bi–Ga bonds are formed, increasing the Bi content at the growing surface. Process 3 shows how these bonds are replaced by stronger As–Ga bonds, reducing the Bi content. Adapted from [7].

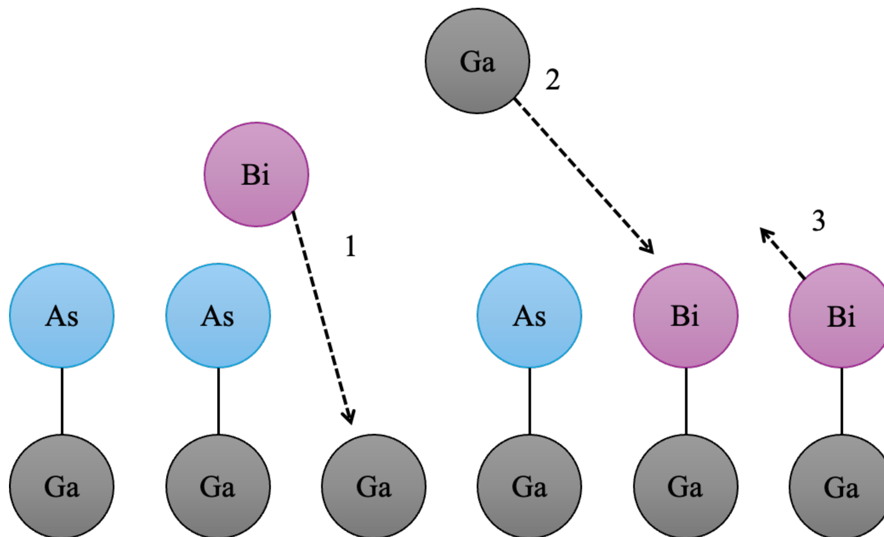


Figure 2.6: Lewis model. Process 1 shows Bi bonding to an exposed Ga site, increasing Bi surface coverage and reducing Ga surface coverage. Process 2 shows a Ga atom bonding to a Bi site as part of the growth process, decreasing Bi surface

coverage. Process 3 shows a Bi atom being thermally ejected to the surface, further reducing Bi surface coverage. Adapted from [6].

2.5 Avalanche multiplication and Impact ionization coefficient

Avalanche multiplication is the carrier-amplification mechanism that underpins both the high internal gain of avalanche photodiodes (APDs) and the breakdown behaviour of Zener or avalanche diodes. When a primary electron or hole is injected into the depleted high-field region of a reverse-biased p-i-n (or p-n) junction, it is accelerated by the local electric field; if the field is sufficiently large, the carrier's kinetic energy after a mean free path exceeds the ionisation threshold and a collision can liberate a valence electron, thereby creating a secondary electron--hole pair. The original and secondary carriers are re-accelerated and may each initiate further ionisation events, giving rise to the branching trajectory sketched in Fig. 2.7(a).

Figure 2.7(b) provides a corresponding reverse-bias band-diagram picture (cf. the reverse-bias band bending introduced in Fig. 2.1(c)): the tilted band edges across the depleted region indicate the strong electric field. Under this field, electrons drift opposite to the field direction towards the n-contact, while holes drift with the field towards the p-contact. Throughout Eqs. (2.7)-(2.9), the coordinate is taken such that $x = 0$ and $x = w$ denote the two edges of the multiplication region of width w. Consistent with the boundary conditions used below, the electron-initiated mean multiplication is $M_e = M(0)$, whereas the hole-initiated mean multiplication is $M_h = M(w)$.

At the microscopic level, an impact-ionisation event requires the initiating carrier to acquire a threshold energy, denoted generically as E_{th} in Fig. 2.7(c). In the dead-space formulation introduced later, this threshold is written more explicitly as $E_{\mathrm{th},e}$ and $E_{\mathrm{th},h}$ for electrons and holes, respectively (see Eqs. (2.14)-(2.15)).

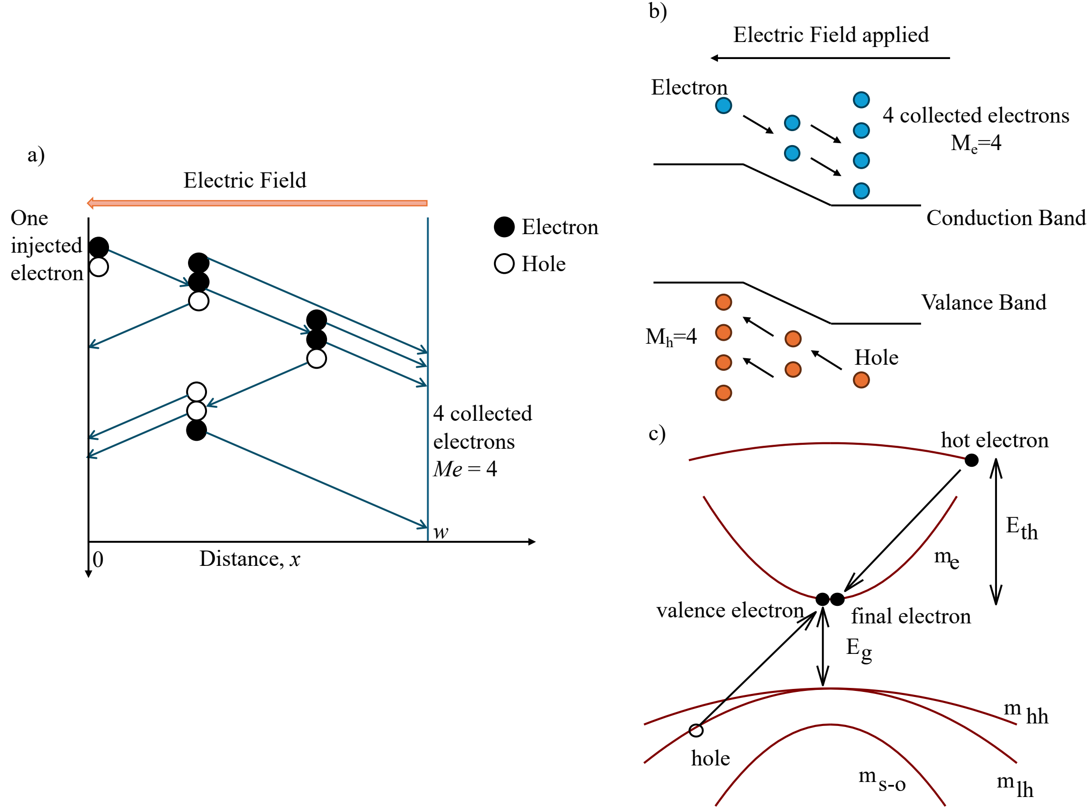


Figure 2.7: Schematic illustration of impact ionisation and avalanche multiplication in a reverse-biased p-i-n junction. (a) Branching of an electron-initiated avalanche within a multiplication region of width w (coordinate x). (b) Reverse-bias band diagram indicating carrier drift: electrons drift towards the n-contact (opposite the field), whereas holes drift towards the p-contact (with the field), motivating the definitions of M_e and M_h. (c) Energy-band picture of an impact-ionisation event showing the band gap E_g and the threshold energy E_{th}.

Stillman and Wolfe employed the carrier-continuity equation to derive an analytical expression for the mean gain, assuming that the ionisation probability depends only on the instantaneous electric field and that carriers drift ballistically between collisions [14]. For a depletion region of width w the mean multiplication for a carrier injected at a depth x is

$$M(x) = \frac{\exp\left[-\int_x^w (\alpha(x') - \beta(x'))\,\mathrm{d}x'\right]}{1 - \int_0^w \alpha(x')\exp\left[-\int_0^{x'} (\alpha(x'') - \beta(x''))\,\mathrm{d}x''\right]\mathrm{d}x'}, \qquad (2.7)$$

where α and β are the electron- and hole-ionisation coefficients, respectively. Setting $x = 0$ yields

$$M_{\mathrm{e}} = M(0) = \frac{1}{1 - \int_0^w \alpha(x') \exp\left[- \int_0^{x'} (\alpha(x'') - \beta(x''))\, \mathrm{d}x''\right] \mathrm{d}x'}, \quad (2.8)$$

whereas a hole injected at $x = w$ gives

$$M_{\mathrm{h}} = M(w) = \frac{1}{1 - \int_0^w \alpha(x') \exp\left[- \int_{x'}^{w} (\alpha(x'') - \beta(x''))\, \mathrm{d}x''\right] \mathrm{d}x'}. \quad (2.9)$$

If the electric field E is assumed spatially uniform, α and β are constant and the integrals integrate directly to

$$M_{\mathrm{e}} = \frac{1}{1 - \frac{\alpha}{\alpha - \beta}\{\exp\left[(\beta - \alpha)w\right] - 1\}}, M_{\mathrm{h}} = \frac{1}{1 - \frac{\beta}{\beta - \alpha}\{\exp\left[(\alpha - \beta)w\right] - 1\}}. \quad (2.10)$$

Equations (2.8)–(2.10) show that the excess noise is minimised when only one carrier type predominates in the ionisation process, because the statistical variance of the multiplication chain is then lowest. Silicon exemplifies this desirable condition: for reverse fields typically employed in APDs $\alpha \gg \beta$, so electron-initiated multiplication dominates and the noise factor $F(M)$ remains modest [14]. In direct-gap III–V materials such as GaAs or GaAsBi, by contrast, the magnitudes of α and β are more comparable (see Figure 2.2, which contrasts the band structures of GaAs and Si and foreshadows their different ionisation behaviour), leading to larger excess-noise coefficients but also enabling higher gain at a given breakdown voltage.

The ionisation coefficients themselves depend sensitively on the electric field, temperature and material band structure. Chynoweth demonstrated empirically that α and β follow exponential field dependencies of the form [15]

$$\alpha = A_n \exp\left[-\left(\frac{B_n}{E}\right)^{c_n}\right], \beta = A_p \exp\left[-\left(\frac{B_p}{E}\right)^{c_p}\right], \quad (2.11)$$

where $A_{n,p}$, $B_{n,p}$ and $c_{n,p}$ are fitting parameters determined for a given semiconductor. Narrow-gap materials, in which carriers require less kinetic energy to surmount the band gap, show larger A and smaller B, yielding higher ionisation rates at modest fields; wide-gap semiconductors display the opposite trend, a property that renders them attractive for high-power electronics where avalanche breakdown must be avoided. Temperature modifies the ionisation rates through two competing mechanisms: increased lattice scattering shortens mean free paths, yet the elevated phonon population lowers the threshold energy needed for impact ionisation; the resultant non-monotonic variation mandates careful thermal design of APD receivers and high-voltage regulators alike.

Material choice thus dictates not only the attainable gain but also the excess-noise characteristics and breakdown voltage. GaAsBi, for instance, inherits the sizeable spin-orbit splitting of bismuth-containing alloys, which can suppress Auger recombination and enable low-noise gain at telecommunication wavelengths, although its α/β ratio remains closer to unity than that of silicon [13]. Accurate values of $\alpha(E,T)$ and $\beta(E,T)$ are normally extracted from time-of-flight or photomultiplication experiments and catalogued in compendia of III–V band parameters [5]; semi-empirical models interpolated from such data are then employed in device-level simulations. As advanced materials such as GaN, SiC or GeSn enter high-speed imaging and power-conversion markets, refinements of Chynoweth-type expressions, possibly augmented by full-band Monte-Carlo calculations, are required to maintain predictive fidelity.

Avalanche multiplication ceases once carriers exit the high-field layer or dissipate their energy through phonon emission. In reverse-biased p–n junctions not intended for multiplication, this termination coincides with the onset of breakdown; uncontrolled current then damages the device unless current-limiting circuitry is in place. Conversely, voltage-regulator diodes exploit the stability of avalanche conduction above breakdown to clamp potentials in power supplies, whereas APDs deliberately operate *below* breakdown so that M remains finite and excess noise controllable. The design trade-

off in every application is therefore a delicate balance among gain, noise, response time and reliability, all of which trace back to the fundamental impact-ionisation coefficients defined in Eq. (2.11).

2.6 Random Path Length (RPL) model

Full-band Monte-Carlo techniques give the most faithful account of carrier transport and scattering in high electric fields, yet their computational cost hampers iterative device optimisation. To circumvent this limitation Hayat and co-workers introduced a non-local recursion formalism in which the probability-density function (PDF) of the ionisation path length replaces explicit trajectory sampling, allowing accurate predictions of gain and excess noise in multiplication regions as thin as 0.1 μm [16]. Building upon that framework, Ong *et al.* proposed the Random-Path-Length (RPL) model, an efficient Monte-Carlo surrogate that draws individual ionisation distances from pre-defined PDFs; by repeatedly displacing the carrier according to these random lengths, the model reproduces the statistics of avalanche multiplication and its attendant noise with negligible loss of accuracy relative to full Monte-Carlo simulations [10]. Subsequent comparisons show that the recursive approach of Hayat and the RPL method of Ong converge to virtually identical mean gains and noise factors for modern III-V APDs, validating the physical equivalence of the two descriptions [17], [18].

In the RPL picture a carrier injected at x_0 travels a stochastic distance x before its first impact ionisation; the PDF $h(x)$ therefore embodies the non-local ionisation probability within a uniform field. For electrons and holes the PDFs are expressed as

$$h_e(x) = \begin{cases} 0, & x < d_e \\ \alpha^* \exp\left[-\alpha^*(x - d_e)\right], & x \geq d_e \end{cases}, h_h(x) = \begin{cases} 0, & x < d_h \\ \beta^* \exp\left[-\beta^*(x - d_h)\right], & x \geq d_h \end{cases}, (2.12 - 2.13)$$

where α^* and β^* are effective ionisation coefficients for electrons and holes, respectively, and d_e and d_h denote the "dead space"—the minimum drift length required for a carrier to acquire the threshold ionisation energy. Within a constant field

ξ these dead spaces are given by

$$d_e = \frac{E_{\text{th,e}}}{\xi}, d_h = \frac{E_{\text{th,h}}}{\xi}, \qquad (2.14-2.15)$$

with $E_{\text{th,e}}$ and $E_{\text{th,h}}$ the threshold energies for electron and hole ionisation, respectively. Equations (2.12)–(2.15) imply that no ionisation can occur until the carrier has traversed at least one dead-space length; once that condition is satisfied, the probability of ionisation decays exponentially with path length, mirroring the memoryless nature of successive scattering events in a homogeneous field.

Implementation of the RPL model proceeds by sampling an ionisation length from h_e or h_h, advancing the carrier by that distance, and spawning secondary electron–hole pairs at the terminal point. Each progeny is then subjected to the same stochastic propagation rule until all carriers exit the high-field layer or recombine. Repeating this procedure for a large ensemble yields the full distribution of multiplication factors M and its first and second moments, from which the excess-noise factor $F(M)$ follows. Because the PDFs in Eqs. (2.12)–(2.13) depend explicitly on the local field through $d_{e,h}$ and implicitly through α^*, β^*, the RPL engine naturally captures field-profile effects, allowing one to evaluate thin APDs with engineered field gradients without resorting to computationally intensive particle tracking.

Comparative studies reveal that the inclusion of dead space suppresses both the mean gain and its variance relative to classical local-field models, an effect that becomes pronounced when the multiplication width approaches a few dead-space lengths—exactly the scale of modern thin III-V photodiodes [17]. Consequently, circuit-level predictions that neglect dead space tend to overestimate responsivity and underestimate excess noise. The RPL formulation, while idealised by its assumption of uniform effective ionisation parameters, remains a valuable compromise between physical fidelity and computational economy, facilitating rapid exploration of device architectures and bias conditions in the design of next-generation low-noise, high-speed APDs.

2.7 Excess noise

Excess noise is an additional fluctuation superimposed on the intrinsic thermal and shot noise of a semiconductor device when carrier multiplication is present. In high-gain structures such as avalanche photodiodes (APDs), these fluctuations arise from the stochastic nature of impact ionisation: each electron or hole accelerated by a strong electric field acquires energy in a probabilistic manner, so the number, timing and spatial location of ionisation events vary from avalanche to avalanche. Because the final current is proportional to the total population of secondary carriers, any randomness in the multiplication chain—however small—broadens the output current distribution and thereby degrades the signal-to-noise ratio (SNR). A clear grasp of the origins and scaling of excess noise is therefore essential for semiconductor components used in optical communication, imaging and precision instrumentation, where faint signals must be distinguished from background fluctuations.

In an idealised device every primary carrier would generate exactly the same number of electron–hole pairs, yielding a fixed multiplication gain M. Reality departs from this ideal because the ionisation process is governed by probability: some carriers fail to reach the threshold energy, others ionise after widely differing path lengths, and secondary carriers do not necessarily replicate the trajectory of their progenitors. Consequently, the current collected after an avalanche event oscillates around an average value, the dispersion of which constitutes excess noise. The magnitude of that dispersion is influenced by three principal factors: the electron and hole impact-ionisation coefficients α and β; the ratio of those coefficients; and the mean multiplication gain itself. Where α and β differ markedly—so that one carrier type overwhelmingly initiates multiplication—the avalanche chain is more deterministic and the noise lower; when $\alpha \approx \beta$, both carrier species contribute almost equally and the variability increases.

A convenient quantitative measure is the excess-noise factor F, defined as the ratio of the total device noise to the noise that would exist if multiplication were noiseless. McIntyre showed that, for a wide class of uniform-field diodes, F can be written approximately as

$$F = kM + (1 - k)\left(2 - \frac{1}{M}\right), \tag{2.16}$$

where M is the average gain and

$$k = \frac{\beta}{\alpha}, \tag{2.17}$$

expresses the relative contribution of holes to electrons in the ionisation process. When k is close to unity the two carriers multiply with similar probability and F increases rapidly with M; when $k \ll 1$ or $k \gg 1$, indicating domination by a single carrier type, the rise in F with M is much slower, permitting higher gain at an acceptable noise penalty.

Material properties strongly modulate these parameters. Silicon, for example, exhibits an electron ionisation coefficient substantially larger than that for holes under typical operating fields, so k is small and silicon APDs achieve high SNR, a critical advantage in telecommunication receivers. III–V compounds such as GaAs or InP, which possess more balanced coefficients, incur higher excess noise but extend spectral sensitivity to longer wavelengths where silicon becomes opaque. The device designer must therefore trade excess-noise performance against wavelength coverage, bandwidth and breakdown voltage when selecting the material platform.

Temperature further complicates the picture. Rising lattice temperature increases phonon scattering, shortening carrier mean free paths and lowering the multiplication gain. At the same time, elevated phonon populations can enhance non-radiative recombination, siphoning carriers from the avalanche chain and adding another layer of variability. Effective noise control in high-temperature environments consequently

demands both material optimisation—favouring high $|\alpha/\beta|$ ratios—and thermal management to stabilise the field profile.

Practical mitigation strategies include tailoring the depletion-layer width so that only one carrier species experiences the highest field, shaping the electric-field gradient to minimise dead-space effects, and selecting materials with inherently disparate ionisation rates. Complementary approaches at the circuit level—such as cryogenic cooling, signal averaging, or analogue/digital filtering—can further suppress the influence of excess noise, but these add complexity and cost. As performance targets tighten for coherent optical links, imaging arrays and scientific detectors, continued refinement of material growth, device architecture and excess-noise modelling remains indispensable.

Reference

[1] W. Shockley, “The theory of p–n junctions in semiconductors and p–n junction transistors,” *Bell System Technical Journal*, vol. 28, no. 3, pp. 435–489, 1949.

[2] S. N. Sze, *Physics of Semiconductor Devices*, 2nd ed., Wiley, 1981.

[3] M. J. Howes and D. V. Morgan, *Gallium Arsenide: Materials, Devices, and Circuits*, Wiley-Interscience, 1985.

[4] L. Wang *et al.*, “Novel dilute bismide: epitaxy, physical properties and device application,” *Crystals*, vol. 7, p. 63, 2017.

[5] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, “Band parameters for III–V compound semiconductors and their alloys,” *Journal of Applied Physics*, vol. 89, no. 11, pp. 5815–5875, 2001.

[6] R. Lewis, M. Masnadi-Shirazi, and T. Tiedje, “Growth of high-Bi-concentration $GaAs_{1-x}Bi_x$ by molecular beam epitaxy,” *Applied Physics Letters*, vol. 101, p. 082112, 2012.

[7] X. Lu *et al.*, “Effect of molecular beam epitaxy growth conditions on the Bi content of $GaAs_{1-x}Bi_x$,” *Applied Physics Letters*, vol. 92, p. 192110, 2008.

[8] R. D. Richards *et al.*, “GaAsBi: from molecular beam epitaxy growth to devices,” *physica status solidi (b)*, p. 2100330, 2021.

[9] R. D. Richards *et al.*, “Molecular beam epitaxy growth of GaAsBi using As_2 and As_4,” *Journal of Crystal Growth*, vol. 390, pp. 120–124, 2014.

[10] C. Foxon and B. Joyce, “Interaction kinetics of As_4 and Ga on (100) GaAs surfaces using a modulated molecular beam technique,” *Surface Science*, vol. 50, pp. 434–450, 1975.

[11] E. Tok, J. Neave, J. Zhang, B. Joyce, and T. Jones, “Arsenic incorporation kinetics in GaAs(001) homoepitaxy revisited,” *Surface Science*, vol. 374, pp. 397–405, 1997.

[12] M. Winter, “The Periodic Table of the Elements,” WebElements (accessed May 5 2024).

[13] A. R. Mohmad *et al.*, “Localization effects and band gap of GaAsBi alloys,”

Physica Status Solidi (b), vol. 251, no. 6, pp. 1276–1281, 2014.

[14] G. E. Stillman and C. M. Wolfe, "Avalanche photodiodes," in *Semiconductors and Semimetals*, vol. 12, pp. 291–393, 1977.

[15] A. G. Chynoweth, "Ionization rates for electrons and holes in silicon," *Physical Review*, vol. 109, no. 5, pp. 1537–1540, 1958.

[16] M. M. Hayat, B. E. A. Saleh, and M. C. Teich, "Effect of dead space on gain and noise of double-carrier-multiplication avalanche photodiodes," *IEEE Transactions on Electron Devices*, vol. 39, pp. 546–552, 1992.

[17] M. A. Saleh *et al.*, "Breakdown voltage in thin III–V avalanche photodiodes," *Applied Physics Letters*, vol. 79, pp. 4037–4039, 2001.

[18] 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," *Journal of Applied Physics*, vol. 83, pp. 3426–3428, 1998.

[19] K. Alberi et al., "Valence-band anticrossing in mismatched III–V semiconductor alloys," Physical Review B, vol. 75, p. 045203, 2007, doi: 10.1103/PhysRevB.75.045203.

[20] Y. Liu et al., "Valence band engineering of GaAsBi for low noise avalanche photodiodes," Nature Communications, vol. 12, article 4784, 2021, doi: 10.1038/s41467-021-24966-0.

Chapter3: Molecular Beam Epitaxy (MBE) Growth Techniques

Molecular Beam Epitaxy (MBE) is a growth technique that has been utilized since the 1970s to produce high-quality semiconductor layers with precise control over thickness [1, 2]. The essential components of this technique include:

A growth chamber that maintains an ultra-high vacuum environment.

Multiple cells containing high-purity elemental sources.

Shutters designed to enable abrupt changes in molecular flux reaching the substrate.

A sample stage equipped with an integrated substrate heater.

Thermocouples and power supplies to monitor and regulate substrate and cell temperatures.

During the growth process, the materials in the cells are heated to temperatures where significant sublimation or evaporation occurs. By opening the shutters, the vaporized material is directed toward the substrate, which is heated on the sample stage. Once on the surface, adatoms briefly diffuse before bonding to exposed lattice sites, forming new semiconductor unit cells. Subsequent sections of this chapter will provide a more detailed exploration of the critical aspects of this growth process.

3.1 Producing and Sustaining a Vacuum

Maintaining an ultra-high vacuum (UHV) in an MBE growth chamber is crucial for producing high-quality semiconductor layers. Reactive species such as $O_2, H_2O, CO,$ and CO_2 can integrate into the growing crystal lattice, causing defects [1]. Achieving and sustaining this vacuum requires a combination of pumping systems:

Scroll or Diaphragm Pumps: Operating between 1000 mbar (atmospheric

pressure) and 5×10^{-3} mbar, these pumps are often referred to as "roughing" or "backing" pumps. They are used initially to reduce chamber pressure after exposure to the atmosphere and subsequently maintain the low exhaust pressure needed for turbo-molecular pumps.

Turbo-Molecular Pumps: These pumps replace roughing pumps to create a high vacuum by utilizing extended mean free paths at low pressure and molecular drag to pump gases efficiently [3]. They operate at high rotational speeds, typically ≥ 750 Hz.

Cryo Pumps or Cryo Shrouds: Utilizing super-cooled surfaces as freezing traps, these systems collect materials from within the chamber [3]. While cryo pumps use liquid helium compressors to achieve temperatures as low as ~ 4 K, cryo shrouds more commonly employ liquid nitrogen (LN_2), offering a larger surface area. These systems typically operate at ~ 80 K but can reach a minimum temperature of ~ 180 K.

Titanium Sublimation Pumps: These pumps utilize titanium (Ti) for its getter effect, trapping gas species within the chamber. Titanium filaments are "fired" using ohmic heating, causing Ti to sublimate and coat the chamber walls. This coating captures loose gas molecules, reducing chamber pressure. A baffle is often installed to prevent Ti deposition on sensitive in-vacuum electronics.

Ion Getter Pumps: These pumps operate in two stages. A high electric field (3–7 kV) ionizes gas species, drawing their electrons toward an anode and their nuclei toward a cathode made of getter material. When nuclei strike the cathode, they implant themselves or sputter cathode material, which captures additional gas species [3].

The MBE system utilized for layer growth in this work is an Omicron MBE-STM reactor. It is equipped with a variety of vacuum pumps, including turbo-molecular

pumps supported by scroll pumps, a titanium sublimation pump, an ion pump, and a cryo shroud lining the interior of the growth chamber. The layout of these pumps within the two primary chambers is illustrated in Figure 3.1.

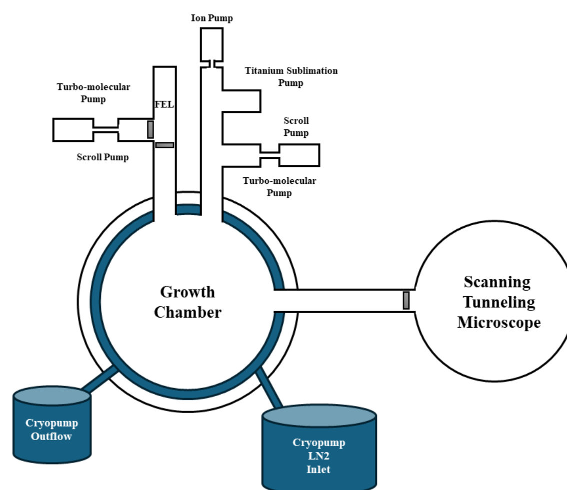


Figure 3.1: A diagram illustrating the general positioning of pumps within the MBE-STM system. Gate valves used to isolate chambers and pumps in the system are represented as solid gray rectangles. The scanning tunneling microscope (STM) connected to the growth chamber was not utilized in this work and is therefore not detailed.

This system comprises three main chambers: the Fast Entry Load-lock (FEL), the growth chamber, and the Scanning Tunneling Microscope (STM). However, as the STM is not currently operational, it will not be discussed further.

The Fast Entry Load-lock (FEL) serves as the entry chamber for the system. It is frequently brought to atmospheric pressure to facilitate the exchange of substrates, with its small volume designed to minimize the time required to pump back down to vacuum. Finished layers are transferred to this chamber after growth is complete, where they are prepared for removal and replaced with fresh substrates.

During the exchange process, the FEL is isolated from the growth chamber and FEL turbo pump using gate valves. Dry nitrogen is then introduced into the FEL, raising the pressure to approximately 1.3 bar. Once the nitrogen stabilizes, the blanking flange

covering the FEL's access port is loosened until the nitrogen begins to escape. The flange is then fully removed to allow the substrate exchange. After the new substrate is in place, the chamber is resealed, the nitrogen line is closed, and the FEL is evacuated back to $\sim 1 \times 10^{-7}$ mbar using a dedicated turbo pump and backing pump.

The growth chamber, as its name suggests, is the site where new materials are epitaxially deposited onto the substrate. It operates under vacuum for extended periods and has the highest number of pumps connected to maintain its conditions.

3.2 Molecular Sources and Generating a Molecular Beam

To produce the molecular fluxes necessary for growth, extremely high-purity elemental sources are required. The purity of the source material is typically expressed as '#N,' where '#' represents the number of significant figures of purity (e.g., 5N = 99.999% pure). Figure 3.2 below provides a general illustration of the arrangement of the cells within the growth chamber relative to the substrate.

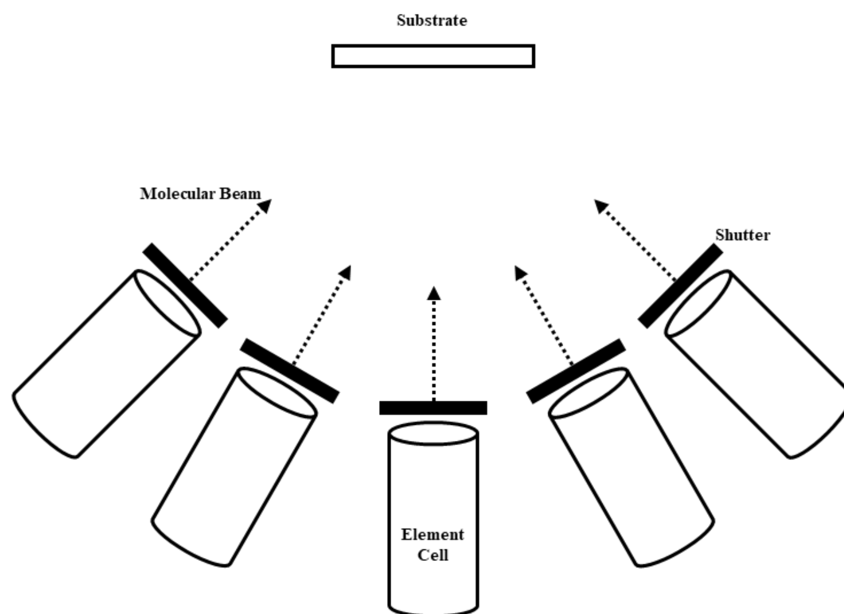


Figure 3.2: Substrate and cell positioning within the MBE reactor. Adapted from [2].

A key requirement for an MBE system is the high stability of the source beams. When source material of sufficient purity is available, the generation of the elemental beam from the bulk material is typically achieved through a thermal process. Heating the material within the vacuum raises its vapor pressure, resulting in a significant flux of evaporated or sublimated material, depending on whether the source is a liquid or a solid at the given temperature. The traditional device used to hold the bulk source material and regulate the beam flux through radiative heating is a Knudsen effusion cell.

For a Knudsen effusion cell with a well-defined geometry, the molecular flux output can be expressed as shown in Equation 3.1 [1]:

$$J(\text{atoms cm}^{-2}\text{s}^{-1}) = 1.118 \times 10^{22} \frac{p \times A}{l^2 \times (M \times T)^{1/2}} \quad (3.1)$$

Where p (Torr) is the source vapor pressure, A (cm^2) is the cell aperture area, l (cm) is the distance between the cell and the substrate, M is the source molecular weight, and T (K) is the source temperature. Models have also been developed to account for non-uniform flux across the substrate surface due to non-normal incidence of impinging fluxes, which has proven useful in mapping the effects of growth conditions on semiconductors grown in stationary positions [4].

Significant effort has been dedicated to the design of the Knudsen effusion cell, a critical component in MBE systems. This includes sophisticated design features to enable operation at high temperatures and to minimize thermal interference between neighboring cells. An annotated cross-section of a Knudsen cell is provided in Figure 3.3.

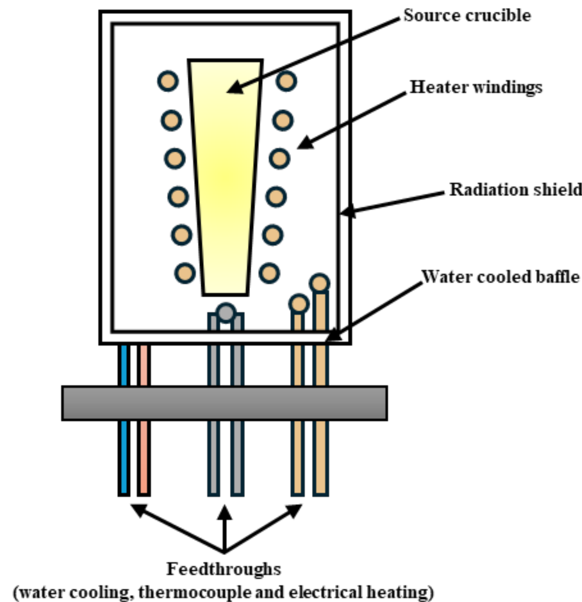


Figure 3.3: A generalized cross-section of a Knudsen effusion cell.

Refractory metals such as molybdenum, tantalum, and tungsten are commonly used in many MBE components. This is because, at temperatures exceeding 150°C, impurities like manganese, magnesium, and chromium are often released from stainless steel [1]. Similarly, strict material selection is applied to electrical insulation, with the preferred material being pyrolytic boron nitride (PBN).

PBN (pyrolytic boron nitride) can be manufactured with impurity levels below 10 ppm, making it highly suitable for use in MBE systems. While PBN begins to degrade at temperatures above 1400°C, such conditions are rarely encountered, and the nitrogen released at these temperatures has not been shown to significantly affect layer properties. As such, PBN remains the preferred material for source crucibles and is also used for electrical insulation around substrate heaters. Other materials, such as graphite and quartz, are selectively used for specific applications.

To measure the in-vacuum temperature of cells during operation, thermocouples are commonly employed. Type-K thermocouples, composed of nickel and chromium, are widely used due to their operational temperature range of 0–1200°C. They generate a voltage output of 0–50 mV, which is converted into a temperature reading using an analog-to-digital converter.

The water-cooled baffle surrounding the cell plays an essential role in capturing and dissipating thermal radiation from the cell, thereby preventing thermal interference with adjacent cells. The water network supplying these baffles operates at low pressures (~1.3 bar) and temperatures between 14–30°C.

Some Knudsen effusion cells are modified to include a high-temperature channel and are referred to as "dissociation cells" or "cracker cells." These are primarily used with group V sources, such as arsenic (As), where the sublimated species from the bulk is the tetramer As_4. In these cells, the As_4 vapor passes through a heated channel maintained at 650–1000°C. During transit, molecular collisions with the channel walls can cause the tetramer to dissociate into As_2 species. By carefully designing the channel to ensure sufficient collisions, cracker cells can reliably produce a beam of either As_2 or As_4, depending on the temperature. While cracker cells are sometimes used for other group V elements, they are not commonly employed for bismuth (Bi) sources due to the natural production of Bi_2 dimers and Bi monomers.

Molecular flux produced by a cell at a specific temperature is measured using an ion gauge, positioned in the direct path of the molecular beam near the substrate's focal point. Flux measurement relies on detecting changes in ionization current through the ion gauge before and after the cell shutter is opened. The ion gauge outputs its readings in pressure units (mbar or Torr), displayed on its controller.

Flux readings, expressed as beam equivalent pressures (BEPs), are not absolute flux measurements, as they vary across materials due to differences in ionization efficiency. However, BEPs provide a comparative measure for individual cells and can characterize their performance within a specific reactor. For group III and Bi cells, BEPs typically range from 1×10^{-8} to 5×10^{-7} mbar. For As cells, BEPs usually fall between 1×10^{-7} and 1×10^{-5} mbar.

3.3 Cell Temperature Control and Outgassing

Leaving material cells at operational temperatures during periods of inactivity, such as overnight or over weekends, would rapidly deplete the material, increasing downtime and maintenance frequency. To avoid this, cells are cooled to a rest temperature, typically 100–600°C below their operational temperature, depending on the material. For molten sources such as aluminum (Al) and bismuth (Bi), the rest temperatures are limited to 850°C and 300°C, respectively, once liquefied.

A challenge arises as gas species within the chamber may condense on cooling cells, and these condensates can release back into the chamber when reheated. If this occurs during growth, these impurities can interact with the growing material, causing defects in the lattice. To mitigate this, cells are preheated to approximately 20°C above their operational temperature and outgassed before substrate loading. This process releases foreign materials, allowing them to be pumped away before growth begins. The slightly elevated temperature ensures contaminants with similar vapor pressures as the source material are quickly released, while those with significantly higher vapor pressures remain unaffected.

Outgassing is particularly crucial after maintenance when the chamber has been exposed to atmospheric pressure and may have absorbed significant reactive species. This process typically involves a 48-hour bake-out at approximately 200°C while being pumped. However, there is debate about whether MBE chambers should undergo bake-outs, as excessive heat can damage delicate equipment or compromise gasket interfaces. The bake-out aims to remove as much moisture as possible, as water desorption rates are exponentially higher at elevated temperatures than at room temperature. Additionally, purging and pumping with dry nitrogen before reopening the reactor can prevent moisture buildup and dilute any residual toxic gases. Post-closure purging is also beneficial for removing condensed moisture within the chamber.

3.4 Automation

The complexities of MBE growth have been significantly reduced by advancements in computing power and automation. Modern systems allow for fully automated control of cell temperatures, shutters, and growth recipes, enabling continuous operation for extended periods without human intervention. Digital logging of measured fluxes and remote desktop access facilitate real-time data recording and analysis without requiring operators to be physically present. On some reactors, fiber-optic sensors and servo motors have been implemented to automate substrate loading and exchange within vacuum systems. However, this advanced level of automation was not available on the MBE reactor used in this work.

3.5 III-V Growth Overview

With the engineering and automation aspects of MBE covered, a summary of III-V compound growth can now be provided.

When the growth temperature is below the congruent sublimation temperature of a III-V compound, it is assumed that the group III material has a sticking coefficient of 1, meaning that only enough group V material adheres to satisfy this supply. Excess molecules are desorbed back into the vacuum [1]. As a result, the growth of III-V compound semiconductors is often group III limited, where the total flux of group III material determines the crystal growth rate. This group III flux control simplifies alloying of group III metals since the solid phase composition of the alloy is primarily dictated by the fluxes of the individual metals. However, alloying group V metals is more challenging, particularly for alloys containing Bi, as discussed previously in Section 2.7.

In addition to its use as a constituent, Bi has also found applications in MBE as a surfactant. For example, Bi has been utilized as a surfactant in InGaAs/GaAs heterostructures [5], where it reduced surface roughness and improved

photoluminescence (PL) intensity. Similarly, in Ge/Si growth [6], Bi was shown to suppress 3D islanding and segregation of Ge. Surfactants are materials that do not typically integrate into crystal growth but significantly influence how a material grows. Essentially, surfactants act as adsorbed species that modify the surface free energy and growth kinetics of a material [7]. They are classified as "reactive" or "non-reactive" based on whether they decrease or increase adatom migration lengths during growth. Reactive surfactants, such as Bi during GaAs growth, interact with lattice sites, whereas non-reactive surfactants act interstitially. Given its role in GaAsBi literature, where Bi segregates to the surface, it is likely that Bi also acts as a surfactant during growth.

3.5.1 Limitations of low-temperature III–V MBE growth

Low-temperature MBE is intrinsically a non-equilibrium growth regime. As the substrate temperature is reduced, desorption is suppressed and surface diffusion lengths shorten, so the surface stoichiometry and morphology become increasingly controlled by kinetics rather than by near-equilibrium reconstruction thermodynamics. This reduced surface mobility can hinder step-flow and layer-by-layer growth, thereby increasing the propensity for kinetic roughening and three-dimensional islanding, even when the nominal flux ratio remains within a conventional “2D growth” window. [29,30]

A second consequence is the enhanced probability of incorporating excess group-V related point defects and background impurities. In GaAs, extensive low-temperature MBE studies show that lowering the growth temperature can strongly increase the incorporation of oxygen and carbon below a critical temperature, with impurity levels rising by orders of magnitude in the low-temperature regime. [31] Such impurities and native point defects can introduce deep levels and act as non-radiative recombination centres, which is directly relevant to device-oriented structures where carrier lifetime, leakage, and noise are critical. [30,31]

For GaAsBi, these generic limitations coexist with the Bi-specific requirement of operating at reduced substrate temperatures to suppress Bi desorption and phase separation. Consequently, achieving high structural and electronic quality at 350–450 °C requires particularly stringent UHV practices (e.g., thorough chamber bake-out and careful outgassing of sample holders), together with growth-condition optimisation (growth rate, As:Ga ratio, and As molecular species) to stabilise a

smooth growth front. [29–31]

3.6 Reflection High-Energy Electron Diffraction (RHEED)

The primary in-situ tool for monitoring substrate conditions during MBE growth is Reflection High-Energy Electron Diffraction (RHEED). This technique utilizes a collimated electron beam accelerated to energies of 10–20 keV, incident on the substrate at a shallow grazing angle of 2–3°. Due to the low incident angle, the electron beam is diffracted predominantly by the top few monolayers of the substrate [8]. The resulting diffraction pattern is displayed on a phosphor screen positioned opposite the electron gun, providing a detailed view of the surface structure.

A diagram illustrating the operation of a RHEED system is shown below in Figure 3.4.

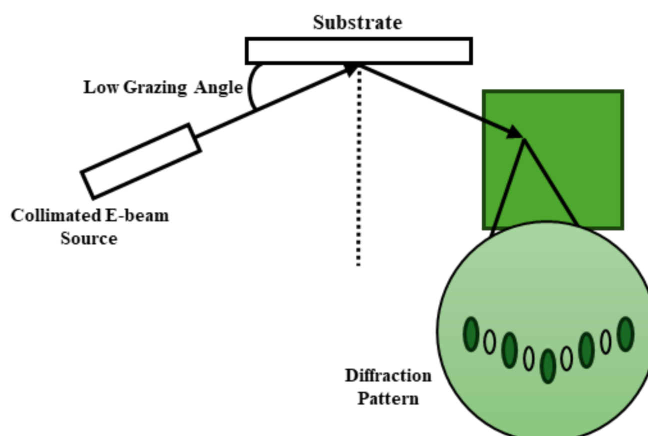


Figure 3.4: RHEED System Diagram. The diffraction pattern (periodic green rods) arises from elastic diffraction collisions, while the green background haze is caused by inelastic collisions and electron scattering within the reactor.

The diffraction pattern produced in RHEED is influenced by the surface morphology of the substrate and the crystal orientation relative to the electron beam. During growth, the surface does not form half-unit cells but instead adopts periodic structures, known as reconstructions, to minimize surface energy. In this work, two primary reconstructions of GaAs were observed: the (2×4) and $c(4 \times 4)$ reconstructions.

In addition to elastic scattering, inelastic scattering events play a crucial role in the RHEED pattern. These events include interactions with phonons or plasmons within the lattice, resulting in a diffuse cone of electrons that are subsequently diffracted by the crystal lattice planes, depending on the beam angle. This leads to the formation of sharp lines, bands, and arcs in the RHEED pattern, commonly referred to as Kikuchi lines [8].

The intensity of Kikuchi lines depends on several factors beyond beam orientation. These include surface morphology (e.g., steps and terraces that cause line broadening), bulk crystallinity (although due to the beam's exit angle, this only probes approximately 10 unit cells), and crystal mass. Heavier materials, such as Si and silicon carbide (SiC), display stronger Kikuchi line intensity. Collectively, these features aid in aligning the crystal azimuth and monitoring thin-film quality based on line intensity. However, since such methods are not employed in this work beyond observation, they will not be discussed further.

RHEED measurements have multiple applications beyond verifying reconstructions and confirming the absence of growth issues. These will be explored in the following subsections.

3.6.1 RHEED Measurements for Substrate Heater Calibration

In an MBE system, measuring the substrate temperature may not always be possible using a pyrometer or thermocouple due to various challenges. Pyrometers can be costly and require calibration with surface parameters such as emissivity. Thermocouples are difficult to place in direct contact with the substrate, limiting their accuracy. Additionally, equipment faults may occur during a growth campaign, complicating temperature monitoring.

If a functioning RHEED system is installed, it is possible to use calibration points to correlate substrate temperature with a heating variable, such as voltage, current, or power. Among these, heater current is the most reliable option. These calibration points

are based on well-established conditions where a temperature change causes a significant alteration in the surface RHEED pattern. For a $GaAs(001)$ surface, three critical calibration points are:

Desorption of an amorphous As cap at approximately 300°C.

The transition between $c(4 \times 4)$ and (2×4) reconstructions at approximately 400°C.

Oxide sublimation from the substrate surface at approximately 590–600°C.

Throughout this work, the GaAs(001) surface reconstructions observed by RHEED were used as practical temperature-calibration markers during oxide desorption and As-cap removal. At elevated temperatures the As-rich c(4×4) phase evolves towards (2×4); however, the apparent transition temperature is not universal and depends strongly on the arsenic chemical potential (i.e., whether an external As flux is present, the As2/As4 species, and the beam-equivalent pressure). Under external As fluxes, (2×4) stability is typically reported for $T \geq 480$ °C, and a c(4×4)↔(2×4) transition near $T \approx 525$ °C has been reported for specific As4 flux conditions. [32–34] In contrast, when the transition is identified in the absence of external flux during pre-growth decapping procedures, the corresponding thermocouple set-point can be significantly lower; values around 400 °C have been used as a practical calibration point. [32] Consistent with this flux-free calibration condition, our system exhibits the c(4×4)→(2×4) transition at a thermocouple temperature of approximately 400 °C; throughout this thesis, such temperatures should be interpreted as thermocouple set-points rather than absolute wafer-surface temperatures. [32–34]

Once measured, these three points can be used to create an exponential trendline, enabling estimation of the heater current required to maintain any temperature between 300–600°C, with an accuracy of ±10°C. However, these calibrations depend heavily on substrate thickness and doping, as these factors influence the absorption of incident radiation and, therefore, the surface temperature. Consequently, calibrations must be repeated whenever the substrate type or supplier changes. To monitor potential drift in heater output, it is also recommended to check these calibrations on a weekly or bi-monthly basis.

3.6.2 RHEED Measurements for Finding Growth Rates

One common application of a RHEED system in MBE is determining growth rates [8]. This involves monitoring the intensity of diffraction spots over time after the growth process begins. During growth, the surface undergoes cycles of roughening and smoothing as reconstructions break and reform unit cells. This process is illustrated in Figure 3.5.

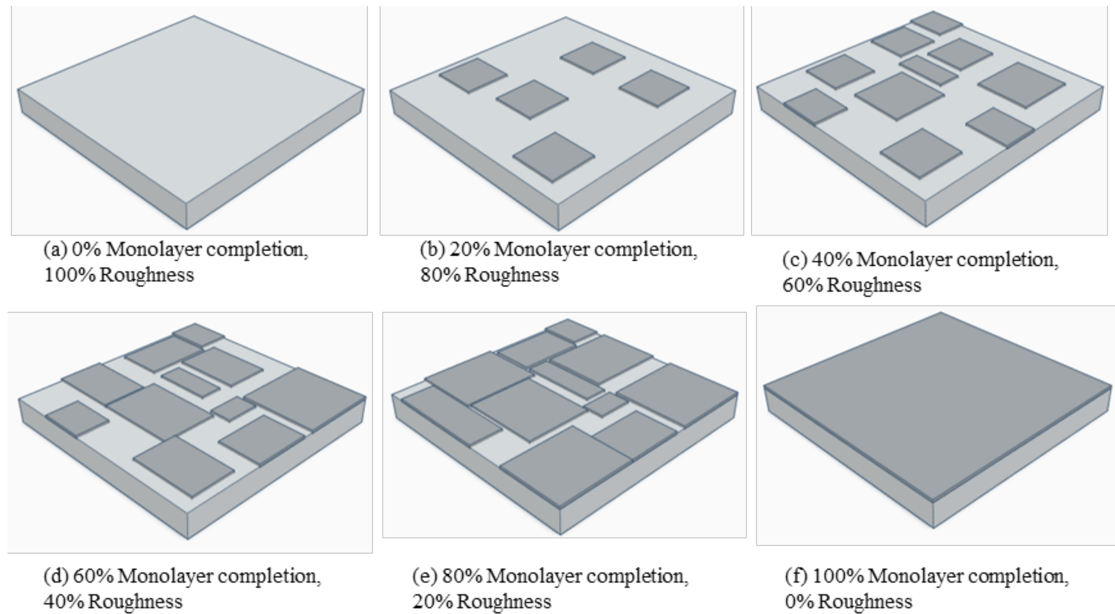


Figure 3.5: Simplified representation of the surface profile during the growth of a single monolayer.

This roughening and smoothing process can also be observed as periodic dimming and brightening of the RHEED pattern, corresponding to variations in layer uniformity. For example, one oscillation would typically represent the growth of a single monolayer [9]. By recording these oscillations over a period of time, using the lattice constant of the semiconductor and calculating the mean oscillations per second, the growth rate under specific growth conditions can be determined in-situ. These measurements should be taken at the onset of growth, as the oscillations tend to dampen over time. For GaAs, this process involves annealing the surface under an As flux, setting up a

data logger to monitor the RHEED intensity, and then opening the Ga shutter to resume growth. An example of growth rate oscillation data will be provided in Section 3.8, along with a description of the hardware used.

3.6.3 RHEED Measurements for As_2:Ga 1:1 Calibration

Another application of a RHEED system is calibrating the 1:1 atomic flux ratio between As_2 and Ga being incorporated into the growing crystal. This ratio determines the atomic flux of As required to match the deposition rate of Ga on the semiconductor surface and the rate of As desorption. This is particularly important because, while a wide permissible range exists for high-temperature GaAs growth, the As overpressure significantly influences Bi incorporation, as discussed in Section 2.7. Richards et al. [10] demonstrated this effect, and the results are illustrated in Figure 3.6 below.

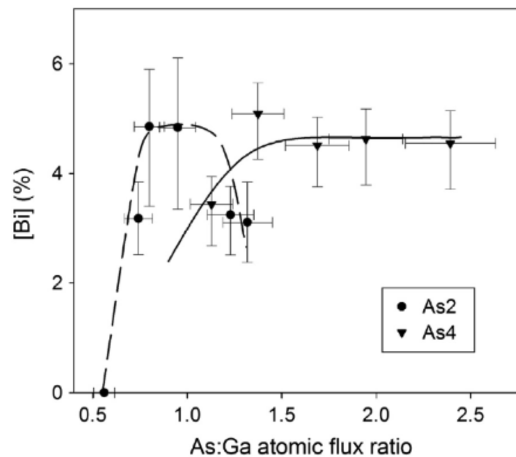


Figure 3.6: Comparison of Bi incorporation for various As_2 and As_4:Ga atomic flux ratios. Reprinted from *Journal of Crystal Growth*, 390, R. D. Richards et al., *Molecular beam epitaxy growth of GaAsBi using* As_2 *and* As_4, 120–124, Copyright (2014) [10].

The process for determining the 1:1 point involves sequentially reducing the As_2 flux and opening the Ga shutter for two-minute intervals while monitoring the RHEED

pattern. Research groups handling large amounts of As for III-V growth often employ valved cracker cells, allowing quicker adjustments of the As flux by changing the position of a needle valve, instead of altering the As bulk material's temperature. Adjustments to the As flux are thus managed by increasing or decreasing the valve position to modify the flux accordingly.

After each two-minute exposure, the surface must be annealed for at least 30 seconds to ensure it stabilizes to its initial state. Sufficient time must also be allowed for the chamber pressure to equilibrate under the new As valve position.

Before opening the shutter, the surface should exhibit a static (2×4) reconstruction, typically formed on a GaAs buffer layer grown at an optimal MBE temperature of approximately 580°C. After opening the shutter, the surface reconstruction may change depending on the As:Ga ratio and substrate temperature, transitioning to one of the reconstructions shown in Figure 3.7. If no change in reconstruction occurs, the As_2 flux is still too high, and the valve position should be further reduced.

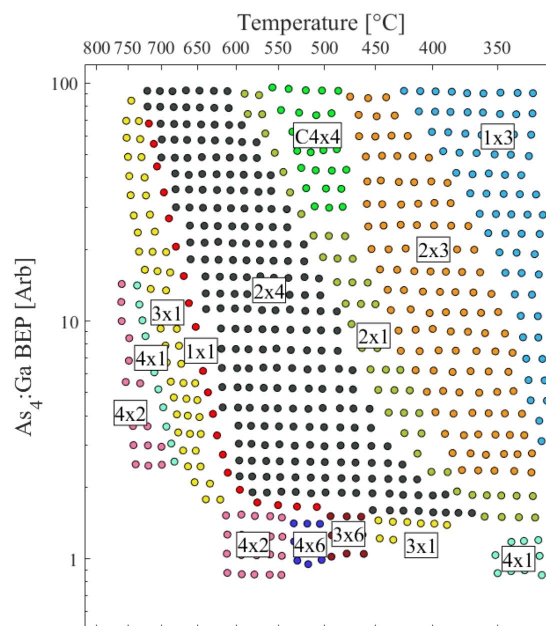


Figure 3.7: Dynamic reconstructions of GaAs during growth. Adapted from [11].

When approaching the 1:1 balance, it is expected to observe the pattern transition to (1×1) and possibly (4×2) or (4×6). The (4×2) and (4×6) reconstructions are Ga-rich structures, indicating an oversupply of Ga relative to As_2. If this occurs, the valve position should be gradually increased in smaller increments, and the test repeated until the RHEED pattern stabilizes at (1×1) without reverting to (4×2) or (4×6) within the allocated time. This point is referred to as the 1:1 balance for As_2. To convert the flux value for As_2 to As_4, the calibration value is doubled, as As_4 has a sticking coefficient of approximately 0.5 [12], compared to around 1 for As_2 when growth is Ga flux-limited [12, 13]. This assumes that the As cracker (used to thermally crack As_4 into As_2) does not affect the total atomic flux incident on the sample.

It is important to note that the reconstruction map in Figure 3.7 serves only as a guide. The As_4:Ga ratios are reported as beam equivalent pressures (BEPs), which depend on ionization efficiencies that vary between materials and can fluctuate over time due to wear and material accumulation on ion gauges. The 1:1 ratio being calibrated refers specifically to the atomic flux ratio rather than the BEP ratio.

3.7 Heater Calibration Process

As outlined in Section 3.6.1, the RHEED system can be used alongside a set of well-defined calibration points to calibrate the substrate heater, particularly when a pyrometer or thermocouple is unavailable on the MBE reactor. The calibrations in this work are based on the supplied current, ensuring that any voltage drops across wire connections can be ignored. This ensures the heating element receives consistent input from the power supply.

The first calibration point for a new substrate type is the desorption of the native oxide, which occurs at approximately 600°C. Before this, the RHEED pattern appears as an amorphous haze caused by the oxide layer covering the substrate. As the oxide

sublimates, a diffraction pattern from the exposed semiconductor surface emerges. However, oxide removal leaves the surface slightly roughened, producing a RHEED pattern with 3D features, as shown in Figure 3.8.

Before further calibrations are performed, a buffer layer is grown to smooth the surface and establish a recognizable reconstruction. This buffer is typically 200–300 nm thick and grown at 580–590°C with an As_2:Ga ratio of ~1.6:1. Under these conditions, the surface forms a (2×4) reconstruction. An image of this reconstruction is provided in Figure 3.9.

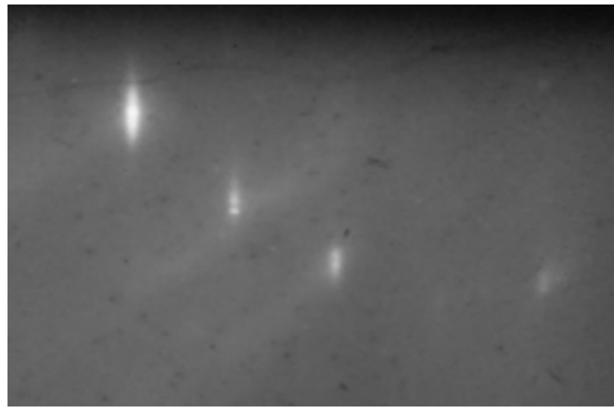


Figure 3.8: Oxide-removed "3D" RHEED pattern.

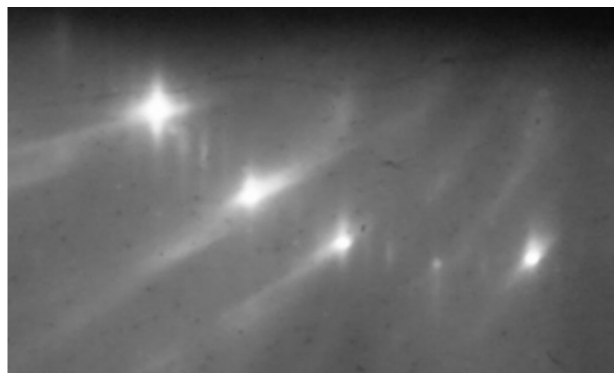


Figure 3.9: (2×4) RHEED pattern. Annotations indicate the integer and fractional order diffractions.

The next calibration point is the removal of the As cap, which occurs at approximately

300°C. This process begins with the deposition of an As cap onto the surface of the buffer layer grown in the previous step. The As valve is fully opened, and the heater current is dropped to 0 A, allowing As to coat the surface in a random configuration. This results in an amorphous As cap, which scatters electrons randomly, producing a uniform green haze in the RHEED pattern, similar to the one observed before oxide removal.

The heater current is then gradually increased from 0 A to around 0.55 A (depending on substrate doping) in small increments every 10 minutes until a $c(4 \times 4)$ RHEED pattern emerges. This transition marks the sublimation of the As cap, revealing the underlying reconstruction. Figures 3.10 and 3.11 depict the RHEED patterns before and after As cap removal.

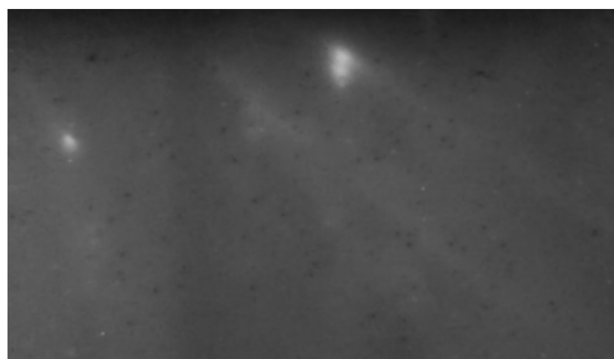


Figure 3.10: Hazy As cap. The incident electron beam is absorbed and scattered by the amorphous As surface, preventing the formation of diffraction peaks.

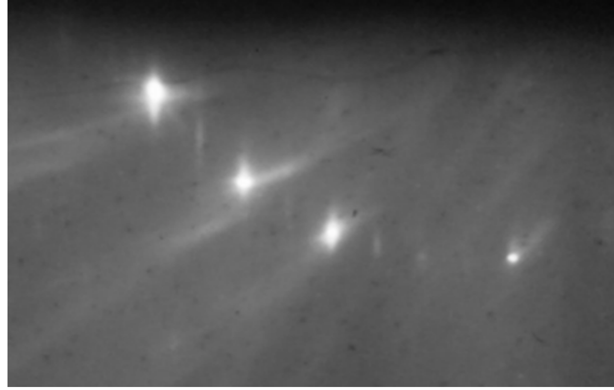


Figure 3.11: $c(4 \times 4)$ Reconstruction after As cap removal. Annotations indicate the integer and fractional order diffractions. Image captured along the $[1\bar{1}0]$ azimuth at 14.5 keV.

The final calibration point involves the transition between the $c(4 \times 4)$ and (2×4) reconstructions, which occurs at approximately 400°C in the absence of an external As flux. As with previous steps, the heater current is gradually increased from below the expected transition current in 10-minute increments. The absence of an As flux is critical since it would alter the transition temperature.

During this process, As sublimates from the surface at the calibration temperature. It is essential to allow the substrate sufficient time at each current setting to confirm the transition. If the substrate is left too long, it may become Ga-rich, leading to Ga droplet nucleation that could irreparably damage the surface.

Once all calibration points have been identified, the heater currents and corresponding characteristic temperatures are plotted, as shown in Figure 3.12. An exponential trend-line is used to illustrate the relationship between substrate temperature and heater current.

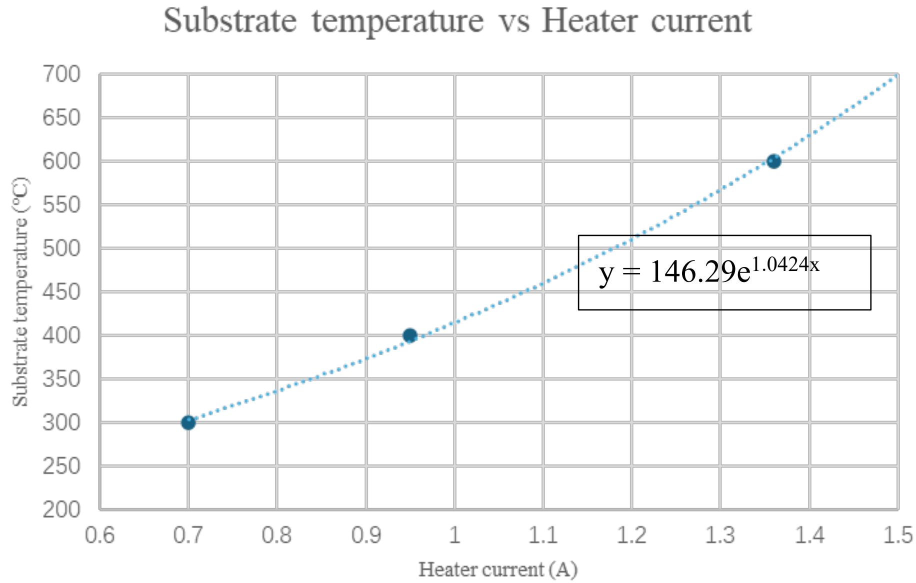


Figure 3.12: Plot of heater calibration points with an exponential trend-line.

Figure 3.12 presents the empirical relationship between the thermocouple readout and the substrate-heater current over the operating range relevant to this work. The dependence is monotonic but noticeably non-linear; therefore, a simple analytical function was introduced to provide a convenient interpolation between calibration points. While an exponential curve provides a compact empirical fit over the limited current range shown, this functional form is used here strictly as an interpolation aid and is not intended to imply a fundamental exponential physical law of the heater–wafer thermal system.

Using the exponential trend-line, an arbitrary point can be mapped onto the curve to estimate the required current for any temperature within the calibrated range. These calibrations are performed for each substrate type and are checked weekly to ensure they remain accurate. To guarantee that samples are fully oxide-free during automated processes, the temperature is set approximately 20°C higher than required.

3.8 Growth Rate Calibration

As discussed in Section 3.6.2, RHEED measurements can also be utilized for growth rate calibration. Rather than performing this manually, a more practical approach involves using a webcam mounted in front of the phosphor screen and image analysis software, such as MATLAB. The live RHEED screen feed is processed in the software, where rectangles are drawn over regions of interest that the software will analyze. When the Ga shutter is opened, growth recommences on the surface, causing oscillations in the intensity of the diffraction pattern as monolayers are completed. The software records the average brightness of these regions over time, resulting in waveforms like those shown in Figure 3.13a.

The intensity oscillations that begin after the Ga shutter is opened at approximately 7 seconds are visually recognizable, but the noise in the data can confuse peak-finding algorithms. To address this, the data is processed through a moving average filter, smoothing the waveforms as seen in Figure 3.13b. Once the noise is removed, the software identifies a representative range of data, counts the number of oscillations within this range, and calculates the duration of these oscillations. This yields a growth rate in monolayers per second ($\mathrm{MLs^{-1}}$).

The growth rate can be converted into micrometers per hour ($\mathrm{\mu mh^{-1}}$) using Equation 3.2, where a is the lattice parameter of the material in meters. For GaAs, the conversion of 1 $\mathrm{MLs^{-1}}$ corresponds to 1.018 $\mathrm{\mu mh^{-1}}$, and these values are considered equivalent for simplicity.

$$GR(\mathrm{\mu m/h}) = \frac{GR(\mathrm{ML/s}) \times a \times 10^{6} \times 3600}{2} \quad (3.2)$$

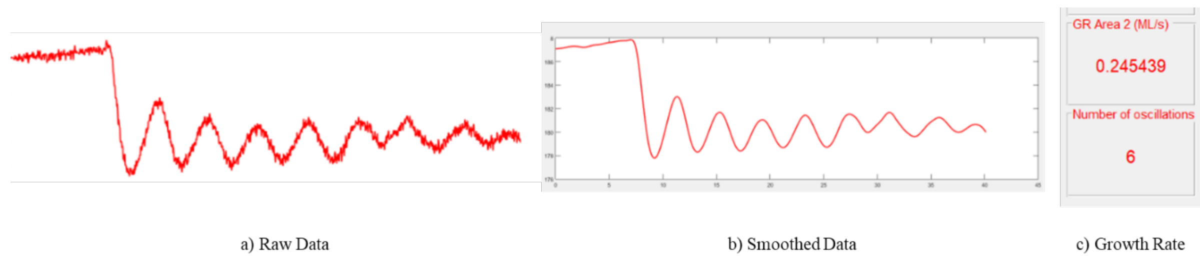


Figure 3.13: Growth Rate Measurements

3.9 GaAs Growth

Growing a high-quality GaAs buffer layer is a prerequisite for producing GaAsBi. Therefore, understanding GaAs growth provides a deeper insight into the mechanisms of MBE growth. Studies have shown that high-quality GaAs is typically grown at substrate temperatures ranging from 600–640°C [14, 15]. If the substrate temperature exceeds 650°C, the GaAs growth rate decreases due to Ga evaporation from the surface [16]. This temperature, known as the congruent evaporation temperature, marks the point above which both group III and V elements become unstable. For optimal growth, the temperature should remain below this critical point.

In our lab, the substrate temperature is regulated by heater current, which introduces slight deviations from the target temperature. Practical observations indicate that temperatures above 600°C result in suboptimal GaAs quality in our setup. Hence, we set the substrate temperature to approximately 580°C to ensure high-quality growth.

There are two arsenic species, As_2 and As_4, that can be used for GaAs growth via MBE. The primary difference lies in their reaction mechanisms with Ga atoms. For As_2, dissociative chemisorption occurs with a single Ga atom, whereas for As_4, pairs of As_4 molecules react with adjacent Ga atoms [17]. Research has shown that As_2 results in lower electron trap concentrations compared to As_4 at 537°C. Even at lower growth temperatures, As_2 demonstrates superior performance, such as brighter photoluminescence (PL) and better interface quality. Thus, our lab utilizes As_2 for

GaAs growth.

To ensure stoichiometric growth, group V elements must be oversupplied during the growth of III-V alloys. Arsenic, being highly volatile, readily evaporates above 300°C. As the temperature increases, the absence of an external As flux can cause surface roughness. The total required arsenic flux can be expressed as:

$$\mathrm{As_{total}} = \mathrm{As_{As}} + \mathrm{As_{Ga}} \tag{3.3}$$

Here, $\mathrm{As_{Ga}}$ represents the As flux needed to compensate for Ga flux to achieve a 1:1 stoichiometric ratio between As and Ga. Additionally, excess As ($\mathrm{As_{As}}$) compensates for arsenic loss through evaporation. The minimum As overpressure depends on the substrate temperature. Maintaining this excess flux ensures the surface remains annealed and flat, while minimizing As waste. Based on these principles, the As:Ga ratio is set to 1.6:1 [18] for GaAs growth in our lab.

3.10 GaAsBi Growth

3.10.1 Review of GaAsBi Growth Investigation

Since the first GaAsBi sample was grown using MOVPE in 1998 [19], numerous studies have explored the growth conditions and material properties of GaAsBi. These investigations have significantly contributed to understanding the growth behavior of Bi and have provided valuable insights into achieving high-quality Bi samples. Key findings include:

- Growth of GaAsBi requires low temperature and stoichiometric arsenic-to-gallium flux conditions.

An early study by Tixier *et al.* [20] described the molecular beam epitaxy (MBE) growth of GaAsBi. They found that Bi incorporation into GaAs is negligible when the substrate temperature exceeds 450°C, even with a high arsenic-to-gallium flux ratio. Furthermore, they observed that maintaining an arsenic-to-gallium flux ratio close to

stoichiometry improves Bi incorporation.

Similarly, Young *et al.* reported that Bi has a strong tendency to segregate on the surface under conventional GaAs growth conditions. To mitigate this, both low growth temperatures and near-stoichiometric arsenic-to-gallium flux ratios are necessary. They successfully grew GaAsBi at temperatures ranging from 320°C to 390°C and employed in-situ light scattering as a tool to monitor and optimize growth conditions [21].

The choice of arsenic species for GaAsBi growth, As_2 or As_4, determines the optimal growth conditions.

Richards *et al.* explored the impact of arsenic species on GaAsBi growth by growing layers under both As_2 and As_4 conditions across a range of temperatures. They concluded that Bi incorporation into GaAs is independent of the arsenic species used. Additionally, the photoluminescence (PL) intensity of the samples remained unaffected by the choice of As_2 or As_4. However, they suggested As_4 as the preferred species due to its broader arsenic-to-gallium flux ratio range, which facilitates achieving optimal growth conditions more easily compared to As_2 [22].

The growth rate impacts bismuth incorporation.

Ptak *et al.* proposed a kinetic-limited growth regime for GaAsBi in 2012 [23]. They conducted experiments at a low growth temperature (315 ± 5°C) to examine the effects of varying growth rates and Bi fluxes. The root-mean-square (RMS) surface roughness was analyzed using atomic force microscopy (AFM). Their findings showed that Bi incorporation is proportional to the Bi flux in the kinetic-limited regime. In this regime, Bi atoms lack sufficient energy to segregate on the surface, resulting in most Bi atoms being incorporated into the layer.

Additionally, the authors highlighted that the growth rate plays a critical role in Bi incorporation. When a high growth rate and low Bi flux are applied, surface roughness increases due to inadequate surfactant effects. Conversely, with the same Bi flux but a lower growth rate, the surface remains smooth. However, applying high Bi flux at a low

growth rate causes significant roughness, attributed to the formation of Bi droplets. The high Bi flux introduces more Bi atoms onto the surface, leading to the formation of a Bi-metallic layer instead of incorporation into the layer.

These results provide a valuable guide for identifying the optimal growth window to achieve smooth, droplet-free GaAsBi growth.

The growth temperature determines the limitation of the dark current in GaAsBi diodes.

Rockett *et al.* investigated the effects of growth conditions on the structural and optoelectronic quality of GaAsBi diodes [24]. They studied samples grown at various temperatures and Bi fluxes within a growth temperature range of 355–405°C. Their analysis revealed that samples grown at higher temperatures exhibited higher room-temperature photoluminescence (RT-PL) intensity but showed non-uniform structural characteristics, as confirmed by X-ray analysis, consistent with previous studies [25].

Furthermore, the current-voltage (I-V) characteristics of the diodes were strongly influenced by the growth temperature. The dark current was found to be inversely proportional to the growth temperature, suggesting that reducing the growth temperature is essential for enhancing the electronic performance of GaAsBi devices.

3.10.2 Summary of GaAsBi Growth Conditions

Based on reviews of GaAsBi growth studies, the growth of GaAsBi must adhere to the following requirements:

1. The growth temperature should be below 450°C with a stoichiometric arsenic-to-gallium flux condition.
2. As_4 is the preferred arsenic species for the growth process.

Since the Bi content in our growth is generally below 5%, a low growth rate is more conducive to maintaining sample quality. This approach helps to minimize surface

roughness, which can occur due to the lack of surfactant effects. However, these methods may not be universally applicable, and additional calibration may be necessary for specific samples to achieve optimal growth conditions.

3.11 System Maintenance (Bake-Out Process)

Regular maintenance and inspection are crucial for ensuring the stability of the MBE system and maintaining high sample growth quality. The bake-out process is an essential procedure, especially after routine maintenance tasks such as refilling cell materials or repairing the substrate heater stage, which expose the system to atmospheric conditions. Following the baking process, material sources must also undergo thorough outgassing to ensure they remain clean.

According to the literature, bake-out procedures can vary between laboratories [26, 27, 28]. Once the system is exposed to the atmosphere, contaminants and moisture from the air adhere to the chamber's interior walls. At room temperature, these impurities take longer to evaporate and be pumped out. To expedite this process, the machine can be enclosed in a thermally insulated stainless steel box and heated to over 100°C for 48 hours, as shown in Figure 3.14.

At such high temperatures, all components must be made from refractory materials to withstand the heat. Any components that cannot endure prolonged high temperatures—such as water-cooling tubes, motorized parts of cells, rubber gaskets, and thermally sensitive connections—should be removed before baking.

In our lab, the bake-out temperature is set to 125°C for 2–3 days, achieving a background pressure below 5×10^{-10} mBar.

Figure 3.13: Bake-out preparation of MBE system

References

[1] E. H. Parker, *The technology and physics of molecular beam epitaxy*. Plenum press New York, 1985.

[2] B. R. Pamplin, *Molecular beam epitaxy*. Elsevier, 2013.

[3] B. D. Power, *High vacuum pumping equipment*. Reinhold Publishing Corporation, 1966.

[4] C. Wang, L. Wang, X. Wu, Y. Zhang, H. Liang, L. Yue, Z. Zhang, X. Ou, and S. Wang, "Molecular beam epitaxy growth of $AlAs_{1-x}Bi_x$," *Semiconductor Science and Technology*, vol. 34, no. 3, p. 034003, 2019.

[5] M. Pillai, S.-S. Kim, S. Ho, and S. Barnett, "Growth of $In_xGa_{1-x}As/GaAs$ heterostructures using Bi as a surfactant," *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena*, vol. 18, no. 3, pp. 1232–1236, 2000.

[6] K. Sakamoto, K. Kyoya, K. Miki, H. Matsuhata, and T. Sakamoto, "Which surfactant shall we choose for the heteroepitaxy of $Ge/Si(001)$? Questionable — Bi as a surfactant with small self-incorporation," *Japanese journal of applied physics*, vol. 32, pp. 204–206, 1993.

[7] E. Tournie, N. Grandjean, A. Trampert, J. Massies, and K. Ploog, "Surfactant-mediated molecular-beam epitaxy of III–V strained-layer heterostructures," *Journal of Crystal Growth*, vol. 150, pp. 460–466, 1995.

[8] A. Ichimiya and P. I. Cohen, *Reflection high-energy electron diffraction*. Cambridge University Press, 2004.

[9] R. A. Stradling, *Growth and Characterisation of Semiconductors: Papers Contributing to a Short Course Organised by the Departments of Materials and*

Physics with... at the Imperial College of Science, Technology and Medicine, University of London. Hilger, 1991.

[10] R. D. Richards, F. Bastiman, C. J. Hunter, D. F. Mendes, A. R. Mohmad, J. S. Roberts, and J. P. David, “Molecular beam epitaxy growth of GaAsBi using As_2 and As_4,” *Journal of Crystal Growth*, vol. 390, pp. 120–124, 2014.

[11] L. Däweritz and R. Hey, “Reconstruction and defect structure of vicinal GaAs (001) and $Al_xGa_{1-x}As$ (001) surfaces during MBE growth,” *Surface Science*, vol. 236, no. 1-2, pp. 15–22, 1990.

[12] C. Foxon and B. Joyce, “Interaction kinetics of As_4 and Ga on (100) GaAs surfaces using a modulated molecular beam technique,” *Surface Science*, vol. 50, no. 2, pp. 434–450, 1975.

[13] E. Tok, J. Neave, J. Zhang, B. Joyce, and T. Jones, “Arsenic incorporation kinetics in GaAs (001) homoepitaxy revisited,” *Surface Science*, vol. 374, no. 1-3, pp. 397–405, 1997.

[14] R. Fischer *et al.*, "Incorporation rates of gallium and aluminum on GaAs during molecular beam epitaxy at high substrate temperatures," *J. Appl. Phys.*, vol. 54, no. 5, pp. 2508–2510, May 1983, doi: 10.1063/1.332317.

[15] H. Morkoç and A. Y. Cho, "High-purity GaAs and Cr-doped GaAs epitaxial layers by MBE," *J. Appl. Phys.*, vol. 50, no. 10, pp. 6413–6416, Oct. 1979, doi: 10.1063/1.325732.

[16] B. A. Joyce, "Molecular beam epitaxy," *Rep. Prog. Phys.*, vol. 48, pp. 1637–1697, 1985.

[17] C. T. Foxon and B. A. Joyce, "INTERACTION KINETICS OF As4 AND Ga ON {100} GaAs SURFACES USING A MODULATED MOLECULAR BEAM TECHNIQUE," *Surf. Sci.*, pp. 434–450, Mar. 1975.

[18] M. Masnadi-Shirazi, D. A. Beaton, R. B. Lewis, X. Lu, and T. Tiedje, "Surface reconstructions during growth of GaAs1−xBix alloys by molecular beam epitaxy," *J. Cryst. Growth*, vol. 338, no. 1, pp. 80–84, Jan. 2012, doi: 10.1016/j.jcrysgro.2011.09.055.

[19] K. Oe and H. Okamoto, "New Semiconductor Alloy GaAs1−xBix Grown by Metal Organic Vapor Phase Epitaxy," *Jpn. J. Appl. Phys.*, vol. 37, no. Part 2, No. 11A, pp. L1283–L1285, Nov. 1998, doi: 10.1143/JJAP.37.L1283.

[20] S. Tixier *et al.*, "Molecular beam epitaxy growth of GaAs1−xBix," *Appl. Phys. Lett.*, vol. 82, no. 14, pp. 2245–2247, Apr. 2003, doi: 10.1063/1.1565499.

[21] E. C. Young, M. B. Whitwick, T. Tiedje, and D. A. Beaton, "Bismuth incorporation in GaAs1−xBix grown by molecular beam epitaxy with in-situ light scattering," *Phys. Status Solidi C*, vol. 4, no. 5, pp. 1707–1710, 2007, doi: 10.1002/pssc.200674291.

[22] R. D. Richards *et al.*, "Molecular beam epitaxy growth of GaAsBi using As_2 and As_4 ," *J. Cryst. Growth*, vol. 390, pp. 120–124, Mar. 2014, doi: 10.1016/j.jcrysgro.2013.12.008.

[23] A. J. Ptak *et al.*, "Kinetically limited growth of GaAsBi by molecular-beam epitaxy," *J. Cryst. Growth*, vol. 338, no. 1, pp. 107–110, Jan. 2012, doi: 10.1016/j.jcrysgro.2011.10.040.

[24] T. B. O. Rockett *et al.*, "Influence of growth conditions on the structural and opto-electronic quality of GaAsBi," *J. Cryst. Growth*, vol. 477, pp. 139–143, Nov. 2017, doi: 10.1016/j.jcrysgro.2017.02.004.

[25] T.-H. Kim, K. Forghani, K. Collar, T. F. Kuech, and A. S. Brown, "Growth of GaAs1−xBix by molecular beam epitaxy: Trade-offs in optical and structural characteristics," *J. Appl. Phys.*, vol. 116, no. 4, p. 043524, Jul. 2014, doi: 10.1063/1.4891874.

[26] J. C. M. Hwang, H. Temkin, T. M. Brennan, and R. E. Frahm, "Growth of high-

purity GaAs layers by molecular beam epitaxy," *Appl. Phys. Lett.*, vol. 42, no. 1, pp. 66–68, Jan. 1983, doi: 10.1063/1.93727.

[27] M. Heiblum, E. E. Mendez, and L. Osterling, "Growth by molecular beam epitaxy and characterization of high purity GaAs and AlGaAs," *J. Appl. Phys.*, vol. 54, no. 12, pp. 6982–6988, Dec. 1983, doi: 10.1063/1.332015.

[28] E. C. Larkins, E. S. Hellman, D. G. Schlom, J. S. Harris, M. H. Kim, and G. E. Stillman, "Reduction of the acceptor impurity background in GaAs grown by molecular beam epitaxy," *Appl. Phys. Lett.*, vol. 49, no. 7, pp. 391–393, Aug. 1986, doi: 10.1063/1.97597.

[29] D. J. Eaglesham, “Semiconductor molecular-beam epitaxy at low temperatures,” *Applied Physics*, vol. 77, no. 8, pp. 3597–3617, 1995, doi:10.1063/1.358597.

[30] M. R. Melloch, “Low-Temperature Grown III–V Materials,” *Annual Review of Materials Science*, vol. 25, pp. 547–600, 1995.

[31] C. H. Goo, W. S. Lau, T. C. Chong, L. S. Tan, and P. K. Chu, “High oxygen and carbon contents in GaAs epilayers grown below a critical substrate temperature by molecular beam epitaxy,” *Applied Physics Letters*, vol. 68, no. 6, pp. 841–843, 1996, doi:10.1063/1.116551.

[32] K. Regiński, J. Muszalski, V. V. Preobrazhenskii, and D. I. Lubyshev, “Static phase diagrams of reconstructions for MBE-grown GaAs(001) and AlAs(001) surfaces,” *Thin Solid Films*, vol. 267, pp. 54–57, 1995, doi:10.1016/0040-6090(95)06665-9.

[33] D. A. Woolf, M. A. Pate, D. J. Robbins, M. S. Skolnick, and L. J. Lever, “The molecular beam epitaxial growth of GaAs/GaAs(111)B: Doping and growth temperature studies,” *Journal of Applied Physics*, vol. 71, no. 10, pp. 4908–4915, 1992.

[34] A. Ohtake *et al*., “Atomic structure of the GaAs(001)-(2×4) surface under As flux,” (reports (2×4) observed in 480–600 °C under As flux), *Physical Review B* (2002).

Chapter 4: Characterization Techniques

This chapter provides a detailed overview of various characterization methods for material and device analysis. These techniques are categorized into three main types: structural, optical, and electrical measurements.

4.1 X-ray Diffraction (XRD)

X-ray diffraction (XRD) is a powerful, non-destructive technique widely used to study crystalline structures. It enables the analysis of properties such as epitaxial layer thickness, strain, relaxation, and lattice constants through X-ray diffraction patterns. X-rays, being high-energy electromagnetic waves, typically have wavelengths in the range of 0.07–0.2 nm [1].

4.1.1 Bragg's Law

When visible light passes through a double slit, a diffraction pattern can be observed on a screen. This phenomenon occurs only when the slit width is comparable to the wavelength of visible light (approximately 10^{-6} m).

For X-rays, the slit analogy applies to the atomic scale. If the slit width (or interatomic spacing) is significantly larger than the X-ray wavelength, diffraction does not occur. In 1912, Max von Laue hypothesized that the crystal lattice, with interatomic spacings comparable to X-ray wavelengths [2], could produce diffraction patterns. This hypothesis was later confirmed by William Lawrence Bragg and William Henry Bragg in 1913 [3].

Bragg's law, which describes this phenomenon, is expressed as:

$$2d \sin\theta = n\lambda \tag{4.1}$$

d: The distance between two adjacent layers, typically half the lattice constant since each unit cell contains two monolayers.

θ: The angle of the incident beam.

n: An integer value representing the diffraction order.

λ: The wavelength of the incoming X-rays.

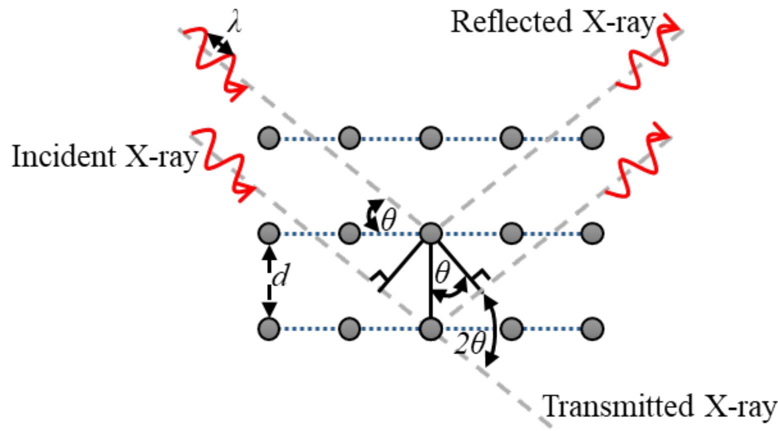


Figure 4.1: Bragg's Law in X-ray Diffraction

Figure 4.1 illustrates how a crystalline structure acts as a grating in X-ray diffraction. According to Bragg's law, constructive interference occurs when the path difference between two parallel beams equals an integer multiple of the wavelength. If this condition is not met, destructive interference occurs, and no signal is detected.

4.1.2 Design of Diffractometer

In this project, a Bruker D8 Discover X-ray diffractometer, which is a triple-axis XRD system, was employed. This design offers an improvement over the double-axis model due to the inclusion of an additional analyzer. The analyzer narrows the range of 2-Theta angles that can pass through it, thereby enhancing the resolution at the expense of reduced intensity.

Figure 4.2 illustrates the schematic setup. The X-ray beam is produced in an X-ray tube, where electrons are accelerated towards a copper target by a high voltage of approximately 30kV. A slit in the path of the X-ray beam helps to collimate it, affecting both the intensity and angular resolution of the beam. A narrower slit increases the intensity of the X-rays that strike the sample but reduces the angular resolution.

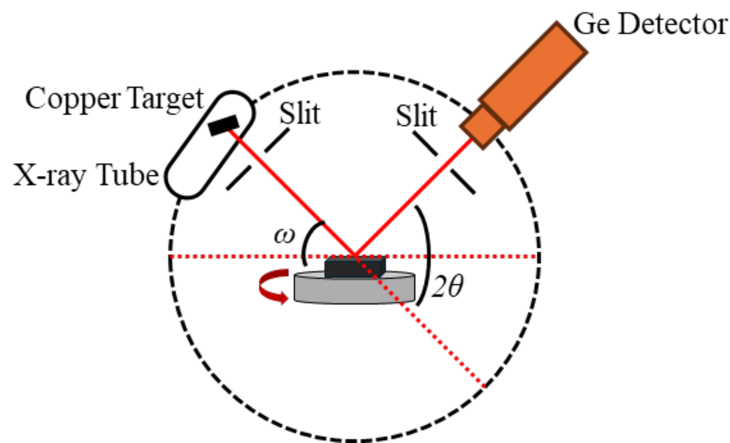


Figure 4.2 Schematic diagram of XRD measurement. Note: The monochromator and analyzer are omitted in this diagram for clarity.

The wafer is secured using a porous ceramic vacuum chuck that holds it in place during the test. This setup allows the stage to tilt vertically, enhancing the test's adaptability. Stepper motors provide precise control over the stage's movements, enabling three-axis motion as well as rotation in azimuth and altitude. This multidirectional capability allows for optimal positioning of the wafer to maximize X-ray intensity capture.

Once the X-rays have been diffracted by the wafer, they are detected by a germanium detector. It's important to note that the Bragg angle varies between different materials, making it essential to determine and set the appropriate Bragg angle for the material under examination prior to initiating measurements.

4.2 Nomarski Microscopy

In the 1950s, George Nomarski first introduced Differential Interference Contrast (DIC) technology in microscopy [5]. As a rapid and effective optical imaging technique, Nomarski microscopy can be used to inspect the wafer surface post-growth. Figure 4.3 will illustrate a schematic diagram showing the basic components of a Nomarski microscope.

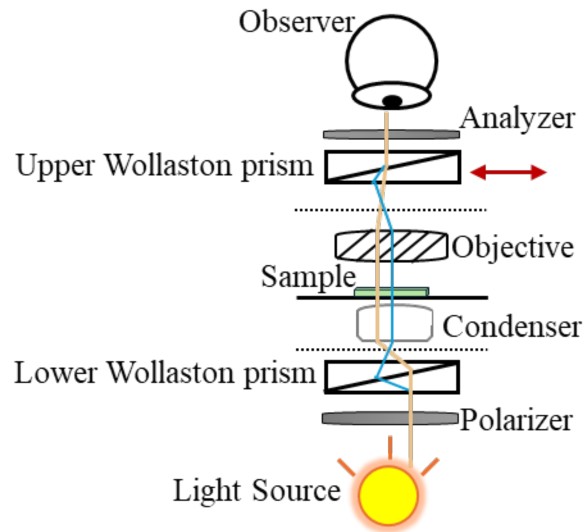


Figure 4.3 Basic Components in a Nomarski Microscope

Light is emitted from a hot-filament bulb and initially polarized at 45° after passing through a polarizing filter. The light then enters a Wollaston prism, which splits it into two rays polarized at 90° to each other, ensuring they do not interfere with one another. To avoid the creation of double images, these rays are separated by a very small "shear" distance, typically about 0.2 μm. Each ray passes through the sample at slightly different positions. If there are variations in the sample such as defects or thickness differences, the rays will traverse different optical paths.

These rays are then refocused by a condenser, pass through the sample, and are subsequently focused by an objective lens before being recombined by an upper Wollaston prism. At this point, the rays may undergo constructive or destructive interference depending on the differences in their optical paths, enhancing the contrast of the resulting image.

This technique allows for a preliminary inspection of the sample's epilayer quality. Nomarski microscopy not only provides detailed information about the surface condition of the epilayer, such as droplets and defects, but also qualitatively assesses strain relaxation through the presence of cross-hatching.

4.3 Photoluminescence (PL)

Photoluminescence spectroscopy is a sensitive and non-destructive optical method used to determine the bandgap and other optical properties of semiconductors. The absorption of light in a semiconductor is influenced by energy levels associated with defects or localized states. To explore these effects, the incident laser power and the temperature of the device can be adjusted, which helps to reveal more about the semiconductor's properties. Additionally, PL serves as an indicator of the optical quality of the sample.

4.3.1 Radiative Recombination

For photon absorption to occur in a semiconductor, the energy of the incoming photon must exceed the material's bandgap.

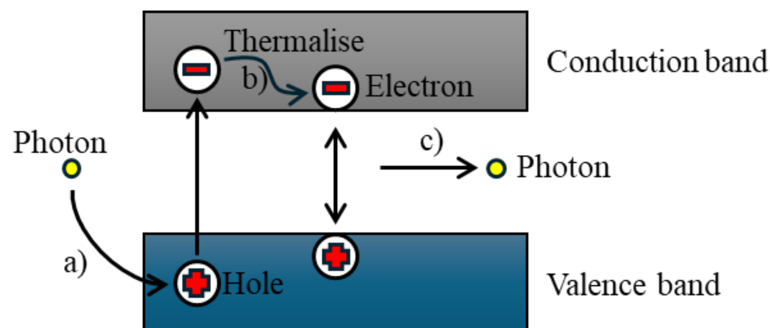


Figure 4.4: The Processes Involved in Photoluminescence (PL) a) An incoming photon excites an electron from the valence band to the conduction band, leaving a hole in the valence band. b) The electron in the conduction band thermalizes to the lower energy edge of the conduction band. c) During the recombination of the electron with the hole, a photon is emitted with the same energy as the material's bandgap.

When light is incident on a semiconductor and its energy exceeds the material's bandgap, the photon is absorbed, elevating an electron from the valence band to the conduction

band. Once in the conduction band, the excess energy of the electron is dissipated as it moves to the lower energy edge, a process known as thermalization. This electron eventually recombines with a hole in the valence band, resulting in the emission of a photon with energy equivalent to the bandgap of the semiconductor. This sequence of events, characteristic of direct bandgap materials, efficiently facilitates the recombination process. In contrast, indirect bandgap materials require the involvement of phonons for recombination. Figure 4.4 illustrates these photoluminescence processes schematically.

4.3.2 Non-radiative Recombination

In addition to band-to-band recombination, non-radiative recombination represents another critical process within semiconductors. This type of recombination typically arises due to the presence of doping-induced defects, impurities, and dislocations within the forbidden energy gap of the semiconductor. Simplified diagrams illustrating these mechanisms are depicted in Figure 4.5 and Figure 4.6.

In Figure 4.5, the non-radiative recombination process occurs through trap states located within the semiconductor's forbidden energy gap, known as Shockley-Read-Hall (SRH) recombination [6]. In one scenario, electrons transition to these trap states before moving to an empty state in the valence band, completing the recombination process. This interaction within the trap states facilitates energy exchange, often leading to lattice vibrations, such as phonons. In a different scenario, both an electron and a hole simultaneously meet in the trap, leading to recombination.

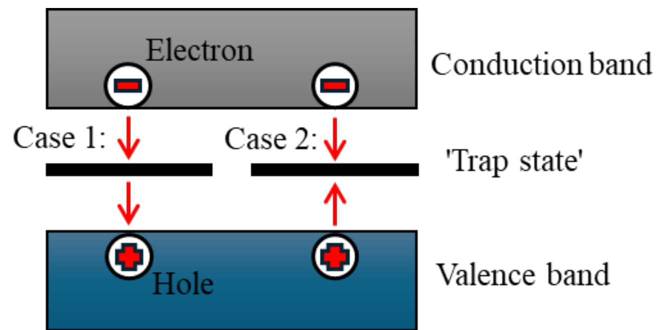


Figure 4.5 Shockley Read-Hall Recombination via Trap State in the Band Gap.

Case 1: Transition of an electron from the conduction band to the valence band through two steps.

Case 2: Both the electron and the hole migrate to the trap state and then recombine.

Auger recombination represents another type of non-radiative recombination process. This mechanism operates similarly to band-to-band recombination but involves energy transfer to a third carrier instead of releasing a photon. Figure 4.6 illustrates two primary forms of Auger recombination: CHCC and CHSH Auger recombination.

CHCC Auger Recombination: An electron in the conduction band recombines with a hole in the heavy hole band. The energy released from this recombination is transferred to another electron, elevating it to a higher energy state within the conduction band.

CHSH Auger Recombination: Similar to CHCC, but instead of an electron, the energy released excites a hole from the valence band to the spin-orbit split-off band.

In both CHCC and CHSH processes, energy and momentum conservation principles are adhered to, ensuring the conservation laws are not violated during the process.

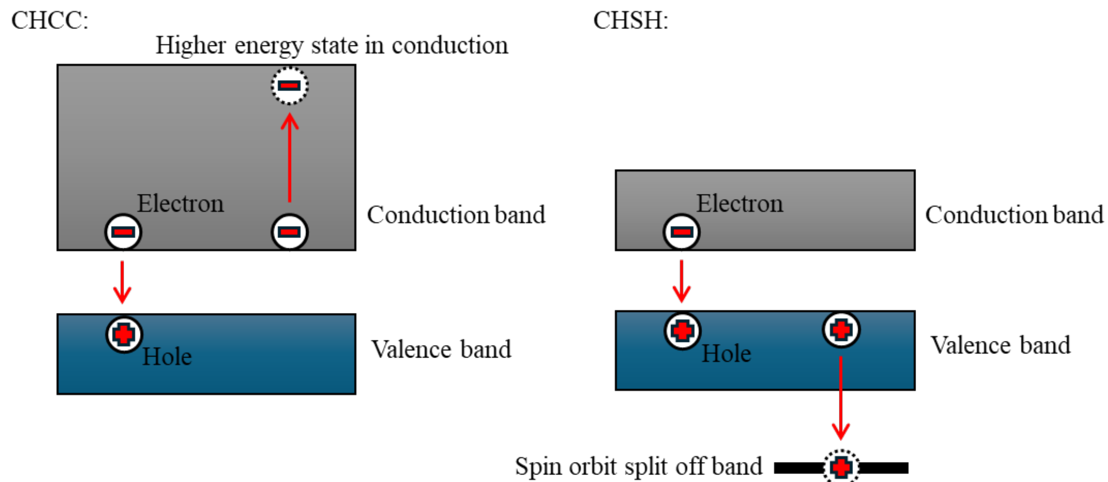


Figure 4.6: Schematic illustration of CHCC (Conduction band to Heavy hole valence band, Conduction band to higher energy state of Conduction band) and CHSH (Conduction band to Heavy hole valence band, Spin orbit split-off band to Heavy hole valence band) Auger recombination processes.

4.3.3 Experimental Setup

The photoluminescence (PL) setup used in this work consists of the following components:

A 532 nm laser,

A chopper,

A Jobin-Yvon Triax 550 monochromator,

A liquid nitrogen-cooled germanium detector, and

A Stanford Research SR830 lock-in amplifier (LIA).

For photon absorption to occur, the laser beam's energy must exceed the bandgap of the semiconductor sample. When using a 532 nm laser, a second-order effect produces an extra peak at the wavelength of 1064 nm. To address this, a long-pass filter is employed to minimize background light and suppress the second-order effect.

A Cassegrain telescope lens collects and focuses emitted photons through the

monochromator slit. A neutral density (ND) filter can also be placed in front of the monochromator to reduce the intensity of the photon beam, ensuring the detector is not saturated.

The monochromator is equipped with three diffraction gratings to cover a wide wavelength range. In this study, a 900 grooves/mm grating with a blaze wavelength of 850 nm is used, as the wavelength of interest for GaAsBi lies between 850 nm and 1400 nm.

The germanium detector collects the emitted photon signals, and their amplitude is displayed on the LIA. To avoid interference from ambient environmental light (e.g., fluorescent lighting operating at 50 Hz mains), the laser light is modulated at 180 Hz using a chopper. A schematic of the setup is provided in Figure 4.7.

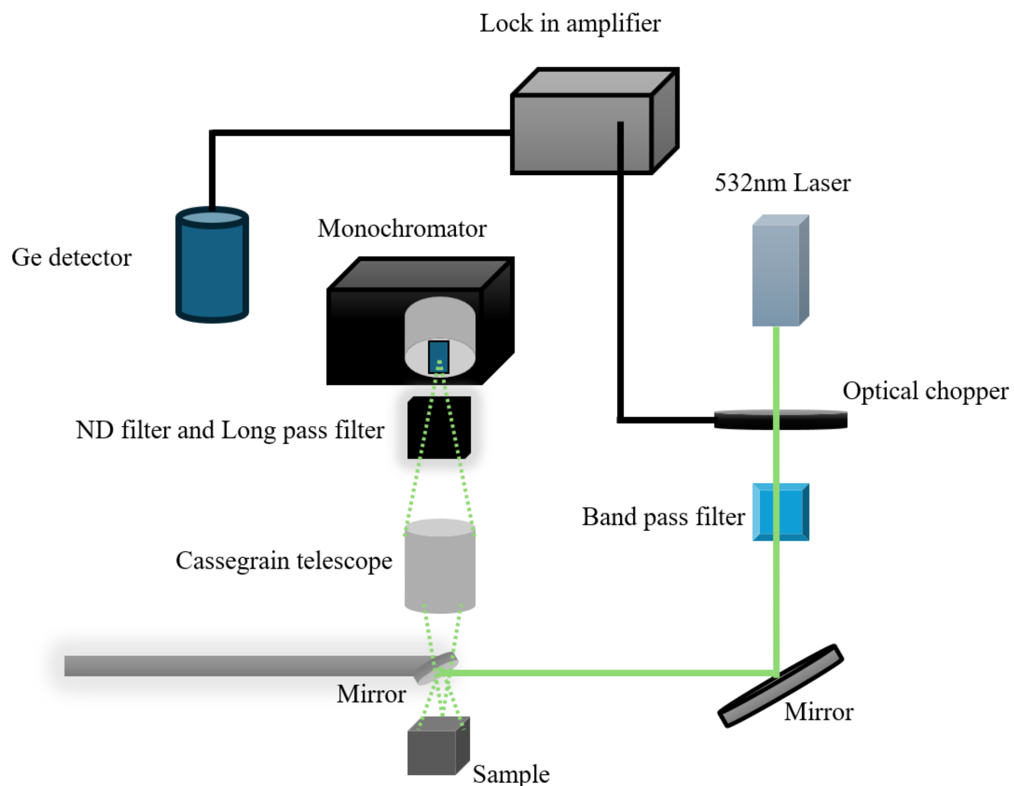


Figure 4.7: The Photoluminescence Setup Used in This Work

4.3.4 Monochromator

The Triax 550 monochromator employed in this study is designed based on the common Czerny–Turner model. This design includes two slits, two mirrors, and a grating. The operational principle of the monochromator is detailed in Figure 4.8.

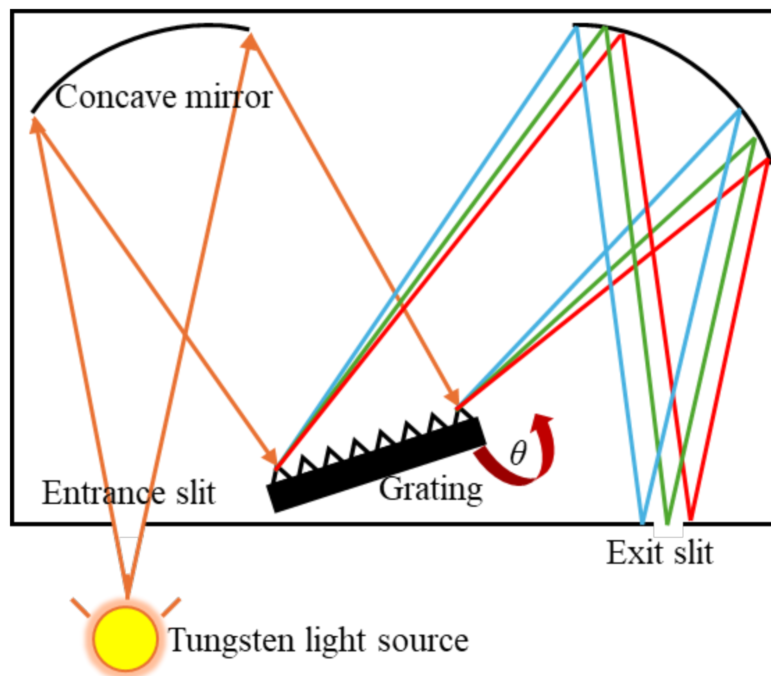


Figure 4.8: Schematic Diagram of a Horiba iHR320 Monochromator

The monochromator utilizes a tungsten bulb as its light source, emitting polychromatic light that encompasses a range of colors. This light first passes through the entrance slit and then reflects off a concave mirror. The primary function of this mirror is to collimate the light, directing it towards the diffraction grating. This grating is characterized by its finely etched grooves set at periodic intervals, and it's coated with a highly reflective material to enhance its reflectivity.

Due to the principles of constructive and destructive interference, the polychromatic light is dispersed by the grating; each wavelength is refracted at a different angle. A second concave mirror focuses the light before it exits through the slit. The grating's angle, θ, is critical as it dictates which wavelength of light is selected to pass through

the exit slit, while other wavelengths are absorbed by the internal black-painted walls of the monochromator.

To optimize the throughput for specific wavelengths, the monochromator is fitted with three different gratings. Each grating is categorized by its groove density and specified blaze wavelength. For the study of GaAsBi material, where the cutoff wavelength is around 1300nm, a 600 grooves/mm grating with a blaze wavelength at 1000nm is employed. This grating effectively handles wavelengths between 667nm and 1500nm, adhering to the $\frac{2}{3} \times \lambda_{\text{blaze}}, \frac{3}{2} \times \lambda_{\text{blaze}}$ rule for usable wavelength range.

4.4 Current-Voltage (IV) Measurement

Following device fabrication, the first step is to conduct a current-voltage (IV) characteristic test of the device. This test involves applying a bias across the diode to generate a current feedback, which helps in assessing whether the diode adheres to expected diode characteristic curves. This IV measurement serves as a preliminary evaluation tool.

Figure 4.9 illustrates the schematic diagram of the IV setup. In this configuration, the device under test (DUT) is connected in series with either a HP4140B picoammeter or a curve tracer. The curve tracer aids in manually checking the device connections and provides insights into determining the breakdown voltage of the diode. Alternatively, a Keithley 236/237 source-measure unit (SMU) can be used in place of the picoammeter. Typically, IV measurements are conducted under dark conditions to avoid interference from ambient light.

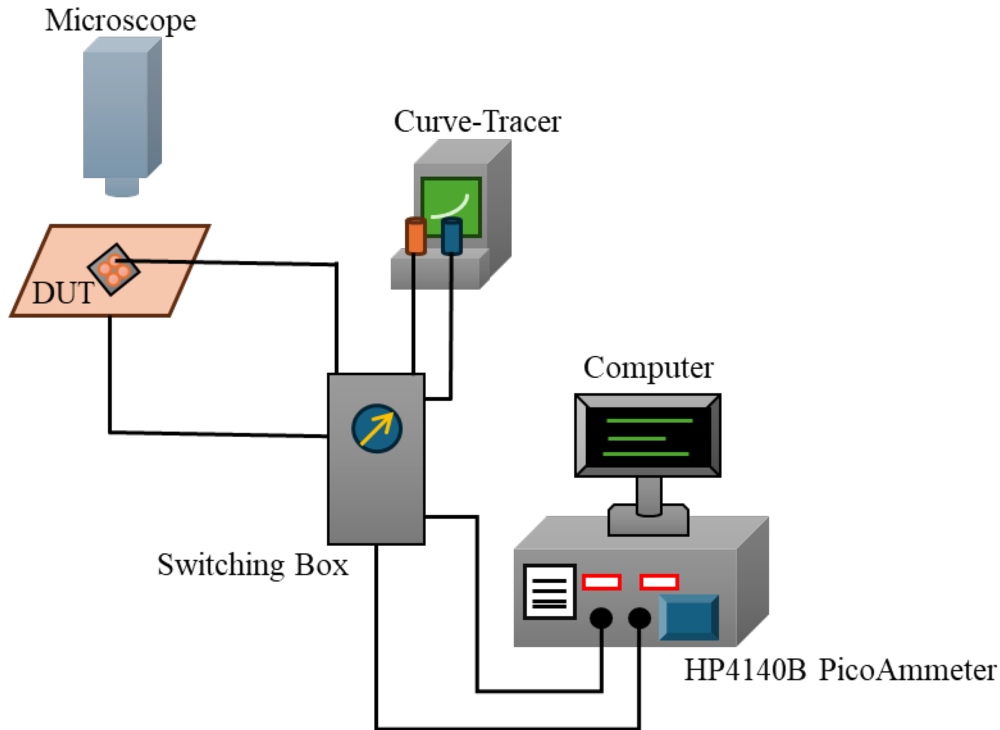


Figure 4.9: Schematic Diagram of the I-V Measurement Setup

Dark current is defined as the current measured in a photodiode under no incident light, regardless of the bias direction. This current depends on the material and the size of the active area and is sensitive to temperature changes, approximately doubling for every 10°C increase. It obeys the Shockley diode equation:

$$J = J_0\left(\exp\left(\frac{qV}{n_0 kT}\right) - 1\right) \tag{4.2}$$

J_0: Saturation current density

q: Electron charge

k: Boltzmann constant

T: Temperature

n_0: Ideality factor, indicating whether the current is predominantly governed by diffusion (1) or generation-recombination mechanisms (2).

Under forward bias, series resistance affects the current, especially as the voltage increases, necessitating inclusion of a resistance parameter R in the equation:

$$J = J_0\left(\exp\left(\frac{q(V-JR)}{n_0 kT}\right)-1\right) \tag{4.3}$$

Series resistance primarily arises from the contact resistance between the metal and the semiconductor, and it can limit the maximum achievable gain in an Avalanche Photodiode (APD). Optimizing metal contacts and annealing processes are essential to minimizing this resistance.

The reverse bias behavior is crucial for evaluating APD performance due to two primary factors: avalanche breakdown near the breakdown voltage and leakage current measurement, which influences the signal-to-noise ratio (SNR). The total dark current under reverse bias consists of bulk (I_{bulk}) and surface (I_{surf}) currents, where their contributions are assessed by:

$$J_{Total} = \frac{I_{bulk}}{A_{bulk}} + \frac{I_{surf}}{A_{surf}} \tag{4.4}$$

Surface current often stems from surface defects and requires passivation to minimize it by terminating the surface dangling bonds.

Surface leakage current often stems from surface defects introduced during mesa definition and is therefore commonly mitigated by surface passivation, which terminates dangling bonds and suppresses surface-state-assisted conduction paths.

Under reverse bias, the depletion region widens and the junction barrier suppresses majority-carrier injection; consequently, the measured reverse leakage current is dominated by defect-assisted carrier generation within the space-charge region and by any residual surface leakage paths. A key bulk contribution is Shockley–Read–Hall (SRH) generation–recombination via deep trap states within the band gap. In the depletion region, traps can alternately capture and emit carriers, giving net recombination under forward bias (often associated with an ideality factor approaching 2 when depletion-region recombination dominates) and net thermal generation under reverse bias. A convenient expression for the SRH current density is [6,10]

$$J_{SRH} = \frac{q\, n_i\, W}{2\, \tau_{\mathrm{eff}}}\left[\exp\left(\frac{qV}{2k_B T}\right) - 1\right] \tag{4.5}$$

Here, q is the electron charge, n_i is the intrinsic carrier concentration, W is the depletion width, and τ_{eff} is an effective lifetime determined by the density and capture properties of the relevant trap states. In reverse bias ($V < 0$), the bracketed term tends to (-1), so the magnitude of the SRH generation current is approximately $|J_{\mathrm{SRH}}| \approx qn_iW/(2\tau_{\mathrm{eff}})$; i.e., the reverse-bias dependence is primarily through $W(V)$, while τ_{eff} captures the impact of trap density and capture cross sections. At sufficiently high reverse bias, direct band-to-band tunnelling can also contribute to leakage current; however, for the bias range considered here it is expected to be negligible and is not discussed further.

4.5 Capacitance-Voltage (CV) Measurement

For room temperature capacitance-voltage (C-V) measurements, a Hewlett-Packard 4275A multi-frequency LCR meter is utilized. This setup involves a circuit where a DC voltage is applied to the device under test (DUT) and superimposed with an AC voltage. The AC voltage helps determine the impedance and phase angle of the device, though it must be small in amplitude compared to the DC voltage to accurately assess the impedance, which is composed of both resistive and reactive (real and imaginary) parts. The reactive component, or reactance, is defined as $\frac{1}{j\omega C}$, linking directly to the frequency of the AC signal.

For effective C-V measurements, selecting an appropriate AC frequency is crucial based on the capacitance value of the device. Typically, capacitance values are in the pico-Farad range, producing an AC signal that is hard to detect due to its small size. Thus, a high frequency is essential to minimize the reactance and enhance the detectability of the AC component; in this setup, 1 MHz is employed as the testing frequency.

To measure impedance accurately using the LCR meter, two models are considered: series or parallel. Figure 4.10 details the equivalent circuits for these configurations and provides a guideline for choosing between them based on the capacitance value and the frequency used. In the series model, a smaller AC signal passes through when the reactance is high, making the parallel model more suitable for larger reactances. Conversely, the series model is preferred when reactance is low. Due to the low capacitance values typical of our devices, the parallel model was selected for optimal measurement accuracy.

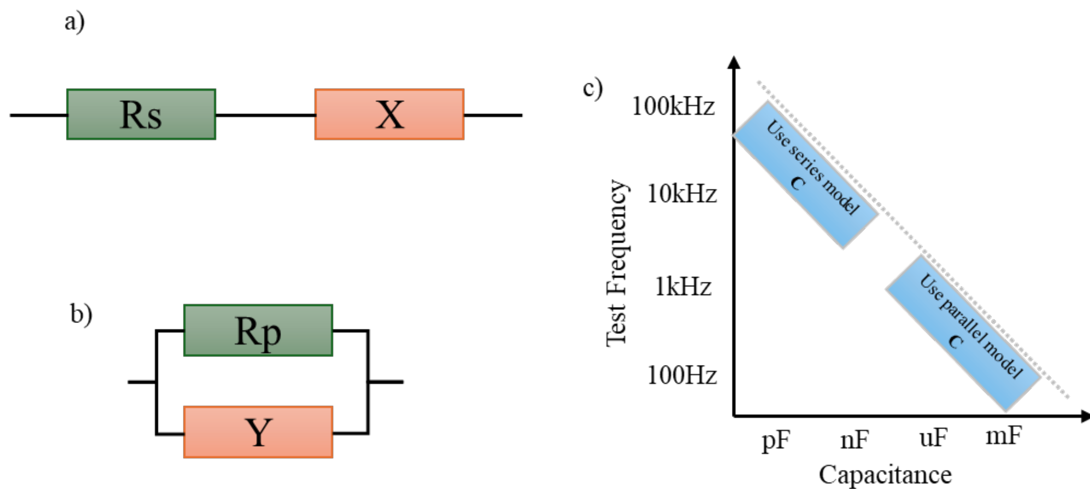


Figure 4.10: a) Equivalent Circuit for Series Model - Showcases the series resistance R_s. b) Equivalent Circuit for Parallel Model - Illustrates the parallel resistance R_p and the reactance x. c) Measurement Guidelines for Capacitance - Provides best practices for capacitance measurement using a standard LCR meter.

The quality factor (Q factor) is crucial for assessing the performance of a capacitor, defined as the ratio between the reactance (X_C) and the series resistance (R_C) of the capacitor, represented mathematically as:

$$Q = \tan\,\theta = \frac{|X_C|}{R_C} = \frac{1}{\omega C R_C} \tag{4.6}$$

Here, θ denotes the phase angle between the current and voltage, ω is the angular

frequency, and C represents the capacitance. The Q factor illustrates the ratio between the energy stored in the capacitor and the energy dissipated as heat through the series resistance. In an ideal capacitor, where no energy is lost, the Q factor would be infinite, aligning with a 90° phase difference typically observed between the current and voltage in capacitors.

In this study, CV measurements were conducted at room temperature with phase angles recorded between 80° and 90°. These measurements help verify whether the bulk properties of the junction capacitance scale correctly with the device area.

CV measurements often contend with two significant errors: offset and gain errors. An illustration of these errors on a logarithmic scale is shown in Figure 4.11, where the x-axis indicates the actual capacitance value and the y-axis displays the measured capacitance. Ideally, if the measurement system were perfect, the results would appear as a straight line at a 45° angle. However, in practice, both gain and offset errors must be considered and corrected during the measurement process.

To minimize the offset error, particularly caused by parasitic capacitances in the probe and cable, corrections such as open-correction for small capacitances (around 10pF) or short-correction for larger capacitances (around 10nF) should be applied before testing. Gain error, which varies with the capacitance magnitude, requires adjustments through load correction. This involves measuring a known standard load to establish a reference, and then calculating the ratio between the measured and actual values to determine the true capacitance.

However, load correction comes with limitations; the selected load must closely match the capacitance range of the device being measured to ensure accuracy.

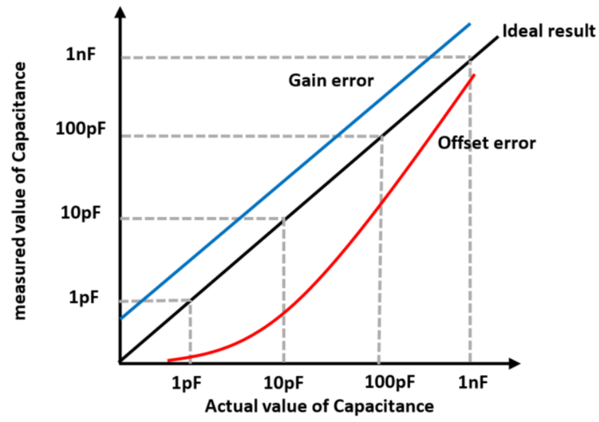


Figure 4.11: Common Errors in Capacitance Measurement

This figure illustrates the typical errors encountered in capacitance measurement. The ideal result is represented by the black line, while the offset error and gain error are depicted by the red line and blue line, respectively, as adapted from reference [9].

4.6 Capacitance-Voltage (CV) Measurement Insights

For a p–i–n diode, the interfaces between the highly doped claddings and the lightly doped (or nominally intrinsic) region can, to first order, be treated as abrupt junctions because of the large carrier-concentration discontinuity. Under the depletion approximation, the small-signal junction capacitance is related to the depletion width by

$$C = \frac{\varepsilon_0 \varepsilon_r A}{W}, \quad (4.7)$$

where ε_0 is the vacuum permittivity, ε_r is the relative permittivity, A is the device area, and W is the depletion width. For an abrupt junction with a uniformly doped lightly doped side of concentration N_b, the depletion width satisfies

$$W = \sqrt{\frac{2\varepsilon_0 \varepsilon_r (V_{\mathrm{bi}} - V)}{q N_b}},$$

giving

$$C = A \sqrt{\frac{q \varepsilon_0 \varepsilon_r N_b}{2(V_{\mathrm{bi}} - V)}}. \quad (4.8)$$

Here, q is the electron charge, V is the applied bias (with reverse bias increasing the net potential drop across the junction), and V_{bi} is the built-in voltage. Squaring Eq.~(4.8) yields the standard linear relation

$$\frac{1}{C^2} = \frac{2(V_{\mathrm{bi}} - V)}{q \varepsilon_0 \varepsilon_r N_b A^2}. \quad (4.9)$$

Therefore, in the voltage range where the depletion region extends through a region of approximately uniform doping, a plot of $1/C^2$ versus V is expected to be linear. The built-in voltage V_{bi} can be estimated by extrapolating this linear fit to $1/C^2 = 0$, while the slope provides the doping density. Equivalently, the doping concentration associated with the depleted region can be extracted from the differential form

$$N_b = -\frac{2}{q \varepsilon_0 \varepsilon_r A^2} \frac{dV}{d(1/C^2)}. \quad (4.10)$$

This expression reduces to a constant N_b when the $1/C^2$–V curve is linear, and yields an apparent doping profile when the doping varies with depth~[10].

Accurate extraction of the intrinsic-region thickness and the cladding/background doping in a multilayer p–i–n heterostructure generally requires going beyond the single-abrupt-junction approximation. In particular, once the space-charge region traverses layers of different doping densities (and potentially different permittivities), the measured $C(V)$ no longer follows a single straight-line $1/C^2$–V relation. In this work, the measured C–V characteristics were therefore analysed using a one-dimensional electrostatic model based on Poisson's equation, in which the layer stack is represented by piecewise-constant doping densities and the depletion width $W(V)$ is obtained self-consistently from the applied bias and boundary conditions. The corresponding model capacitance is then computed via $C(V) = \varepsilon_0 \varepsilon_r A / W(V)$, enabling simultaneous fitting of the intrinsic-region thickness and the relevant doping parameters. The extracted profiles are discussed in Sec. 5.5.2.

4.7 Photomultiplication

The electron and hole ionization coefficients of a material can be determined from gain measurements. Figure 4.12 illustrates the experimental setup for these measurements. In this setup, a laser serves as the light source. A monochromator is not used because the laser light is inherently monochromatic.

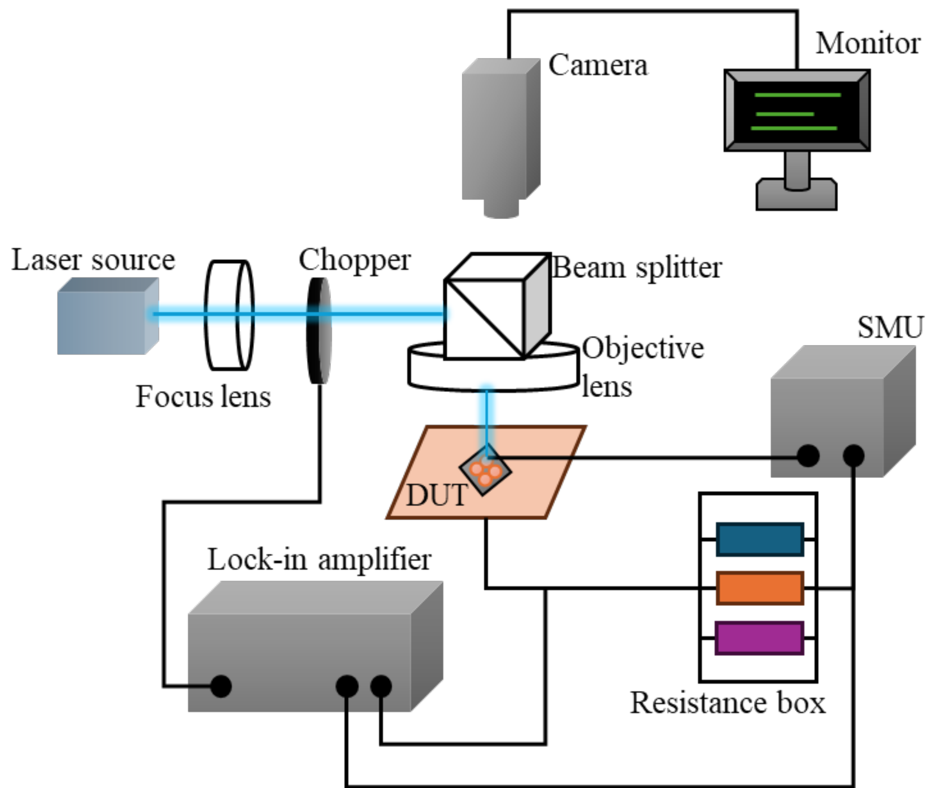


Figure 4.12: Schematic Diagram of the Photomultiplication Setup

This diagram illustrates the setup used to measure photocurrent versus reverse bias. Different wavelength lasers can be utilized depending on the requirements of the experiment.

Photomultiplication is evaluated using a setup that involves focusing laser light onto a small area of the sample, controlled by an optical chopper at 180 Hz and a Keithley 236/237 source-measure unit (SMU) for reverse biasing. High leakage currents, characteristic of the device, necessitate the use of a lock-in amplifier (LIA) connected to a resistance box to record photocurrents accurately.

The significant leakage current can influence multiplication measurements, particularly at high voltage biases, due to a voltage drop across the resistor. This drop causes the actual voltage across the device to be lower than the applied bias. To minimize this, resistors are chosen to ensure the voltage drop does not exceed 0.1 volts, maintaining the measurement accuracy independent of the optical laser power and preventing

potential heating effects.

For precision, power-dependent measurements use a variable neutral density filter. Each device size undergoes at least three repetitions of the measurement process to confirm gain uniformity. The process includes adjusting the laser spot's position across the device to account for non-uniform etching effects and their impact on multiplication results.

Avalanche Gain Calculation

The avalanche gain in these measurements is calculated using the formula:

$$M(V) = \frac{I_{pi}(V)}{I_{pr}} \tag{4.11}$$

where $M(V)$ is the avalanche gain as a function of reverse bias, $I_{pi}(V)$ is the total photocurrent, and I_{pr} is the primary current, also known as the light-produced initiating current.

Photocurrent Modeling Before Multiplication

It is recognized that photocurrent increases with reverse bias due to the expansion of the depletion region. Woods et al. [11] suggest modeling the onset of multiplication by solving the current diffusion equation for an abrupt p-n junction:

$$I_{pr} = \frac{qG_0}{\cosh\,(L/L_{dif})} \tag{4.12}$$

where G_0 is the generation rate per unit area of hole-electron pairs at the cladding layer, L is the distance from the depletion region edge to the sample surface, and L_{dif} is the minority carrier diffusion length. For p-i-n structures where light absorption occurs near the cladding layer, if the linear increase of the depletion region into the cladding layer with reverse bias voltage is considered, the model simplifies to:

$$I_{pr} = aV + b \tag{4.13}$$

where a and b are constants. This linear model helps fit the primary current, correcting for the increased efficiency of injected carriers with bias. Figure 4.13 shows

the fit and the resultant photocurrent modelling using this approach.

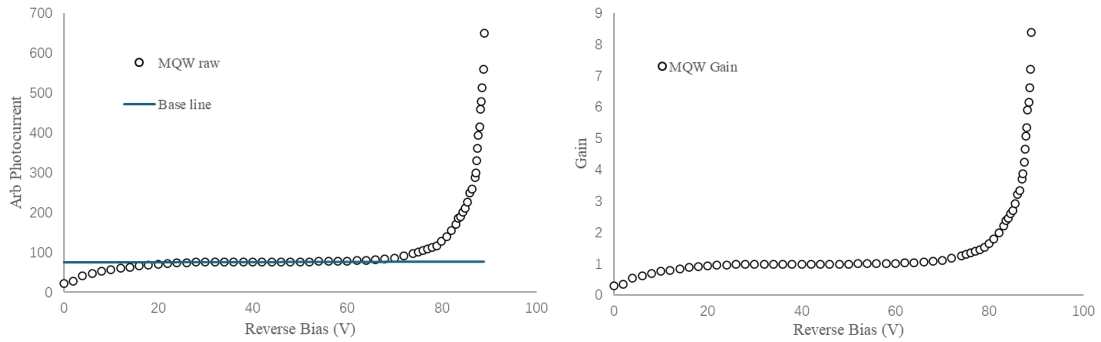


Figure 4.13: Baseline Correction and Gain Calculation for 3100nm GaAsBi p-i-n MQWs Diode

To determine both the hole and electron multiplication factors, M_h and M_e, Stillman and Wolfe have outlined several key experimental procedures [12]. Initially, to achieve pure electron and hole injection, devices with p-i-n and n-i-p structures of consistent intrinsic region thickness and doping concentration are utilized. An appropriate laser wavelength is selected based on the material and thickness of the cap layer to ensure that over 80% of the light is absorbed, facilitating pure injection.

Second, the primary photocurrent needs to be measured accurately before the onset of multiplication to establish a reliable baseline. Third, the gradient of the electric field across the intrinsic region should be gradual, which is generally controlled by adjusting the doping concentrations in the P and N regions.

Fourth, the electric field profile of the device is determined by analyzing the capacitance-voltage (C-V) characteristics, which help assess whether the electric field distribution is uniform or tapered across the intrinsic region.

Lastly, the uniformity of the avalanche gain across the device must be verified by conducting multiplication tests at various points on the device, as consistent avalanche

behavior is crucial for reliable device performance.

4.8 Photocurrent

The setup for measuring photocurrent is detailed in Figure 4.14. This setup utilizes a 100 W white tungsten lamp as the light source, which is spectrally analyzed through a monochromator managed by a computer software. The light of a chosen wavelength is then concentrated at the monochromator's exit slit and mechanically chopped at a frequency of 180 Hz. This chopped signal is transmitted as a phase locking signal to a lock-in amplifier (LIA). The diode involved in the setup is connected in series with a load resistor, while the LIA is arranged in parallel to this resistor. This ensures that only the voltage drop across the resistor is assessed by the LIA. Depending on the dark current readings, various resistor values might be employed during the experiments. The phase-sensitive detection feature of the LIA enhances the photocurrent signals while simultaneously minimizing noise and dark current to nanoampere levels, ensuring that only the signal due to the photocurrent is processed and displayed on the control PC.

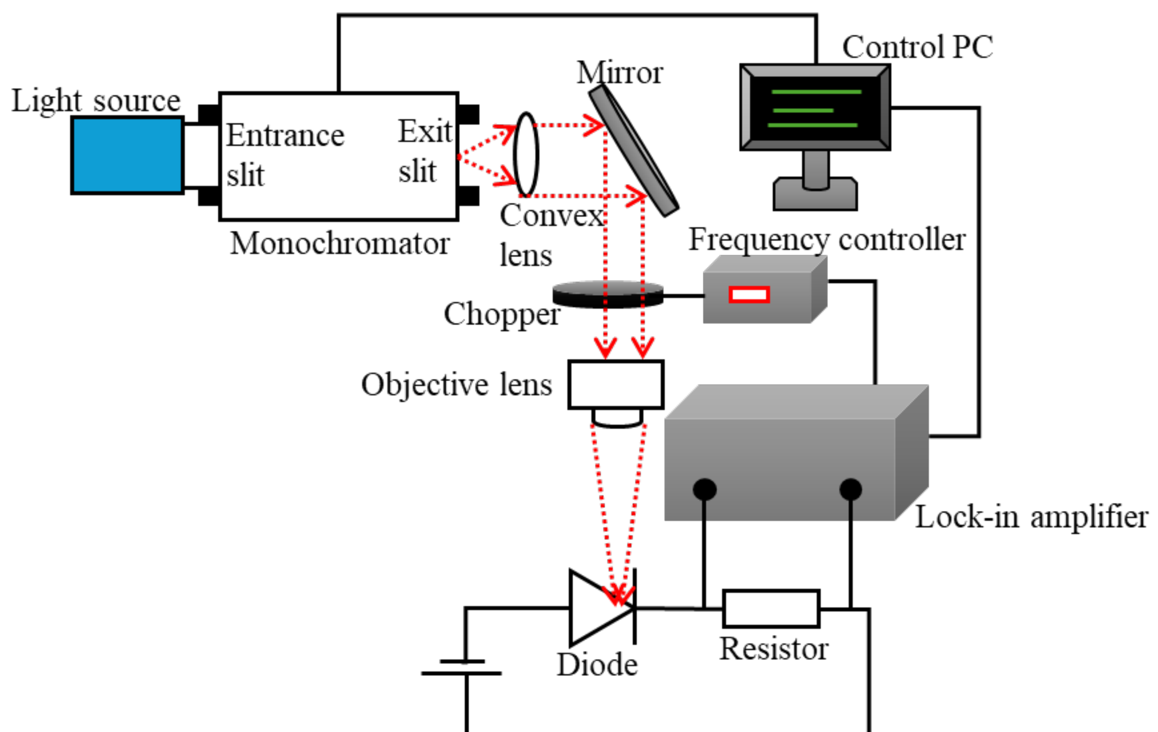


Figure 4.14: Schematic diagram of the photocurrent measurement setup.

The arbitrary units displayed in the photocurrent spectrum on the control PC represent the conversion of the LIA's voltage drop readings across the resistor. The sensitivity settings on the LIA define the precision and accuracy of the measured photocurrent. The relationship between the LIA's conversion value and the diode's voltage drop is presented as follows:

$$\text{LIA reading}(V) = \frac{\text{conversion} \times \text{Sensitivity}}{10} \tag{4.14}$$

$$I_{ph} = \frac{\text{LIA reading}(V)}{0.45 \times R} \tag{4.15}$$

From the photocurrent data, the responsivity R of the devices can be calculated, which is a measure of detection efficiency and is highly valued in photovoltaic device design. This responsivity is derived by normalizing the photocurrent to the incident light's power at a specific wavelength:

$$R = \frac{I_{ph}}{P} \tag{4.16}$$

For each wavelength, the actual power P of the monochromatic light impacting the device needs to be ascertained since its intensity differs from the light source due to losses in the monochromator system. A commercial photodiode (PD) is utilized to measure each wavelength's power, where the photocurrent $I_{ph}(PD)$ and the responsivity $R(PD)$ from the datasheet are employed to compute the power:

$$P_{PD} = \frac{I_{ph}(PD)}{R(PD)} \tag{4.17}$$

Given that the optical signal is mechanically chopped and the photocurrent is detected using a lock-in amplifier, the power used for responsivity normalisation must be consistent with the lock-in definition of amplitude. For a 50% duty-cycle chopped beam that switches between 0 and P_{PD}, the first-harmonic (fundamental) component at the chopping frequency has a sinusoidal peak amplitude of $(2/\pi)P_{PD}$ (i.e., $2/\pi \approx 0.64$). The lock-in amplifier reports the RMS value of this sinusoidal component [13],

introducing an additional factor of $1/\sqrt{2}$ (i.e., $1/\sqrt{2} \approx 0.7$). Consequently, the effective incident optical power associated with the lock-in-measured photocurrent is

$$P = \frac{1}{\sqrt{2}} \cdot \frac{2}{\pi} P_{\mathrm{PD}} \approx 0.7 \times 0.64 \times P_{\mathrm{PD}}, \tag{4.18}$$

$$P \approx 0.45 \times P_{\mathrm{PD}}. \tag{4.19}$$

This definition ensures that P is expressed in the same RMS (fundamental) convention as the lock-in-derived photocurrent used in Eq. (4.16), so that the calculated responsivity is internally consistent.

References

[1] D. K. Bowen and B. K. Tanner,
High Resolution X-Ray Diffractometry AndTopography. CRC Press, 1998.

[2] M. Eckert, “Max von Laue and the discovery of X-ray diffraction in 1912,” *Ann. Phys.,* vol. 524, no. 5, pp. A83–A85, May 2012, doi: 10.1002/andp.201200724.

[3] “The reflection of X-rays by crystals,” p. 11.

[4] Gymnos, “Modern Physics: Mastering Physics: Bremsstrahlung RadiationConceptual Question,” *Modern Physics,* Feb. 13, 2011.http://aeroja.blogspot.com/2011/02/mastering-physics-bremsstrahlung.html

[5] M. Pluta, “Nomarski’s DIC microscopy: a review,” Warsaw, Poland, May 1994, pp. 10–25, doi: 10.1117/12.171873.

[6] W. Shockley and W. T. Read, “Statistics of the Recombinations of Holes and Electrons,” *Phys. Rev.,* vol. 87, no. 5, pp. 835–842, Sep. 1952, doi: 10.1103/PhysRev.87.835.

[7] W. Shockley, “The theory of p-n junctions in semiconductors and p-n junction transistors,” *Bell Syst. Tech. J.,* vol. 28, no. 3, pp. 435–489, Jul. 1949, doi: 10.1002/j.1538-7305.1949.tb03645.x.

[8] B. T. Marozas, W. D. Hughes, X. Du, D. E. Sidor, G. R. Savich, and G. W. Wicks, “Surface dark current mechanisms in III-V infrared photodetectors [Invited],” *Opt. Mater. Express,* vol. 8, no. 6, p. 1419, Jun. 2018, doi: 10.1364/OME.8.001419.

[9] L. Stauffer and K. Instruments, “C-V Measurement Tips, Tricks, and Traps,” p. 14.

[10] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices.* John Wiley & Sons, 2006.

[11] M. H. Woods, W. C. Johnson, and M. A. Lampert, “Use of a Schottky barrier to measure impact ionization coefficients in semiconductors,” *Solid-State Electron.,* vol.

16, no. 3, pp. 381–394, Mar. 1973, doi: 10.1016/0038-1101(73)90013-0.

[12] W. T. Tsang, *Semiconductors and Semimetals: Lightwave CommunicationsTechnology v. 22A.* Orlando: Academic Press Inc, 1985.

[13] Stanford Research Systems, “Model SR830 DSP Lock-In Amplifier: Operating and Programming Manual”, Revision 2.5, Oct. 2011.

Chapter 5: Characterisation of GaAs/GaAsBi Heterostructures

5.1 Introduction

Semiconductor-based Avalanche Photodiodes (APDs) are commonly employed instead of standard photodiodes when the available photon flux is limited, as they enhance the sensitivity of optical systems [1]. APDs improve the Signal-to-Noise Ratio (SNR) by means of internal multiplication (M), which arises from the impact ionization of optically generated carriers in a semiconductor under high electric fields. However, the resulting gain is generally accompanied by an additional "excess" noise due to the stochastic nature of the impact ionization process. In 1966, McIntyre introduced the excess noise factor (F), which depends on M as follows [2]:

$$F(M) = kM + (1-k)\left(2 - \frac{1}{M}\right), \tag{5.1}$$

where k is the ratio of the impact ionization coefficients for holes (β) and electrons (α), denoted $k = \beta/\alpha$. This excess noise constrains the maximum useful multiplication in a device before the SNR is significantly degraded. Hence, APDs for high-sensitivity applications require a material with a very low k for electron-initiated multiplication.

Semiconductor materials such as HgCdTe [3] and InAs [4] exhibit negligible hole ionization, offering nearly ideal gain with very low excess noise. However, their narrow bandgaps demand cryogenic operation to mitigate thermally generated dark currents. Silicon [5] and AlGaAsSb [6][7], with their wider bandgaps, can operate at room temperature with comparatively low dark currents and modest k values.

Efforts to overcome the limitations of materials whose electron (α) and hole (β) ionization coefficients are similar have centered on modifying the material properties to favor one carrier type over the other. Approaches include the incorporation of multiple quantum wells (MQWs) [8], "staircase" structures where band discontinuities provide carriers with additional energy [9], quantum dot avalanche regions [10], and nano-structuring [11]. Although these techniques have shown promise, they often require careful design, complex growth, or advanced fabrication processes.

More recently, studies have revealed that incorporating the large Group V element bismuth (Bi) into GaAs can significantly reduce the hole ionization coefficient while leaving the electron ionization coefficient largely unchanged [12]. This effect is attributed to the strong band anticrossing interaction induced by Bi, which increases the spin–orbit splitting energy (Δso) in GaAs [13]. Since hole ionization in GaAs depends on transitions from the heavy- and light-hole bands into the split-off band—where holes can more readily acquire sufficient energy for impact ionization [14]—increasing Δso lowers the hole population in the split-off band, thereby reducing β.

A key challenge in thick GaAsBi layers is the buildup of compressive strain as Bi content increases. Once the strain surpasses the critical layer thickness [15], misfit dislocations can form, raising dark currents and degrading device performance. Even though relaxed GaAsBi tends to have a smoother surface compared to relaxed InGaAs [16], the relaxation process remains detrimental to GaAsBi-based devices [17]. Identifying alternative ways to incorporate Bi is thus crucial for reducing k by suppressing β without compromising structural quality.

In this chapter, it is demonstrated that thick, bulk GaAsBi structures are not required to reduce β. Incorporating GaAs/GaAsBi heterostructures within a GaAs matrix can also enhance the performance of avalanching structures. A systematic study of the avalanche multiplication of a series of GaAs/GaAsBi heterostructures grown in *p-i-n* or *n-i-p* configuration is undertaken, and their ionization behaviors are investigated through photomultiplication measurements. The β/α ratio in GaAs was decreased by suppressing hole impact ionization through modifications to the valence band structure.

Because Bi is one of the largest atoms that can be introduced into GaAs, its strong electronegativity difference relative to As produces a significant perturbation in the valence band structure. This leads not only to a notable narrowing of the bandgap via a band anticrossing interaction [18] but, more importantly for our purposes, to an increase in the spin–orbit splitting energy. The consequent reduction in β underpins the strategy for achieving lower k in GaAsBi-containing APDs.

5.2 Layer details

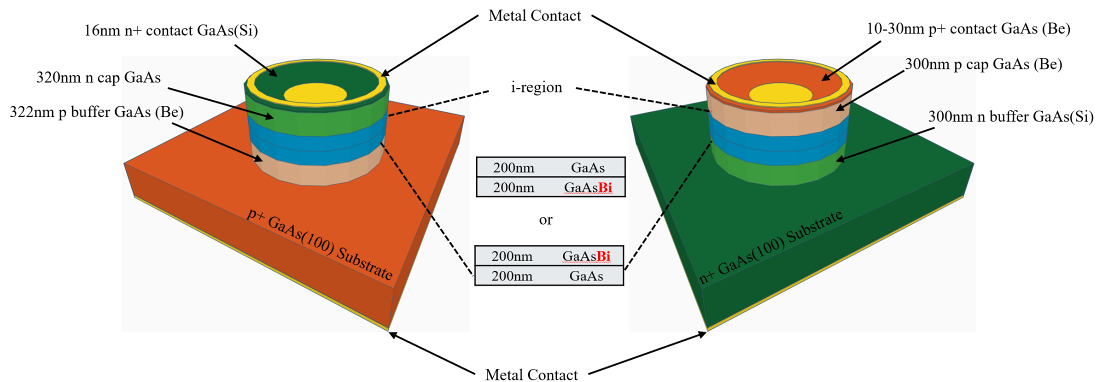


Figure 5.1 Schematic of the n-i-p and p-i-n heterostructure devices structure

Table 5.1 Details of the heterostructure devices

	Configuration	Bi layer	Bi layer %
STK2D	PIN	up	3.12%
STK2J	PIN	down	3.84%
STK62	NIP	up	1.47%
STK64	NIP	down	1.22%

In this study, a series of GaAsBi/GaAs heterojunction p–i–n and n–i–p structures were grown on GaAs substrates, as shown schematically in Figure 5.1. Each device incorporates a nominally 200 nm GaAsBi active region, with bismuth contents ranging from approximately 1.22% to 4.4%, along with 200 nm of GaAs in the i-region. The choice of a 200 nm GaAsBi layer reflects a balance between ensuring sufficient Bi incorporation for band-structure modification and avoiding strain-induced lattice defects. Specifically, Bi contents and layer thicknesses were determined using X-ray diffraction (XRD) ω–2θ scans and photoluminescence (PL) measurements, while SIMS was used as a qualitative/semi-quantitative check on layer placement and dopant

profiles. These complementary techniques confirmed the uniformity of the GaAsBi layer and verified that the targeted range of Bi composition was successfully achieved without introducing excessive dislocations.

Table 5.1 summarizes the bismuth-containing layer's location within the i-region for each heterostructure, alongside the specific Bi concentrations used. Beyond composition and thickness, the axial placement of the GaAsBi layer within the intrinsic region provides an additional design degree of freedom. Under reverse bias, the electric-field profile across a practical p–i–n / n–i–p stack is not strictly uniform once the space-charge region begins to extend into the doped claddings; therefore, moving the GaAsBi layer within the i-region changes the local field experienced by the GaAsBi/GaAs heterointerfaces. Because GaAsBi modifies the valence-band edge and introduces finite band offsets relative to GaAs, its position can influence the partitioning of the potential drop, carrier collection pathways, and the spatial overlap between the GaAsBi layer and the high-field region relevant to avalanche initiation. Consequently, otherwise similar heterostructures can exhibit measurably different leakage and multiplication characteristics when the GaAsBi layer is relocated within the intrinsic region. [2,11] By precisely controlling the thickness and composition, we aimed to manipulate the electronic and optical properties of the device in a manner conducive to enhanced photodetection and reduced noise. Finally, the topmost layer of each structure consists of a 10–30 nm p+ GaAs contacting layer, which serves to facilitate reliable electrical contacts.

5.3 SIMS of heterostructures

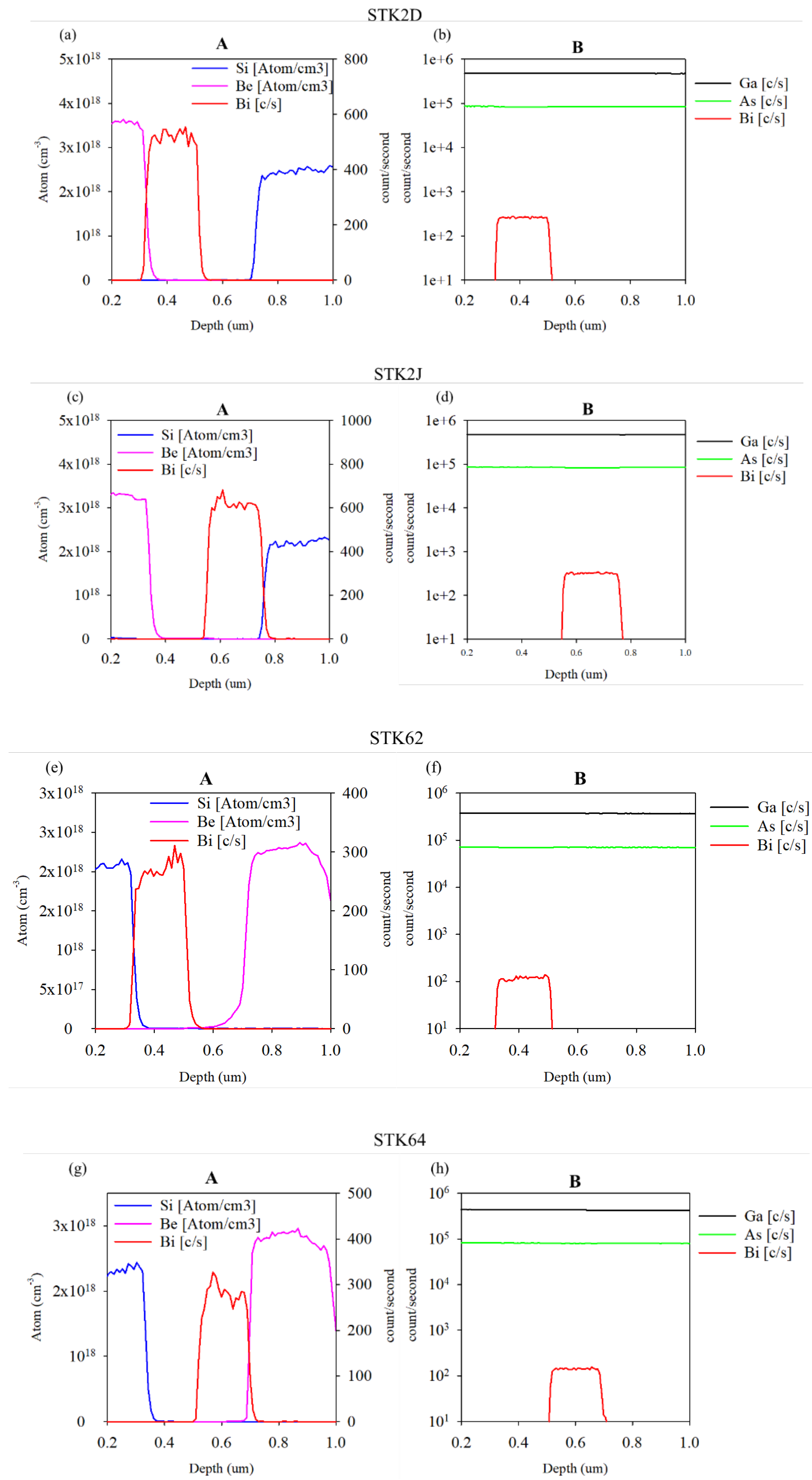


Figure 5.2 SIMS results of STK heterostructures

The SIMS depth profiles are used in this thesis primarily to verify the presence, axial position, and thickness of the GaAsBi insertion layers within the heterostructures, rather than to provide an absolute determination of the Bi mole fraction. In general, SIMS count rates depend on instrumental settings and matrix-dependent sputter and ionisation yields; therefore, quantitative conversion from SIMS intensity ratios to composition requires calibration via relative sensitivity factors (RSFs) and suitable standards. [32,33]

Accordingly, the Bi compositions reported in this chapter are taken from XRD analysis (with consistency checks against the optical response where applicable), while SIMS is interpreted qualitatively and semi-quantitatively. Importantly, across the four samples investigated here, the background-subtracted Bi SIMS signal within the GaAsBi layer increases monotonically and is approximately proportional to the Bi% trend reported for the sample set. This proportionality indicates that, under identical measurement conditions, SIMS provides a reliable measure of relative Bi incorporation and layer placement across the compared heterostructures, while XRD remains the basis for absolute Bi% values.

Four GaAs-based heterojunction photodiodes, designated STK2D, STK2J, STK62, and STK64, were examined using secondary ion mass spectrometry (SIMS) to verify the intended p–i–n or n–i–p design and the incorporation of bismuth in the i-region. In each of the corresponding SIMS depth profiles, panel A illustrates the atomic concentrations of Si and Be (in atoms/cm^3) together with the bismuth signal (in counts/second), while panel B shows the count-rate variations for Ga, As, and Bi. The depletion region (i-region) and the transition from p+ (or n+) into the intrinsic layer are identified by the abrupt changes in dopant levels.

STK2D and STK2J (p–i–n structures):

For STK2D and STK2J, Be (pink trace) is present at high concentrations near the surface, indicating the p+ contact and p cap layers. At greater depths, Si (blue trace) begins to dominate, marking the transition to the n+ buffer and substrate. Within the intervening region of roughly 0.2–0.6 µm or 0.3–0.7 µm, the doping level drops sharply—consistent with a 400 nm i-region.

The bismuth signal (red) peaks within one half of this i-region, in agreement with the

design that specifies a 200 nm GaAsBi layer (with Bi content of approximately 3.12% for STK2D or 3.84% for STK2J). The measured depth and amplitude of the Bi peak line up well with the expected location in the intrinsic region, confirming that the p–i–n structure is properly realized.

STK62 and STK64 (n–i–p structures):
In STK62 and STK64, the SIMS data show a high concentration of Si near the top surface (blue trace), corresponding to the n+ contact and n cap layers. At deeper depths, Be (pink trace) becomes dominant, indicating the p buffer and p+ substrate region. As with the p–i–n samples, the middle ~400 nm of low dopant concentration constitutes the i-region.
In STK62, the top 200 nm of the i-region is GaAsBi with ~1.47% Bi, while in STK64, the bottom 200 nm contains ~1.22% Bi. SIMS confirms this arrangement: the bismuth signal peaks in either the upper or lower portion of the intrinsic layer, matching the design placement of the GaAsBi segment.

Concordance with Design Specifications:

Four GaAs-based heterojunction photodiodes (STK2D, STK2J, STK62, and STK64) were examined by SIMS to verify the intended p-i-n or n-i-p doping sequence and the axial placement of the nominal 200 nm GaAsBi insertion within the ~400 nm i-region. In the SIMS depth profiles, the abrupt Be/Si transitions identify the doped claddings and the low-doped i-region, while the Bi signal (counts/s) indicates the position of the GaAsBi insertion.

For the p-i-n samples (STK2D and STK2J), Be is present at high concentration near the surface and Si dominates at greater depths, consistent with the intended layer ordering. The Bi signal is confined to the expected portion of the i-region, indicating that the GaAsBi insertion is positioned as designed. For the n-i-p samples (STK62 and STK64), Si dominates near the surface and Be becomes dominant at depth, again consistent with the intended polarity, and the Bi signal peaks in the upper or lower half of the i-region in accordance with the intended GaAsBi placement.

Quantitative conversion of SIMS count rates to an absolute Bi mole fraction is not attempted here, because SIMS intensity ratios depend on instrumental settings and

matrix-dependent sputter/ionisation yields and require calibration via relative sensitivity factors (RSFs) and suitable standards. Therefore, the absolute Bi compositions quoted for the heterostructures are taken from XRD analysis (with optical consistency checks where applicable), while SIMS is used to confirm layer placement and to provide a consistent measure of relative Bi incorporation across the sample set under identical SIMS measurement conditions.

5.4 XRD of heterostructures

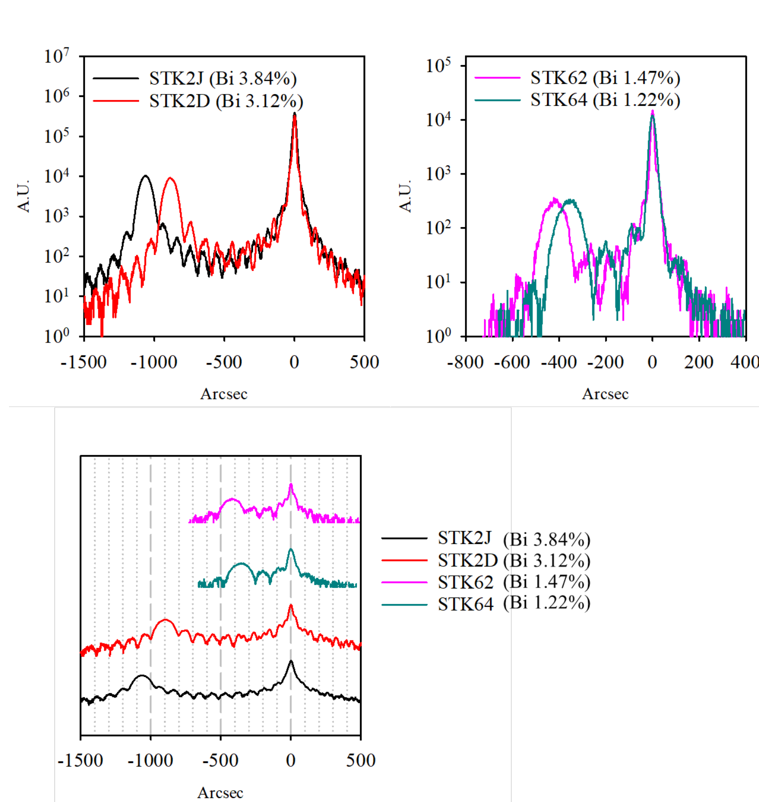


Figure 5.3 X-ray diffraction results of STK heterostructures

XRD results Analysis

X-ray diffraction (XRD) is critical for evaluating the lattice mismatch, strain relaxation, and interface quality of GaAsBi layers [19] [20] [21] [25]. Once the epitaxial layer is several hundred nanometers thick and the Bi fraction is high, it may exceed the critical thickness, leading to strain relaxation or dislocations [21] [24].

The ω–2θ scans around the GaAs (004) reflection show a systematic shift of the GaAsBi peak towards lower 2θ as the nominal Bi content increases, consistent with lattice expansion in dilute GaAsBi. The larger GaAs–GaAsBi peak separation observed for STK2D compared with STK62/STK64 is therefore consistent with the higher Bi fraction targeted in STK2D.

In addition to the peak shift, the GaAsBi feature in STK2D is broader and of reduced intensity. Such peak broadening can arise from a combination of finite layer thickness, composition/strain inhomogeneity, and increased mosaicity or defect-related disorder. Importantly, a single symmetric ω–2θ scan does not uniquely distinguish coherent strain from partial relaxation, and quantitative determination of relaxation generally requires reciprocal-space mapping (RSM) and/or asymmetric reflections. [24,34] Therefore, in the present work the broadened GaAsBi peak in STK2D is interpreted as a qualitative indicator of increased structural inhomogeneity/disorder (and a potentially higher propensity for relaxation at higher Bi fraction), rather than as a direct measure of a specific relaxation fraction.

Differences in substrate doping for p-i-n vs. n-i-p also alter the surface potential at the onset of growth [26] [27]. Locating GaAsBi in the "upper 200 nm" or "lower 200 nm" can influence the local growth temperature profile and Bi adsorption. Accurate fitting of local shoulders or fringe damping often requires a multi-layer or graded-strain approach rather than a simple uniform-layer model [21] [24] [25].

Device Characteristics and Further Validation

As photodiodes, these GaAsBi layers directly determine the bandgap and absorption edge [22] [23]. Each additional 1% Bi typically reduces the bandgap by about 60–90 meV. STK2J (3.84% Bi) may thus enable longer-wavelength infrared detection but could experience higher defect densities and dark current. STK62 (1.47% Bi) and STK64 (1.22% Bi) would cover shorter wavelengths yet might achieve more stable device performance.

Further confirmation of Bi content or bandgap characteristics can come from photoluminescence (PL), photocurrent, or electromodulation spectroscopy [22] [23]. In I–V and spectral response measurements, heavily relaxed high-Bi samples can exhibit

increased nonradiative recombination and reduced efficiency [19] [21] [24]. For lower-Bi materials, bandgap shifts are smaller but crystal quality is more easily maintained, leading to lower noise. By comparing these four samples, one can explore GaAsBi growth and performance under different thicknesses, doping orders, and layer placements, thus gaining a more comprehensive insight into the narrow growth window, strain management, and ultimate photoelectronic behavior [19] [20] [21] [26] [27].

All four heterojunction photodiodes incorporate a 400 nm i-region with a GaAsBi sub-layer, whose Bi fraction spans 1.22% to 3.84%. In p-i-n or n-i-p layouts, higher Bi (>3%) more readily surpasses the critical thickness, causing broader diffraction peaks and more extensive strain relaxation, potentially harming device operation. Lower Bi (1–2%) yields less mismatch but also less bandgap tuning.

By combining optical measurements (e.g., PL, photocurrent) with multi-layer XRD fitting, one can more precisely clarify the actual Bi distribution and strain state, which is valuable for device design and epitaxial optimization [22] [23] [25] [26] [27].

5.5 Electrical characterisation of heterostructures

5.5.1 IV measurements

Four GaAs-based heterostructure photodiodes (STK2D, STK2J, STK62, STK64) were investigated to assess their dark current characteristics. Each sample incorporates a 400 nm intrinsic region (i-region), partially composed of GaAsBi with different Bi contents and placed either in the top or bottom 200 nm of the i-region. Specifically, STK2D and STK2J adopt a p–i–n configuration (Bi ≈ 3.12% or 3.84%), while STK62 and STK64 employ an n–i–p structure (Bi ≈ 1.47% or 1.22%). Figures 5.4–5.7 compare their reverse and forward dark currents for devices of various diameters (e.g., 70 μm, 120 μm, 220 μm, 420 μm).

Although device-to-device variation can include contributions from sidewall/surface

leakage associated with mesa etching and passivation, it is still possible to discuss the expected diameter dependence using a simple geometry argument. For circular mesas, the perimeter-to-area ratio is $P/A = 4/D$. If the reverse leakage contains a bulk (area-related) component and a sidewall (perimeter-related) component, then $I_{\text{dark}}(V)$ can be written phenomenologically as

$$I_{\text{dark}}(V) = J_{\text{bulk}}(V)A + J_{\text{surf}}(V)P,$$

which implies a current density

$$J_{\text{dark}}(V) = \frac{I_{\text{dark}}(V)}{A} = J_{\text{bulk}}(V) + \frac{4}{D}J_{\text{surf}}(V).$$

This relation shows that any sidewall contribution increases systematically in J_{dark} as the diameter decreases. Therefore, the comparison of reverse dark current density across multiple diameters provides a practical diagnostic for assessing whether the measured leakage is bulk-dominated or perimeter-dominated: a near-collapse of $J_{\text{dark}}(V)$ across diameters indicates that sidewall leakage is well suppressed over the measured bias range, whereas a consistent elevation of $J_{\text{dark}}(V)$ for smaller mesas is consistent with a non-negligible surface leakage contribution. In Figs.~5.4-5.7, traces with the same colour correspond to the same nominal diameter, while solid and dashed lines represent measurements from different nominally identical devices of that same diameter, thereby visualising device-to-device spread at fixed geometry.

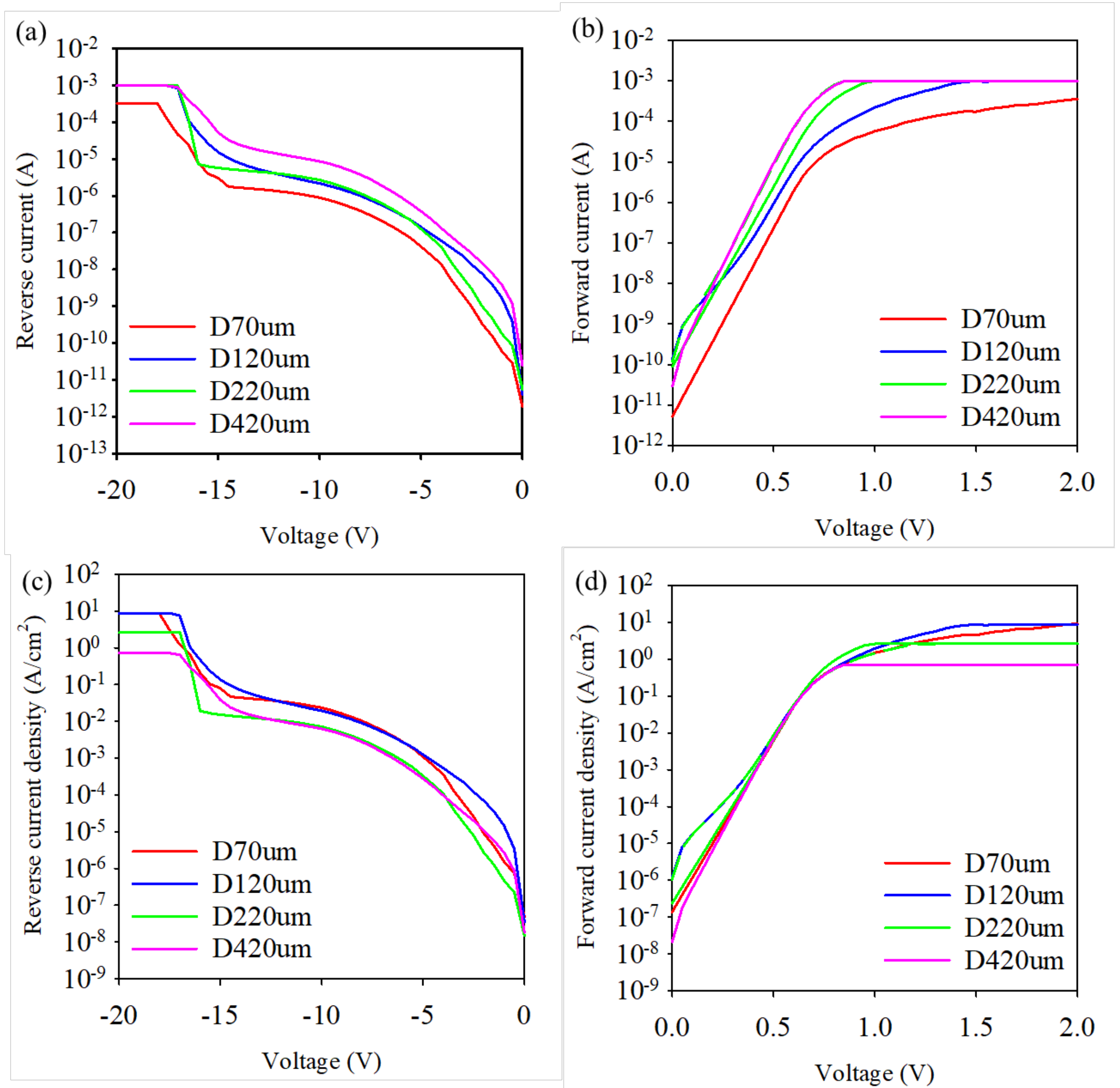


Figure 5.4 (a) Reverse dark current, (b) Forward dark current, (c) Reverse dark current density, (d) Forward dark current density of STK2D

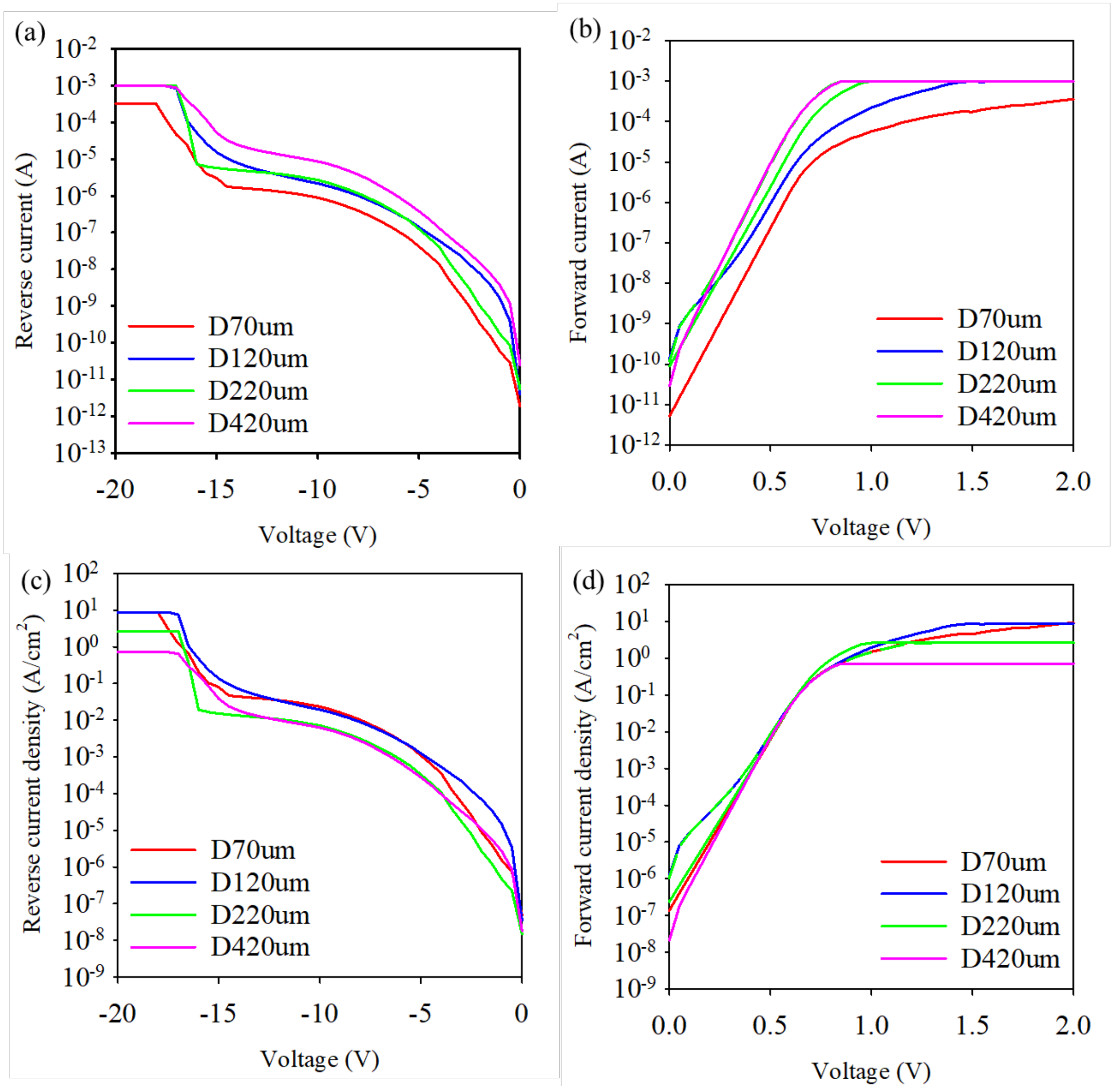


Figure 5.5 (a) Reverse dark current, (b) Forward dark current, (c) Reverse dark current density, (d) Forward dark current density of STK2J

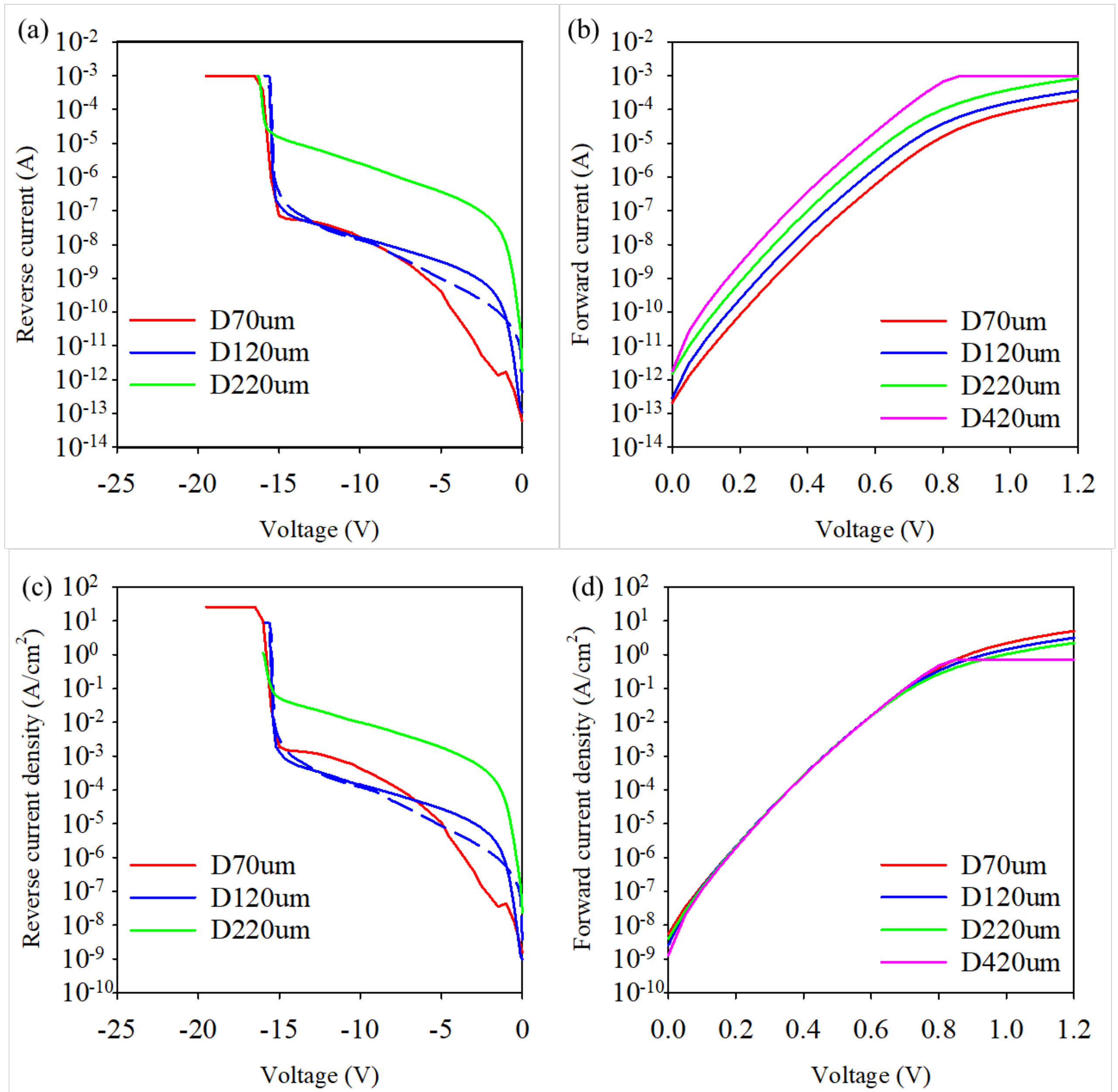


Figure 5.6 (a) Reverse dark current, (b) Forward dark current, (c) Reverse dark current density, (d) Forward dark current density of STK62

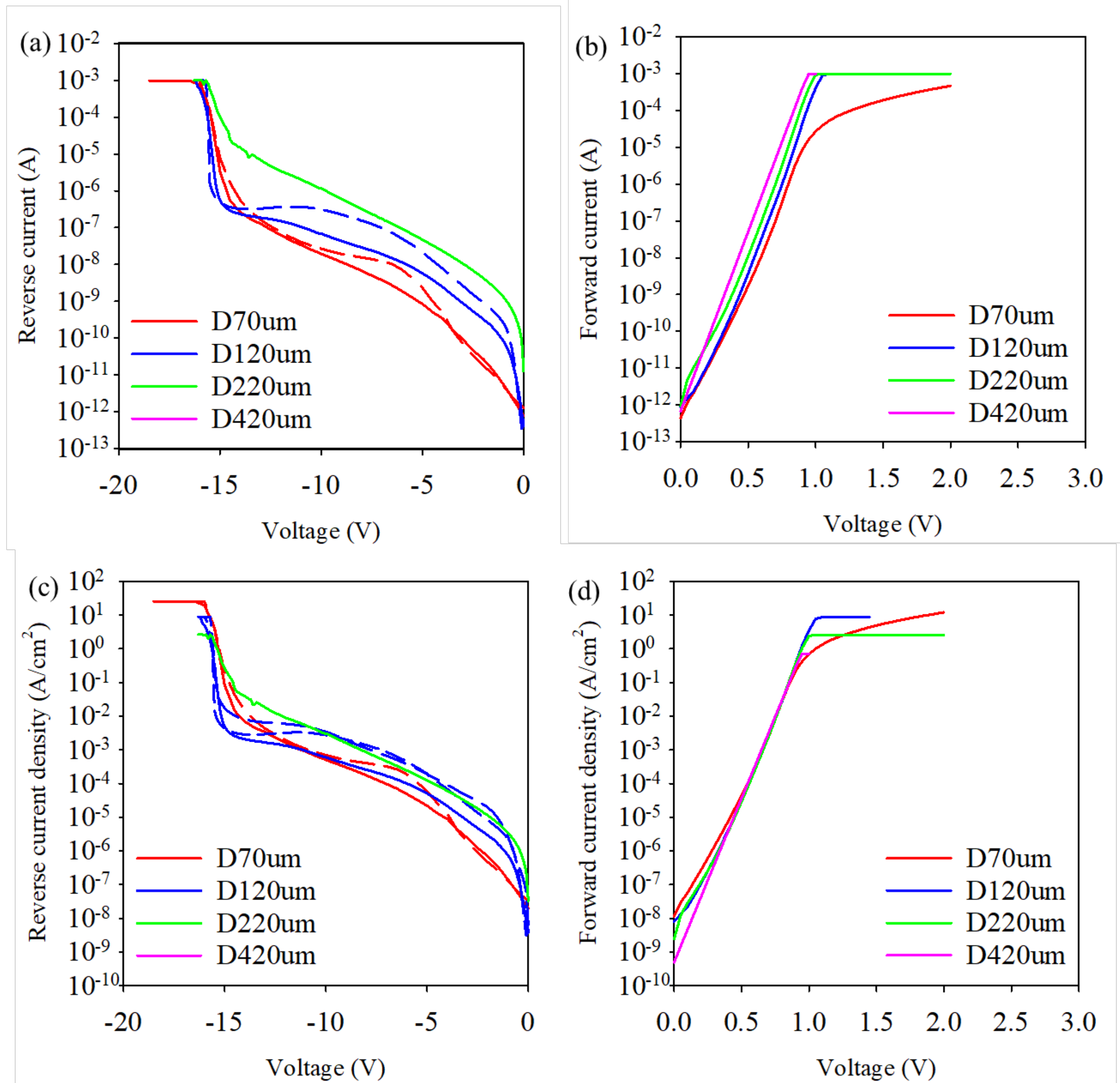


Figure 5.7 (a) Reverse dark current, (b) Forward dark current, (c) Reverse dark current density, (d) Forward dark current density of STK64

Figure 5.4(a) shows the reverse dark current for STK2D at different device diameters. The current increases slightly with increasing device area, but the slopes do not always match perfectly in log scale. In general, the breakdown voltage appears near –20 V for the smaller devices, though the larger diameter device sometimes shows a slightly different breakdown profile due to increased bulk current or series resistance.

In Figure 5.4(b), the forward dark current rises from ~10^{-13} A at 0 V to above 10^{-3} A near 2 V in the largest device. The smaller device (70 μm) exhibits a comparable trend but slightly higher current at moderate voltages, potentially reflecting variations in contact resistance or sidewall leakage. When normalized by area, Figure 5.4(c)-(d) show that, after normalisation by mesa area, the reverse current density traces for different diameters are broadly comparable over the measured bias range. This near-collapse is consistent with the leakage being predominantly area-related (bulk-dominated) in these devices, with any perimeter-related contribution being secondary within the bias range studied. The remaining separation between nominally identical devices of the same diameter is therefore interpreted as device-to-device spread at fixed geometry, which can arise from a combination of fabrication- and measurement-related variations (e.g., small differences in effective mesa area, contact quality, and local sidewall/passivation conditions), rather than from a systematic diameter dependence alone.

Figure 5.5 displays the I–V curves for STK2J, which also has a p–i–n layout but with a higher Bi fraction (~3.84%) located in the bottom half of the i-region. As in STK2D, the reverse dark current in Figure 5.5(a) increases with device diameter, though not strictly proportional over the entire voltage range. Meanwhile, the forward I–V in Figure 5.5(b) typically shows higher current than STK2D at similar biases, possibly due to the higher Bi content facilitating increased carrier generation or a slightly lower effective bandgap in the active region.

The corresponding current densities in Figure 5.5(c)–(d) largely overlap for the 120 μm and 220 μm devices, indicating that the epitaxial uniformity is reasonably maintained for those die sizes. However, the 70 μm and 420 μm devices deviate somewhat, suggesting that edge effects, contact quality, or doping gradients might be influencing the overall dark current.

Figure 5.6(a) illustrates the reverse dark current for STK62, which follows an n–i–p structure with ~1.47 % Bi in the top half of the i-region. The breakdown voltage (–15 V to –20 V range) is slightly lower than that of STK2D or STK2J, consistent with different doping profiles or avalanche thresholds for the n–i–p device. At reverse biases just below breakdown, current scaling with area is fairly consistent among the 70 µm, 120 µm, 220 µm, and 420 µm devices, though minor discrepancies suggest local variations in the p buffer or n cap.

Under forward bias (Figure 5.6(b)), the dark current spans 10^{-13} A to 10^{-3} A from 0 V up to ~1.2 V, reflecting the diode conduction regime. The area-normalized data in Figure 5.6(c)–(d) imply that the i-region doping and Bi incorporation do not significantly degrade the uniformity across devices of different sizes, although the largest diode diameter tends to show a slightly higher forward current density above 0.8 V, possibly due to lateral conduction paths or contact resistance.

Figure 5.7 (not shown in detail here, for brevity) corresponds to STK64, another n–i–p configuration but with Bi (~1.22%) placed in the lower half of the i-region. Overall behavior resembles STK62, with the reverse leakage current and breakdown appearing at similar magnitudes. Forward conduction in STK64 is comparable to that of STK62, though subtle differences around 1.0–1.5 V might indicate slight variations in doping profiles or in the GaAsBi layer thickness.

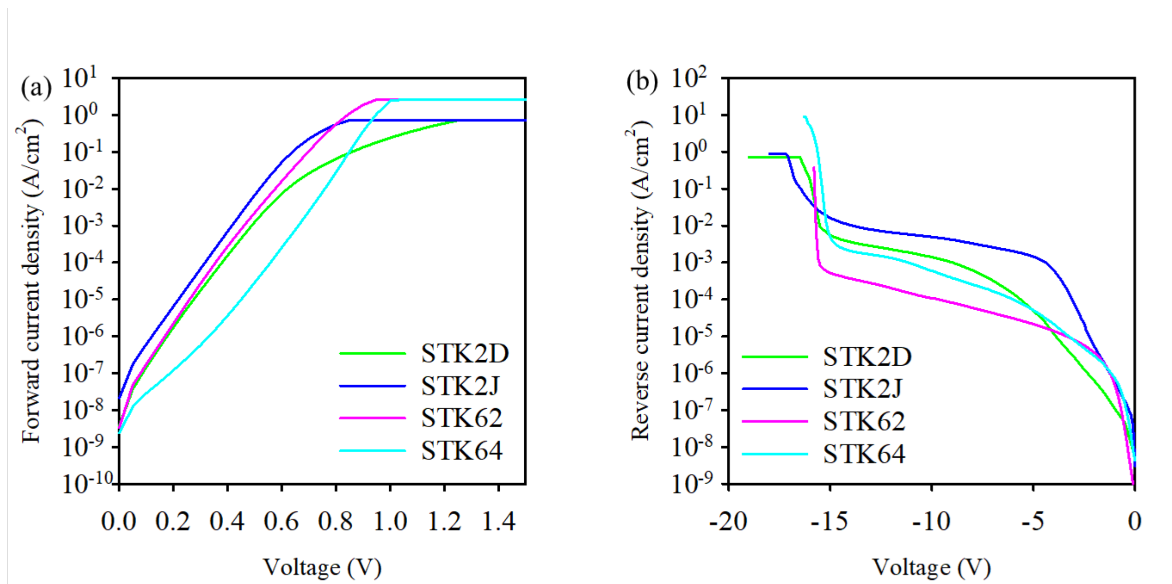


Figure 5.8 a) Reverse dark current densities and (b) Forward dark current densities of heterostructure devices

To further compare these four samples (STK2D, STK2J, STK62, and STK64), Figure 5.8 plots their reverse and forward dark current densities. Over the entire bias range, both p–i–n devices (STK2D, STK2J) display somewhat different slopes than the n–i–p samples (STK62, STK64), likely reflecting the band structure modifications by placing GaAsBi either on the top or bottom of the i-region. Table 5.2 summarises the extracted saturation current density J0 and ideality factor n for the heterostructure devices. The values vary substantially across the sample set, indicating that the dark current is sensitive to multiple coupled factors, including junction polarity, axial placement of the GaAsBi insertion, and process-dependent leakage pathways (e.g., mesa sidewall condition and passivation), in addition to Bi fraction. Consequently, within the present dataset a monotonic dependence of dark current density on Bi composition cannot be established, and the results highlight the importance of co-optimising epitaxial quality and device processing when pushing towards higher Bi incorporation.

Table 5.2 Dark current density and ideality factor of heterostructure devices

Samples	J0(A/cm^2)	Ideality Factor n
STK2D	1.9×10^{-8}	1.79
STK2J	7.12×10^{-8}	1.75
STK62	1.23×10^{-7}	2.04
STK64	9.5×10^{-11}	1.63

In Table 5.2, the extracted saturation current densities (J_0) span roughly 10^{-8} to 10^{-11} A cm^{-2}, with ideality factors (n) of 1.6–2.0. These values align with standard diode behavior, although they vary slightly across different Bi fractions and doping layouts.

From the above I–V measurements, all four heterostructure photodiodes exhibit diode-like conduction in forward bias and a reasonably defined breakdown region in reverse bias. Variations in current scaling with area indicate that sample-to-sample differences in doping uniformity, contact resistance, and sidewall passivation can influence the dark current—particularly near breakdown or above 1 V forward bias. Nonetheless, the overall performance is in line with the intended design of placing 200 nm GaAsBi layers at various positions within a 400 nm i-region, confirming that the epitaxial growth and doping profiles are suitably configured for subsequent device evaluation.

5.5.2 CV measurements

To further evaluate the depletion-region characteristics and doping profiles of the four GaAs-based heterostructure photodiodes (STK2D, STK2J, STK62, STK64), we performed capacitance–voltage (C–V) measurements on devices with various diameters (e.g., 70 µm, 120 µm, 220 µm, 420 µm). Each diode has a 400 nm i-region containing a 200 nm GaAsBi layer at either the upper or lower half, with different Bi contents. Figures 5.9–5.12 show the measured capacitance and the capacitance per unit area over a range of applied voltages, providing insight into the depletion behavior under reverse (or small forward) biases.

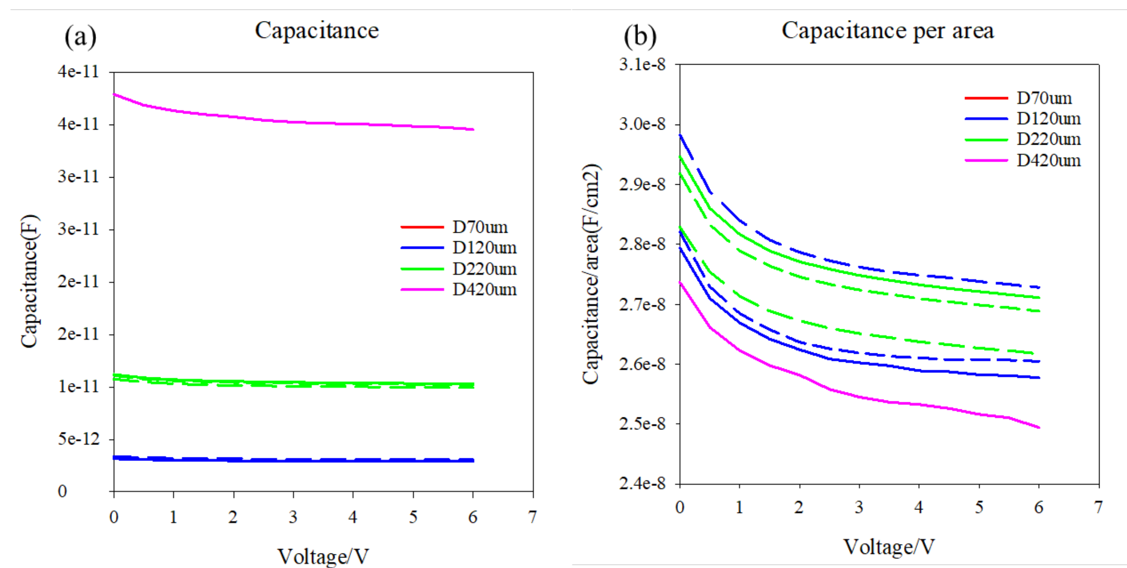


Figure 5.9 (a) CV of STK2D; (b) Capacitance per area of STK2D

In Figs. 5.9–5.12, colour is used to indicate device geometry: traces with the same colour correspond to devices with the same mesa size (and hence nominal area). For each size, the solid and dashed curves represent measurements from different nominally identical devices, thereby visualising device-to-device variation at fixed geometry.

As illustrated in Figure 5.9(a), STK2D exhibits a gradually decreasing capacitance from 0 V to about 7 V, with the 420 µm device showing a significantly larger absolute capacitance than the 70 µm device. After normalizing by area (Figure 5.9(b)), the

curves for different diameters become more comparable but still exhibit a degree of variation.

Such discrepancies typically indicate that the nominal and actual device areas differ slightly, or that the cladding/substrate doping is not perfectly uniform. Overall, STK2D's C–V curve decreases monotonically with increasing voltage, consistent with the gradual expansion of the depletion region in a p–i–n configuration.

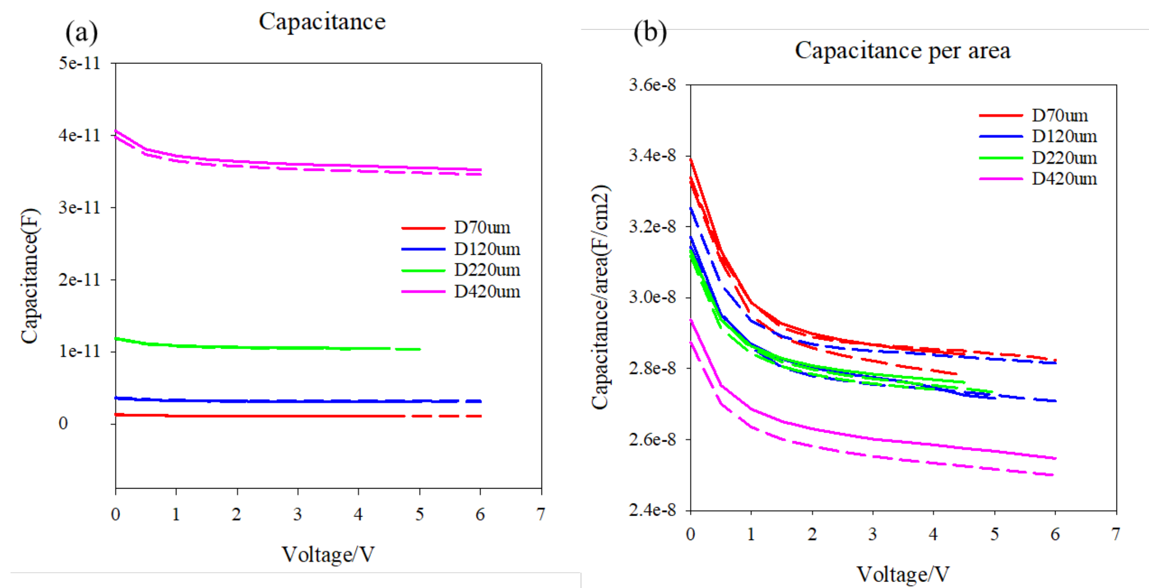


Figure 5.10 (a) CV of STK2J; (b) Capacitance per area of STK2J

Figure 5.10(a) shows that STK2J's capacitance likewise declines as the voltage rises, but the slope near 0 V is slightly less steep than in STK2D. This device incorporates a higher Bi fraction (~3.84%) at the bottom 200 nm of the i-region, potentially affecting its band structure. The area-normalized data in Figure 5.10(b) reveal more pronounced differences between device diameters at lower voltages, suggesting that edge or contact effects play a larger role when the depletion region is shallower.

However, by around 5 V, the diode is nearly fully depleted, and the unit-area capacitances for most diameters converge, indicating that the depletion zone is

extending between the p+ and n buffer layers as intended.

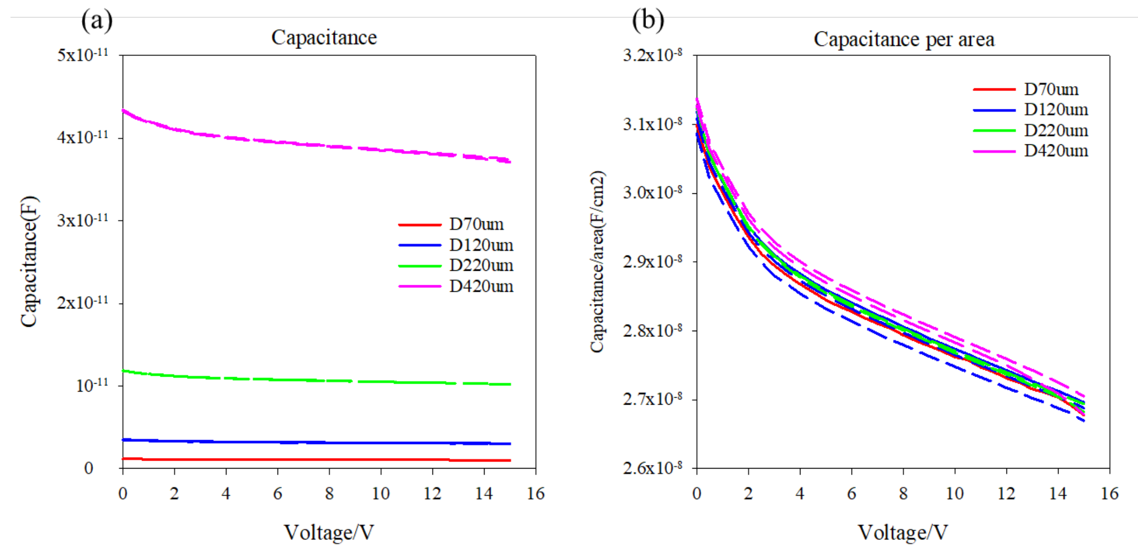


Figure 5.11 (a) CV of STK62; (b) Capacitance per area of STK62

STK62 adopts an n–i–p configuration, with ~1.47 % Bi in the top 200 nm of the i-region. As shown in Figure 5.11(a), its capacitance decreases steadily with voltage from 0 V up to ~15 V. The absolute values scale with device diameter, with the 70 µm diode having the lowest total capacitance and the 420 µm diode the highest. In the normalized curves (Figure 5.11(b)), the traces generally overlap in the mid-voltage range (2–10 V), though some spread at the low- and high-voltage ends indicates variations in doping uniformity or lateral extension of the depletion region.

Compared to the p–i–n diodes, STK62 can be biased up to ~15 V before exhibiting breakdown, possibly reflecting a higher allowable depletion width in its n–i–p configuration.

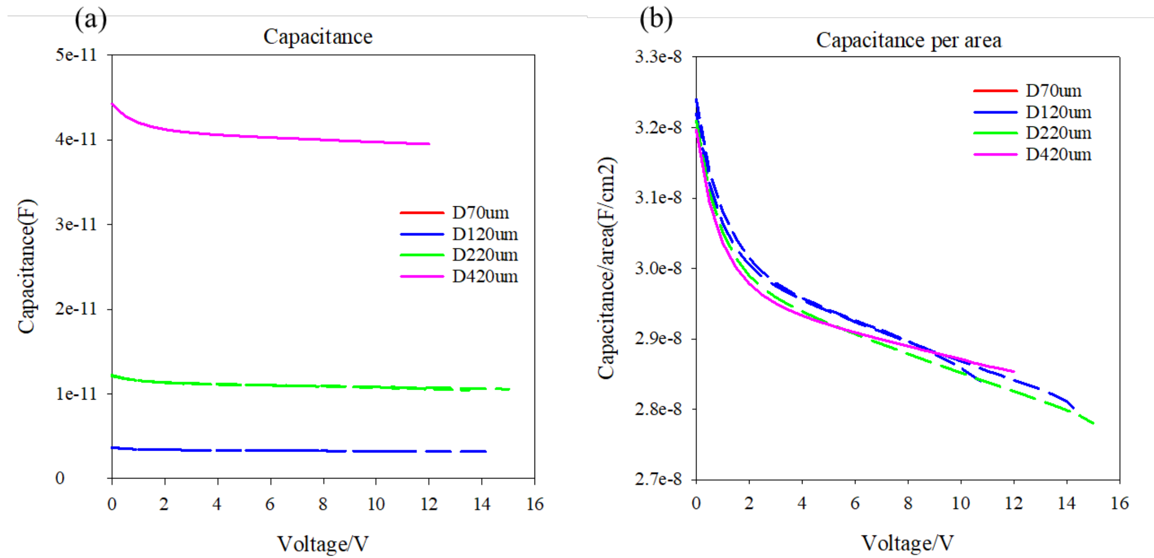


Figure 5.12 (a) CV of STK64; (b) Capacitance per area of STK64

STK64 is also n–i–p, but places ~1.22% Bi in the bottom 200 nm of the i-region. Figure 5.12(a) shows that its absolute capacitance is slightly lower than STK62 at comparable voltages, suggesting that under the same bias, the depletion region may extend further or that the band offsets differ.

In Figure 5.12(b), the unit-area capacitance converges among most device diameters in the intermediate voltage range (5–10 V), becoming more dispersed only at very low or very high voltages—again pointing to minor geometry or doping variations. Overall, STK64 exhibits a C–V profile similar to STK62, though the specific Bi content and placement lead to distinct depletion characteristics near the extremes of the bias range.

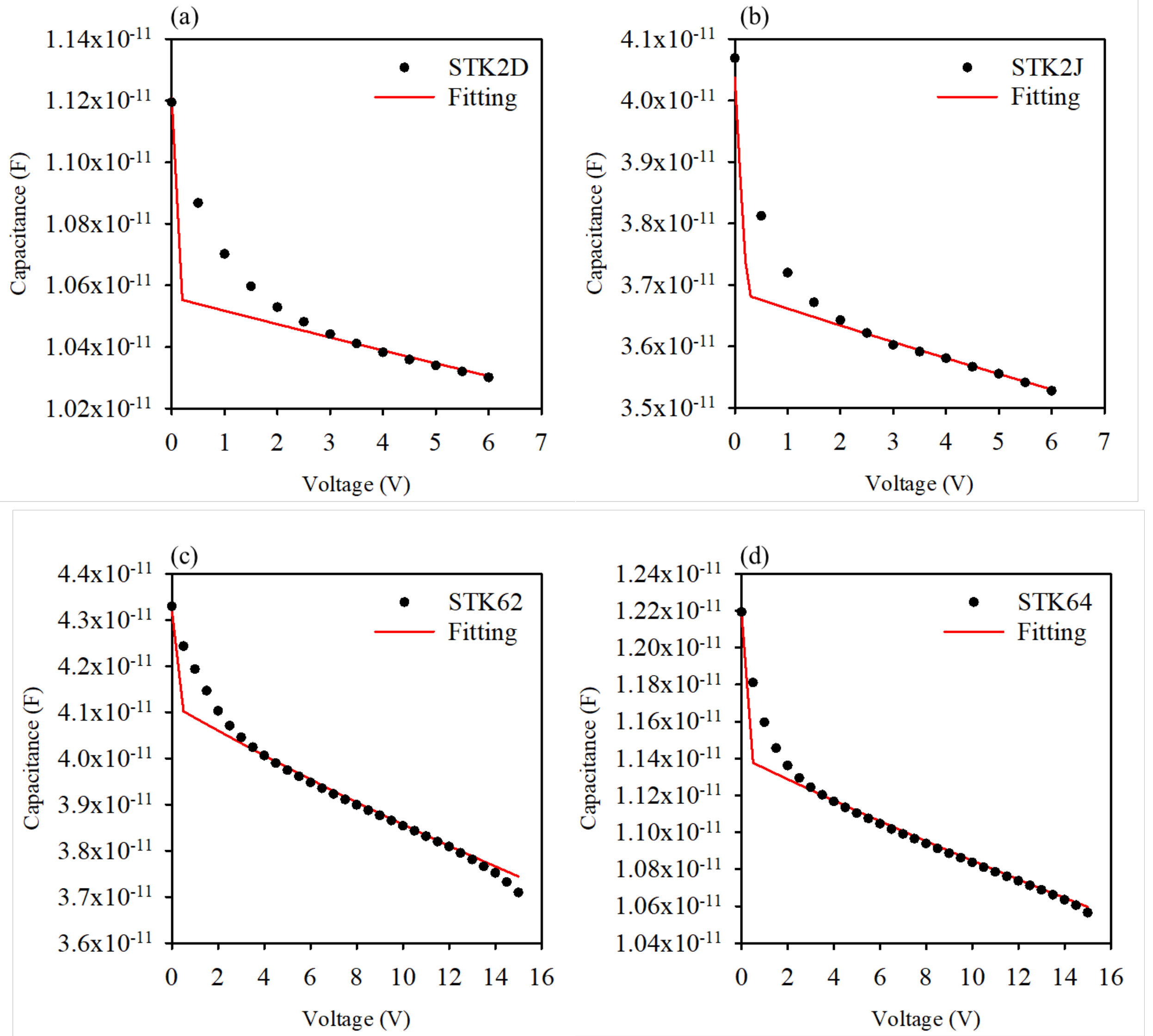


Figure 5.13 (a)-(d) Measured CV of heterostructure devices, black dots are raw data and red lines are fitting from Poisson's equation

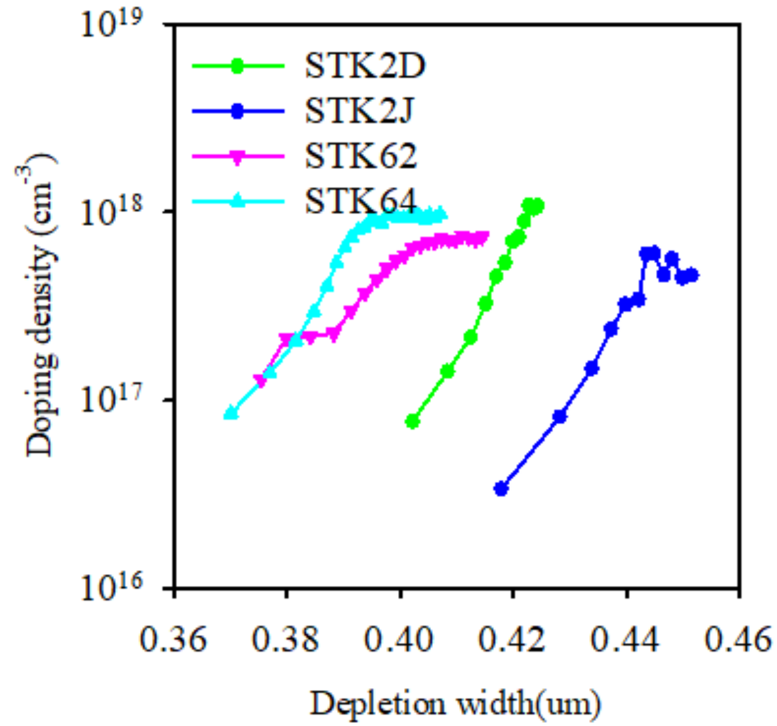


Figure 5.14 Doping profile of heterostructure devices

Table 5.3 heterostructure CV model parameters compared with nominal thickness

Samples	Nominal i-region thickness(nm)	i-region thickness(nm)	Error (%)
STK2D	400	437	+9.25
STK2J	400	422	+5.5
STK62	400	397	-0.75
STK64	400	389	-2.75

In Figure 5.13(a)–(d), the measured C–V data (black dots) are compared with Poisson-based fitting curves (red lines). This comparison enables estimation of the actual i-region thickness, as well as the doping distribution via approaches like the $1/C^2$ method. Figure 5.14 plots the doping concentration (~10^{17-10}19 cm^{-3}) versus depletion width for the four diodes, illustrating how the doping gradients differ among p–i–n

(STK2D, STK2J) and n–i–p (STK62, STK64) structures.

Table 5.3 lists the i-region thickness extracted from the C–V fits against the nominal 400 nm value. The discrepancies generally remain within ±10%, consistent with acceptable process tolerances.

From these C–V measurements, we see that the capacitance of STK2D, STK2J, STK62, and STK64 does not always scale perfectly with device area—reflecting potential differences in actual geometry or doping uniformity. Nevertheless, once fully depleted at moderate or higher biases, the unit-area capacitance for each diameter tends to converge, indicating that the i-region thickness primarily governs the device's depletion behavior.

Furthermore, comparison with the SIMS-derived thickness confirms that the i-region length obtained from C–V fitting is in good agreement with the SIMS results. This corroboration underscores the reliability of the molecular beam epitaxy (MBE) growth process, demonstrating that the GaAsBi layers—despite varying Bi contents and positions—were successfully incorporated as designed. Consequently, the p–i–n / n–i–p heterostructures maintain well-defined depletion regions, providing a solid basis for subsequent investigations into their optical response and potential avalanche multiplication behaviour.

5.6 Photo spectrum of heterostructures

To investigate the optical absorption and response of four heterostructure photodiodes (STK2D, STK2J, STK62, and STK64) under varying bias conditions, photocurrent-versus-wavelength measurements were performed at applied voltages ranging from 0 V to 4 V. Each device features a 400 nm i-region with a 200 nm GaAsBi layer (differing in Bi content and placement). Figures 5.15(a)–(d) show the measured photocurrent spectra for STK2D, STK2J, STK62, and STK64, respectively, while Figure 5.15(e) compares their absorption coefficients as a function of photon energy at 0 V.

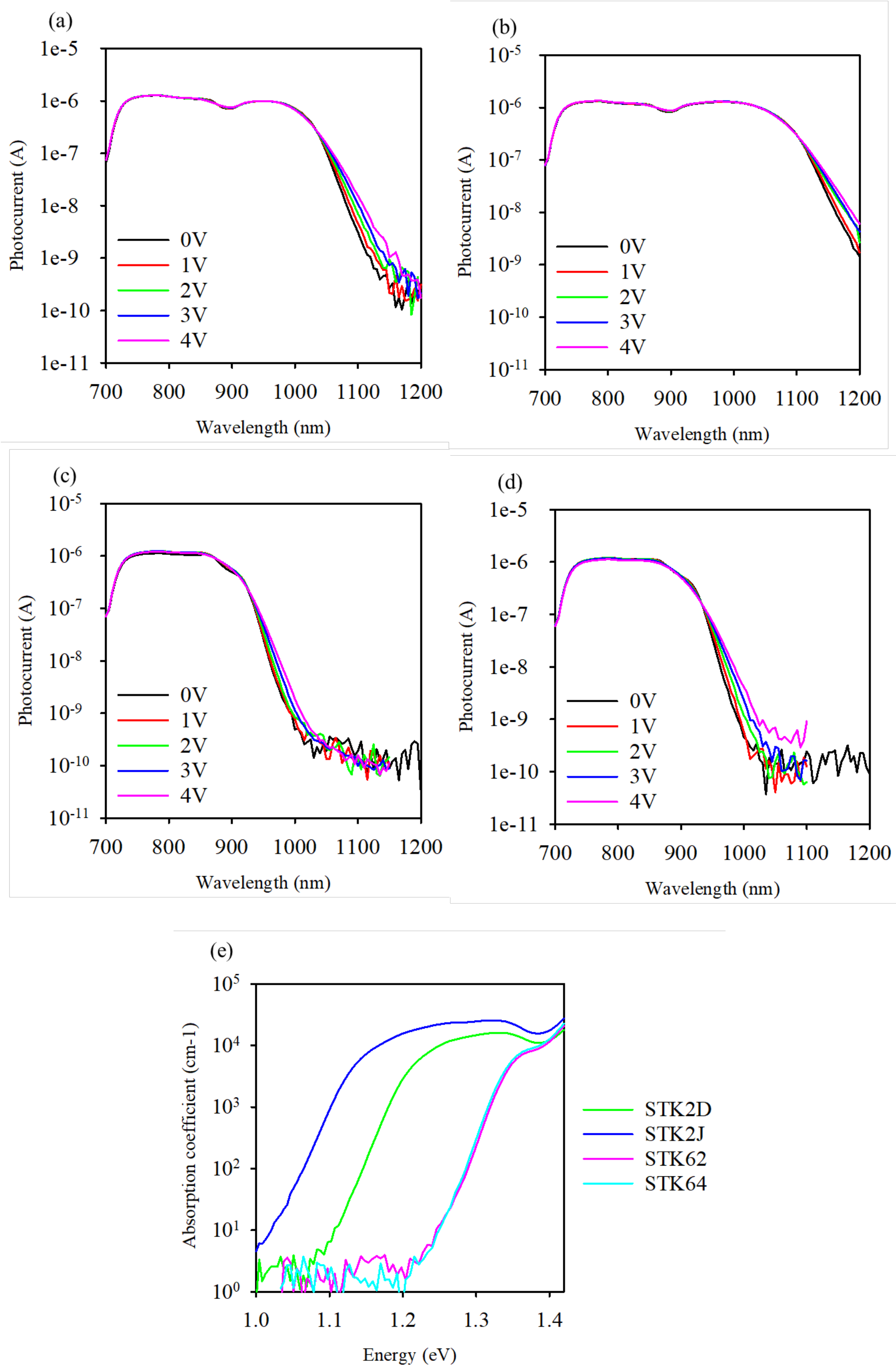


Figure 5.15 (a)-(d) Bias dependent photocurrent of heterostructure devices; (e)Absorption coefficient of heterostructure devices (@0V)

In Figure 5.15(a), STK2D exhibits photocurrent from 700 nm to 1200 nm. As the bias increases from 0 V to 4 V, notable features appear near ~830 nm (GaAs band edge) and ~1000 nm (GaAsBi band edge). The enhancement or shift around 900–1100 nm suggests that the depletion region and Franz–Keldysh effects modify carrier collection under an increasing forward bias. Since STK2D is a p–i–n device with ~3.12% Bi in the top 200 nm of the i-region, the extended photocurrent beyond 1000 nm indicates a red-shifted effective bandgap.

Figure 5.15(b) presents STK2J, which is also p–i–n but with 3.84% Bi in the lower half of the i-region. Compared to STK2D, the higher Bi content yields a more pronounced shift at ~1000–1100 nm under biases up to 4 V, reflecting stronger absorption in this near-infrared range. The photocurrent increases at longer wavelengths as the bias grows, pointing to enhanced carrier extraction aided by the electric field. Overall, STK2J's larger Bi fraction and deeper GaAsBi placement lead to a distinctly shaped spectrum compared to STK2D.

Figure 5.15(c) shows the spectrum for STK62, an n–i–p device with ~1.47% Bi in the top 200 nm. Although its absorption edge still shifts past 900 nm, the long-wavelength photocurrent beyond 1000 nm is somewhat weaker compared to the p–i–n devices. As the bias increases from 0 V to 4 V, subtle changes occur around 830 nm (GaAs band gap) and 1000 nm, highlighting the Franz–Keldysh effect and possible changes in depletion width. The overall photocurrent spans ~10^{-5} A to 10^{-11} A, reflecting five orders of magnitude in response.

Similarly, Figure 5.15(d) depicts STK64, another n–i–p structure but with ~1.22% Bi in the lower 200 nm of the i-region. The spectrum resembles that of STK62, with a notable absorption tail extending to ~1000 nm, though the overall photocurrent at longer wavelengths is lower, consistent with the smaller Bi fraction. Raising the bias from 0 V to 4 V again modifies the absorption edge via the Franz–Keldysh mechanism, though the changes are generally less dramatic than in the higher-Bi-content diodes.

Figure 5.15(e) compares the absorption coefficients of these four devices at 0 V as a function of photon energy. Their band-edge transitions fall between ~1.1 eV and ~1.3 eV, corresponding to wavelengths near 950–1100 nm. STK2D and STK2J, with higher Bi contents, show more prominent red shifts (lower bandgap), while STK62 and STK64, with lower Bi contents, exhibit higher-energy absorption edges. These distinctions align well with the photocurrent tails observed in Figures 5.15(a)–(d).

All four devices reveal spectral features around ~830 nm (GaAs) and ~1000 nm (GaAsBi). The changes observed under different biases underscore the Franz–Keldysh effect and variations in carrier collection.The higher-Bi p–i–n devices (STK2D, STK2J) achieve stronger near-infrared response, whereas the lower-Bi n–i–p samples (STK62, STK64) display less pronounced red-shifting, yet still extend the absorption beyond that of standard GaAs. The 0 V absorption coefficient curves confirm that STK2D and STK2J have narrower effective bandgaps, while STK62 and STK64 remain at slightly higher energies, matching their lower Bi percentages.

Overall, these results demonstrate that placing a GaAsBi layer of various Bi fractions and positions within the i-region enables tunable photoresponse from 700 nm to 1200 nm. Under biases up to 4 V, each diode exhibits a rational shift in band-edge absorption and photocarrier collection, thus paving the way for future investigations into avalanche multiplication and impact ionization within these heterostructures.

5.7 Photomultiplication characterisation of heterostructures

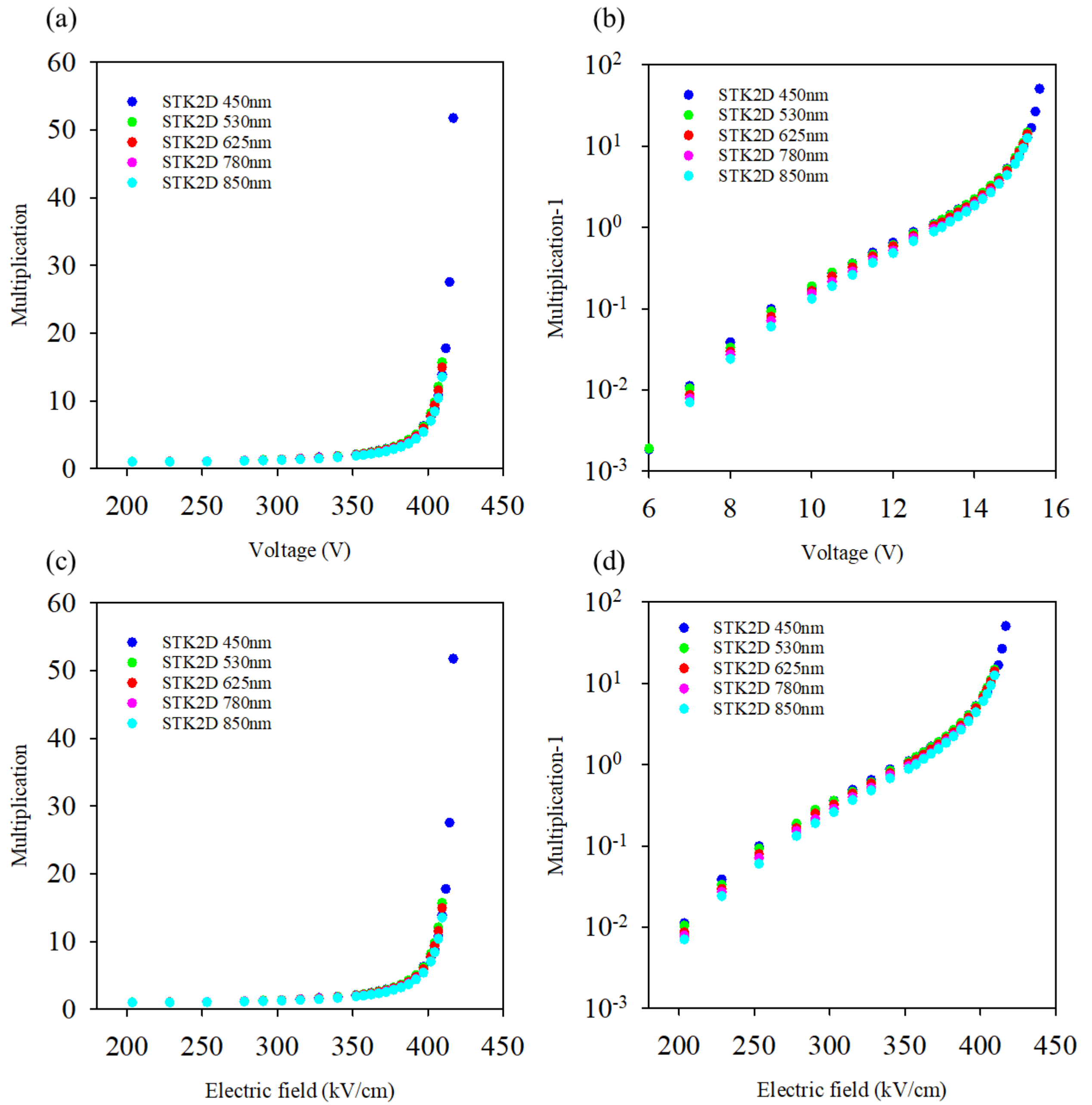


Figure 5.16 (a) M_e (Blue dots), M_{530} (Green dots), M_{625} (Red dots), M_{780} (Pink dots), M_{mix} (Cyan dots) of STK2D with at different voltage; (b) M_e -1 (Blue dots), M_{530} -1 (Green dots), M_{625} -1 (Red dots), M_{780} -1 (Pink dots), M_{mix} -1 (Cyan dots) of STK2D with at different voltage; (c) M_e (Blue dots), M_{530} (Green dots), M_{625} (Red dots), M_{780} (Pink dots), M_{mix} (Cyan dots) of STK2D with at different electric field; (d) M_e -1 (Blue dots), M_{530} -1 (Green dots), M_{625} -1 (Red dots), M_{780} -1 (Pink dots), M_{mix} -1 (Cyan dots) of STK2D with at different electric field

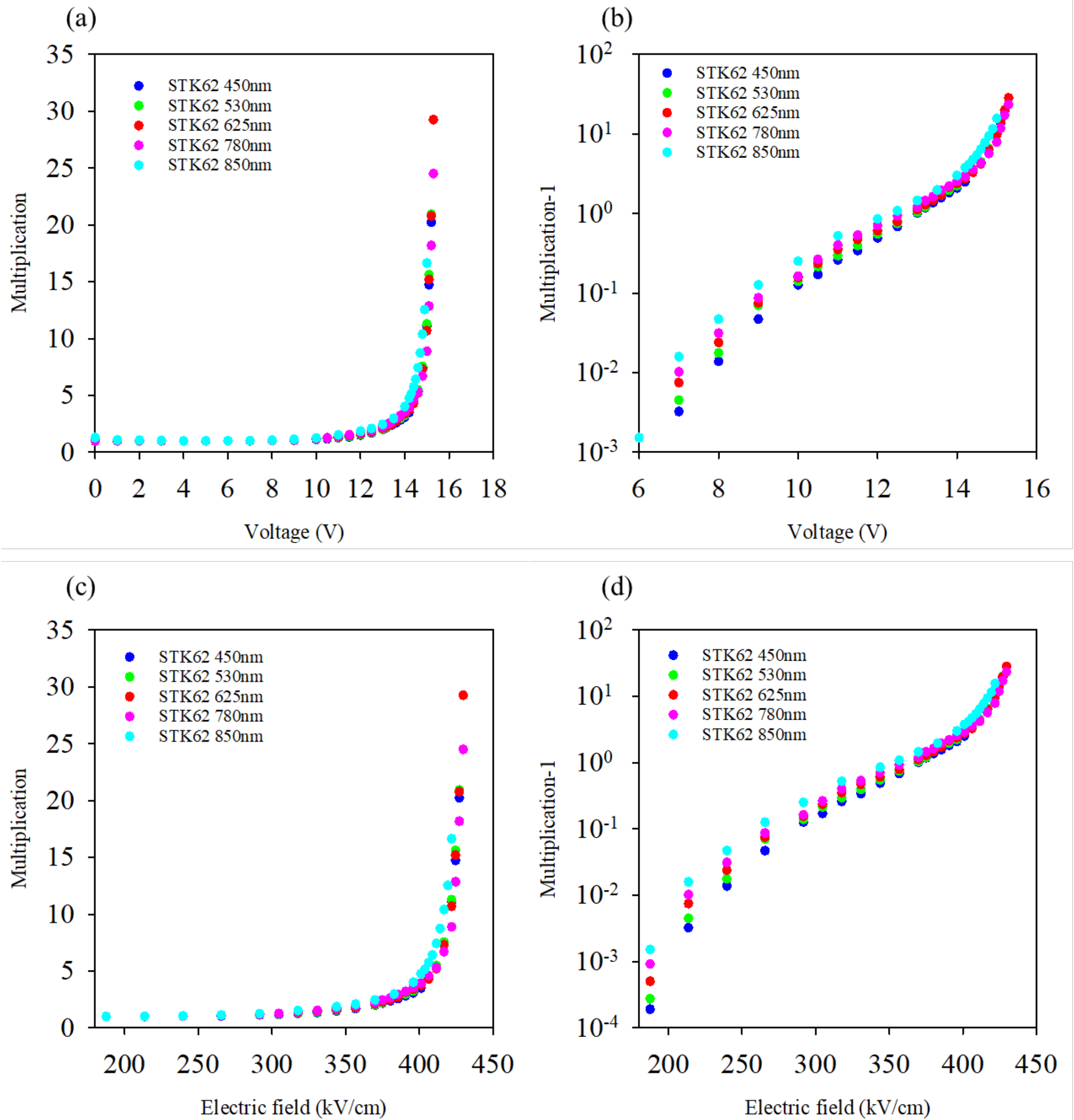


Figure 5.17 (a) M_h (Blue dots), M_{530} (Green dots), M_{625} (Red dots), M_{780} (Pink dots), M_{mix} (Cyan dots) of STK62 with at different voltage; (b) M_h -1 (Blue dots), M_{530} -1 (Green dots), M_{625} -1 (Red dots), M_{780} -1 (Pink dots), M_{mix} -1 (Cyan dots) of STK62 with at different voltage; (c) M_h (Blue dots), M_{530} (Green dots), M_{625} (Red dots), M_{780} (Pink dots), M_{mix} (Cyan dots) of STK62 with at different electric field; (d) M_h -1 (Blue dots), M_{530} -1 (Green dots), M_{625} -1 (Red dots), M_{780} -1 (Pink dots), M_{mix} -1 (Cyan dots) of STK62 with at different electric field

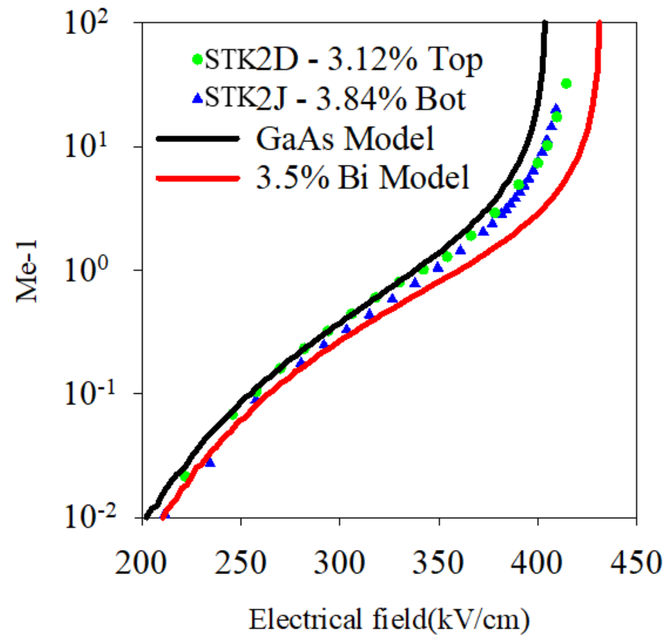


Figure 5.18 M_e -1 (dots and triangles) verse electric field of heterostructures with RPL simulation (solid lines)

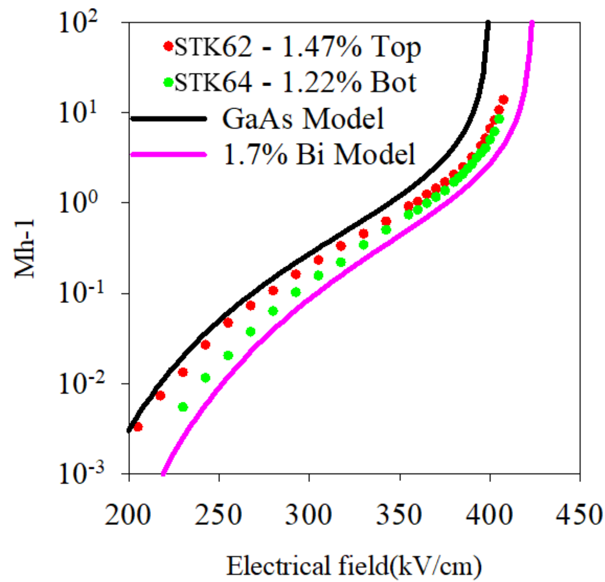


Figure 5.19 M_h -1 (dots) verse electric field of heterostructures with RPL simulation (solid lines)

Figure 5.18

Black line: Simulation for a 400 nm i-region of pure GaAs

Red line: Simulation for a 400 nm i-region of GaAsBi with 3.5 % Bi

Measured data: $(M_e - 1)$ versus electric field for STK2D (top 3.12 % Bi) and STK2J (bottom 3.84 % Bi)

Figure 5.19

Black line: Simulation for a 400 nm i-region of pure GaAs

Pink line: Simulation for a 400 nm i-region of GaAsBi with 1.7 % Bi

Measured data: $(M_h - 1)$ or $(M_{\text{mix}} - 1)$ versus electric field for STK62 (top 1.47 % Bi) and STK64 (bottom 1.22 % Bi)

All four photodiodes have 400 nm i-regions, half (200 nm) of which is GaAsBi. STK2D and STK2J are p–i–n devices, while STK62 and STK64 are n–i–p devices, differing in both Bi fraction and layer placement (top or bottom). Because shorter-wavelength excitation (M_e) often emphasizes electron-initiated avalanching, and longer-wavelength (M_h or M_{mix}) highlights hole-involved processes, the resulting curves reveal how GaAsBi modifies electron/hole transport under different bias conditions [12].

Convergence at High Bias (Large Electric Fields)

In Figures 5.18 and 5.19, one observes that, at sufficiently high bias—often above ~300 kV/cm—pairs of devices with similar Bi contents (STK2D vs. STK2J, STK62 vs. STK64) converge toward each other and closely track their corresponding GaAsBi simulation curves. For instance:

Figure 5.18:

Both STK2D (3.12 % Bi) and STK2J (3.84 % Bi) show steep rises in $(M_e - 1)$ above ~350–400 kV/cm, lying near the red "3.5 % Bi" curve and deviating distinctly from the black GaAs reference.

Figure 5.19:

STK62 (1.47 % Bi) and STK64 (1.22 % Bi) similarly cluster around the pink "1.7 % Bi" line for fields ~300–400 kV/cm, moving away from the black GaAs simulation.

This indicates that under high fields, the ultimate avalanche level is primarily dictated by Bi concentration: devices with near-equal Bi percentages exhibit nearly the same high-field multiplication profile, matching the GaAsBi model more closely than GaAs. In this regime, whether Bi is placed at the top or bottom becomes secondary, as carriers with sufficient energy can traverse the entire i-region and fully experience the GaAsBi band-structure modifications.

Position Dependence at Low Bias (Smaller Electric Fields)

At lower fields (~200–300 kV/cm), however, bottom-Bi devices tend to align more closely with the GaAsBi simulation, whereas top-Bi devices more closely match the pure GaAs model. This can be seen in both figures:

Figure 5.18:

STK2J (3.84 % Bi at the bottom) hugs the red "3.5 % Bi" curve at ~250–300 kV/cm.

STK2D (3.12 % Bi at the top) remains nearer the black GaAs line until higher fields.

Figure 5.19:

STK64 (1.22 % Bi at the bottom) more closely follows the pink "1.7 % Bi" curve at lower fields.

STK62 (1.47 % Bi at the top) stays nearer the black GaAs baseline until the bias increases further.

Hence, at moderate or low electric fields, if the GaAsBi layer is at the bottom of the i-region, hole-involved or mixed multiplication will reflect the GaAsBi bandgap properties earlier (i.e., at lower bias). Conversely, if Bi is at the top, at weak fields carriers cannot fully traverse that Bi region, and the device initially behaves more like pure GaAs [28], [29], [30], [31]. Only once the bias rises sufficiently do the carriers experience the Bi-rich layer, shifting the multiplication curve toward the GaAsBi model.

Conclusions and Implications

Dominance of Bi Fraction at High Bias

Under large fields, what most matters is how much Bi is present overall; devices with similar Bi contents show nearly identical high-voltage avalanche slopes, closely matching the corresponding GaAsBi simulation line.

Sensitivity to Bi Placement at Low Bias

Placing the Bi-rich layer at the bottom of the i-region reveals its band effects earlier in the avalanche process, whereas top-Bi devices act more like GaAs at low bias and only shift toward GaAsBi-like behavior at higher fields [12].

Device-Design Perspective

To exploit GaAsBi advantages at relatively low bias, one might place Bi in the lower portion of the i-region so that hole carriers (or combined carriers) quickly encounter the modified band structure. Otherwise, top-Bi arrangements could require higher voltages before the avalanche gains significant GaAsBi character.

Overall, Figures 5.18 and 5.19 illustrate that high-field avalanche gains are primarily set by the Bi fraction, while at low fields the difference between a top-Bi and bottom-Bi i-region has a more pronounced effect. This behavior highlights the importance of

layer positioning within the i-region for fine-tuning avalanche onset, gain slope, and potential noise performance in GaAsBi heterostructure photodiodes.

5.8 Discussion

In this chapter, we integrate the findings from the previous sections (5.1–5.7) to discuss the effects of incorporating GaAsBi layers of varying Bi content and placement within p–i–n / n–i–p GaAs structures. We compare the observed material and device performance with existing literature on avalanche photodiodes (APDs) and bandgap-engineered semiconductors, highlighting the potential of GaAsBi for reducing hole impact ionization coefficients (β) and extending near-infrared detection.

5.8.1 GaAsBi Epitaxy and Material Quality

From the SIMS and XRD analyses in Sections 5.2–5.4, it is clear that:

Bi Incorporation Accuracy and Uniformity

The Bi concentration in each 200 nm GaAsBi layer matches the design values (1.22%–3.84%), indicating a high level of control during molecular beam epitaxy (MBE). This agrees with previous reports that dilute bismide alloys can be precisely tuned under appropriate growth conditions [19].

Strain and Lattice Mismatch

For samples with relatively high Bi fractions (>3%), the XRD peaks shift further from the GaAs substrate peak, suggesting increased strain and partial relaxation [15][16][27]. Although no severe surface degradation was observed in this work, further increasing the Bi content or layer thickness would require more elaborate strain management (e.g., graded buffers or superlattice structures) to avoid dislocation-related leakage currents.

5.8.2 Electrical Characteristics and Depletion Region Control

Section 5.5 presented the current–voltage (I–V) and capacitance–voltage (C–V) measurements on these heterostructure photodiodes:

Dark Current and Breakdown Voltage

All devices exhibit typical diode-like behavior, and the reverse dark current scales reasonably with device area. The breakdown voltage (ranging from about –10 V to –20 V, or higher) varies with Bi composition and doping profiles, consistent with reported values for GaAs-based APDs [2][28].

C–V Analysis and i-Region Thickness

As reverse bias increases, the depletion region expands to approximately 400 nm, in good agreement with the device design. The Poisson-based fitting aligns within ±10% of the nominal i-region thickness, confirming the reliability of the MBE growth and doping control. While high-Bi samples may exhibit slight local non-uniformity, overall depletion profiles are sufficiently well-defined for avalanche operation.

5.8.3 Optical Response and Bandgap Tuning

In Section 5.6, photocurrent spectra from 700 to 1200 nm demonstrate:

Extended Near-Infrared Absorption

Absorption edges near ~830 nm (GaAs) and ~1000 nm (GaAsBi) were observed, with higher Bi fractions producing stronger red-shifts and enhanced near-IR response [12][19]. This outcome corroborates the band anticrossing mechanism and increased spin–orbit splitting energy characteristic of dilute bismide alloys [13].

Franz–Keldysh Effects Under Bias

Forward biases from 0 to 4 V lead to notable absorption tail extensions (900–1100 nm range), reflecting enhanced carrier collection and possible shifts in the effective band edge. This behavior is more pronounced in high-Bi samples, where the reduced bandgap accentuates the Franz–Keldysh effect.

5.8.4 Avalanche Multiplication and Hole Ionization Suppression

Section 5.7 detailed the photomultiplication behavior under high electric fields:

Dominance of Bi Content at High Fields

Above ~300 kV/cm, pairs of devices with similar Bi contents (e.g., STK2D vs. STK2J, STK62 vs. STK64) tend to converge toward their respective GaAsBi simulation curves, deviating from the pure GaAs reference line. This highlights the role of Bi in suppressing hole impact ionization coefficients (β), thereby potentially reducing the excess noise factor k [2][14][31].

Position Dependence at Low to Moderate Fields

Bottom-Bi Devices (STK2J, STK64): At lower biases (~200–300 kV/cm), the avalanche process more quickly adopts GaAsBi-like characteristics, since carriers encounter the Bi-rich region earlier.

Top-Bi Devices (STK2D, STK62): At weaker fields, the device initially behaves closer to pure GaAs, and only as the voltage rises sufficiently do carriers traverse the Bi-rich layer, shifting the multiplication curve toward the GaAsBi model.

This confirms that not only does Bi composition affect high-field avalanche gains, but the placement of the Bi layer within the i-region also influences the onset of avalanche multiplication at lower biases.

5.8.5 Significance, Limitations, and Future Work

Implications for Low-Noise APDs

Incorporating Bi into GaAs can significantly raise the spin–orbit splitting energy, suppressing hole-initiated impact ionization [12][14]. Compared to designs relying on quantum wells or nanostructures, GaAsBi doping offers a more direct materials-based approach to reducing β, thus lowering the excess noise factor. The ability to vary Bi fraction and position within the i-region adds flexibility for tailoring gain slope, breakdown voltage, and operating bias.

Limitations

Strain and Defects: Higher Bi contents increase the risk of dislocation formation and strain relaxation, potentially elevating dark current.

Device Uniformity: I–V and C–V variations among different chip sizes hint at edge effects, doping gradients, or sidewall leakage, which require more refined fabrication processes.

Measurement and Modeling: High-field device behavior may involve localized heating and complex carrier transport. Monte Carlo simulations or transient characterization could clarify the interplay between fields, carriers, and defects.

Prospects for Future Research

Strain Management and Multi-Layer Structures: Implementing buffer or superlattice designs could enable higher Bi fractions (>4%) without compromising crystal quality [27].

High-Speed or Large-Area APDs: To maintain low noise under high-speed or large-area operation, more detailed multi-physics simulation and process optimization will be required.

Multi-Band Detection and Integration: With further bandgap engineering (e.g., dual-layer GaAsBi), broadband detection or integrated photonic applications could be envisioned, leveraging both the NIR extension and suppressed hole ionization.

Overall, the present study confirms that embedding GaAsBi layers in GaAs can effectively reduce hole impact ionization while extending the infrared detection range to ~1000 nm. By tuning the Bi fraction and placing it at different depths within the i-region, one can optimize avalanche characteristics across low and high bias regimes. Hence, GaAsBi-based heterostructures represent a promising pathway to achieve low-noise, high-sensitivity APDs with relatively straightforward epitaxial control.

References

[1] G. P. Agrawal, *Fiber – optic communication systems*, Jun. 2021. doi: 10.1002/9781119737391

[2] R. J. McIntyre, “Multiplication noise in uniform avalanche diodes,” *IEEE Transactions on Electron Devices*, vol. ED-13, no. 1, pp. 164–168, Jan. 1966. doi: 10.1109/t-ed.1966.15651

[3] J. Beck *et al.*, “The HgCdTe Electron Avalanche photodiode,” *Journal of Electronic Materials*, vol. 35, no. 6, pp. 1166–1173, Jun. 2006. doi: 10.1007/s11664-006-0237-3

[4] A. R. Marshall, J. P. David, and C. H. Tan, “Impact ionization in InAs electron avalanche photodiodes,” *IEEE Transactions on Electron Devices*, vol. 57, no. 10, pp. 2631–2638, Oct. 2010. doi: 10.1109/ted.2010.2058330

[5] “Si APD S16453-10K: Hamamatsu Photonics,” *Si APD S16453-10K | Hamamatsu Photonics*, https://www.hamamatsu.com/jp/en/product/optical-sensors/apd/si-apd/S16453-10K.html (accessed Jan. 15, 2025).

[6] X. Jin *et al.*, “Very low excess noise Very Low Excess Noise $Al_{0.75}Ga_{0.25}As_{0.56}Sb_{0.44}$ Avalanche Photodiode,” *Optics Express*, vol. 31, no. 20, p. 33141, Sep. 2023. doi: 10.1364/oe.500169

[7] S. Lee *et al.*, “High Gain, Low Noise, Room Temperature 1550 nm GaAsSb/AlGaAsSb Avalanche Photodiodes,” *Optica*, vol. 10, no. 2, p. 147, Jan. 2023. doi: 10.1364/optica.476963

[8] F. Capasso, Won-Tien Tsang, and G. F. Williams, “Staircase solid-state photomultipliers and avalanche photodiodes with enhanced ionization rates ratio,” *IEEE Transactions on Electron Devices*, vol. 30, no. 4, pp. 381–390, Apr. 1983. doi: 10.1109/t-ed.1983.21132

[9] M. Ren *et al.*, "AlInAsSb/GaSb staircase avalanche photodiode," *Applied Physics Letters*, vol. 108, no. 8, Feb. 2016. doi: 10.1063/1.4942370

[10] Y. Ma *et al.*, "Enhanced carrier multiplication in InAs quantum dots for bulk avalanche photodetector applications," *Advanced Optical Materials*, vol. 5, no. 9, Mar. 2017. doi: 10.1002/adom.201601023

[11] A. C. Farrell *et al.*, "Plasmonic field confinement for separate absorption-multiplication in InGaAs nanopillar avalanche photodiodes," *Scientific Reports*, vol. 5, no. 1, Dec. 2015. doi: 10.1038/srep17580

[12] Y. Liu *et al.*, "Valence band engineering of GaAsBi for low noise avalanche photodiodes," *Nature Communications*, vol. 12, no. 1, Aug. 2021. doi: 10.1038/s41467-021-24966-0

[13] M. Usman, C. A. Broderick, A. Lindsay, and E. P. O'Reilly, "Tight-binding analysis of the electronic structure of dilute bismide alloys of GaP and GaAs," *Physical Review B*, vol. 84, no. 24, Dec. 2011. doi: 10.1103/physrevb.84.245202

[14] İ. H. Oğuzman, Y. Wang, J. Kolník, and K. F. Brennan, "Theoretical study of hole initiated impact ionization in bulk silicon and GaAs using a wave-vector-dependent numerical transition rate formulation within an ensemble Monte Carlo Calculation," *Journal of Applied Physics*, vol. 77, no. 1, pp. 225–232, Jan. 1995. doi: 10.1063/1.359374

[15] J. W. Matthews and A. E. Blakeslee, "Defects in epitaxial multilayers: III. Preparation of almost perfect multilayers," *Journal of Crystal Growth*, vol. 32, no. 2, pp. 265–273, Feb. 1976. doi: 10.1016/0022-0248(76)90041-5

[16] R. France, C.-S. Jiang, and A. J. Ptak, "In situ strain relaxation comparison between GaAsBi and GaInAs grown by molecular-beam epitaxy," *Applied Physics Letters*, vol. 98, no. 10, Mar. 2011. doi: 10.1063/1.3562952

[17] R. D. Richards *et al.*, “Photovoltaic characterisation of GaAsBi/GaAs multiple quantum well devices,” *Solar Energy Materials and Solar Cells*, vol. 172, pp. 238–243, Dec. 2017. doi: 10.1016/j.solmat.2017.07.029

[18] K. Alberi *et al.*, “Valence band anticrossing in GaBixAs $_{1-x}$,” *Applied Physics Letters*, vol. 91, no. 5, Jul. 2007. doi: 10.1063/1.2768312

[19] S. Tixier *et al.*, “Molecular beam epitaxy growth of GaAs1−xBix,” *Appl. Phys. Lett.*, vol. 82, no. 14, pp. 2245–2247, Apr. 2003, doi: 10.1063/1.1565499.

[20] M. Henini *et al.*, “Molecular beam epitaxy of GaBiAs on (311)B GaAs substrates,” *Appl. Phys. Lett.*, vol. 91, no. 25, p. 251909, Dec. 2007, doi: 10.1063/1.2827181.

[21] J. Puustinen *et al.*, “Variation of lattice constant and cluster formation in GaAsBi,” *J. Appl. Phys.*, vol. 114, no. 24, p. 243504, Dec. 2013, doi: 10.1063/1.4851036.

[22] M. Beaudoin *et al.*, “Bandedge optical properties of MBE grown GaAsBi films measured by photoluminescence and photothermal deflection spectroscopy,” *J. Cryst. Growth*, vol. 425, pp. 245–249, Sep. 2015, doi: 10.1016/j.jcrysgro.2015.01.019.

[23] F. Dybala *et al.*, “Electromodulation spectroscopy of heavy-hole, light-hole, and spin-orbit transitions in GaAsBi layers at hydrostatic pressure,” *Appl. Phys. Lett.*, vol. 111, no. 19, p. 192104, Nov. 2017, doi: 10.1063/1.5002622.

[24] C. Ryang Wie, “High resolution x-ray diffraction characterization of semiconductor structures,” *Mater. Sci. Eng. R Rep.*, vol. 13, no. 1, pp. 1–56, Sep. 1994, doi: 10.1016/0927-796X(94)90008-6.

[25] H. Achour, S. Louhibi, B. Amrani, A. Tebboune, and N. Sekkal, “Structural and electronic properties of GaAsBi,” *Superlattices Microstruct.*, vol. 44, no. 2, pp. 223–229, Aug. 2008, doi: 10.1016/j.spmi.2008.05.004.

[26] D. F. Reyes *et al.*, “Bismuth incorporation and the role of ordering in GaAsBi/GaAs structures,” *Nanoscale Res. Lett.*, vol. 9, no. 1, p. 23, Dec. 2014, doi: 10.1186/1556-276X-9-23.

[27] T. B. O. Rockett *et al.*, "Influence of growth conditions on the structural and opto-electronic quality of GaAsBi," *J. Cryst. Growth*, vol. 477, pp. 139–143, Nov. 2017, doi: 10.1016/j.jcrysgro.2017.02.004.

[28] A. G. Chynoweth, "Ionization rates for electrons and holes in Silicon," *Physical Review*, vol. 109, no. 5, pp. 1537–1540, Mar. 1958. doi: 10.1103/physrev.109.1537

[29] M. M. Hayat, B. E. Saleh, and M. C. Teich, "Effect of Dead Space on Gain and Noise of Double-Carrier-Multiplication Avalanche Photodiodes," *Optical Society of America Annual Meeting*, 1991. doi: 10.1364/oam.1991.fu2

[30] R. J. McIntyre, "A new look at impact ionization-PART I: A theory of gain, noise, breakdown probability, and frequency response," *IEEE Transactions on Electron Devices*, vol. 46, no. 8, pp. 1623–1631, 1999. doi: 10.1109/16.777150

[31] S. A. Plimmer et al., "Investigation of impact ionization in thin GaAs diodes," *IEEE Transactions on Electron Devices*, vol. 43, no. 7, pp. 1066–1072, Jul. 1996. doi: 10.1109/16.502416

[32] C. W. Magee, "Secondary ion mass spectrometry quantification: Do we need more standards?", Journal of Vacuum Science & Technology B, vol. 41, no. 3, p. 030803, 2023, doi: 10.1116/6.0002466.

[33] International Organization for Standardization, "ISO 18114:2021 Surface chemical analysis – Secondary-ion mass spectrometry – Determination of relative sensitivity factors from ion-implanted reference materials," ISO, 2021.

[34] P. E. Fewster, "Reciprocal space mapping," Critical Reviews in Solid State and Materials Sciences, vol. 22, no. 2, pp. 69–110, 1997, doi: 10.1080/10408439708241259.

Chapter 6: Characterisation of GaAs/GaAsBi MQW photodiodes

6.1 Introduction

Various material systems have been studied in recent years. For example, HgCdTe [3] and InAs [4] display negligible hole ionization and thus achieve near-ideal internal gain with minimal excess noise, but their narrow bandgaps necessitate cryogenic operation to suppress dark currents [5]. By contrast, wide-bandgap semiconductors such as silicon [5] and AlGaAsSb [6][7] can operate at room temperature while maintaining relatively low dark current and moderate excess noise. To reduce the intrinsic similarity between α and β in certain materials, researchers have adopted multiple approaches, including embedding multiple quantum wells (MQWs) [8], employing “staircase” band-edge designs [9] to provide extra carrier energy, or using quantum dots [10] and nanostructures [11] in the avalanche region to favor one type of carrier. However, these strategies often yield only incremental improvements and demand stringent control over material growth and device design.

Notably, incorporating the large group-V element bismuth (Bi) into GaAs has been identified as a relatively straightforward yet highly effective method to suppress hole ionization while leaving the electron ionization coefficient almost unchanged [12]. This is attributed to a band anti-crossing interaction in the valence band of GaAs upon Bi doping, which increases the spin-orbit splitting energy (Δ_{SO}) and consequently reduces the number of holes accessing the split-off band for ionization [13][14]. Nevertheless, excessive Bi incorporation raises compressive strain, potentially exceeding the critical layer thickness [15] and generating dislocations that increase dark currents [16][17].

The present study focuses on introducing GaAsBi quantum wells (QWs) with the same well design but varying numbers of periods, stacking them onto a GaAs substrate to improve and enhance avalanche performance. Compared to growing thick GaAsBi layers in a single step, this multilayer QW method aims to precisely control the overall strain while strategically modifying the local valence band structure to reduce the

probability of hole impact ionization. Because Bi has a large atomic radius and a significant electronegativity contrast with As, it strongly perturbs the valence band of GaAs; this not only narrows the bandgap via a band anti-crossing mechanism [18], but also notably increases the spin-orbit splitting energy, thereby lowering the hole ionization coefficient β. In this chapter, we demonstrate a systematic avalanche multiplication study of multiple GaAsBi/GaAs MQWs, each with different stacking periods in a p-i-n structure, and evaluate their impact ionization characteristics via photomultiplication measurements.

6.2 Layer details

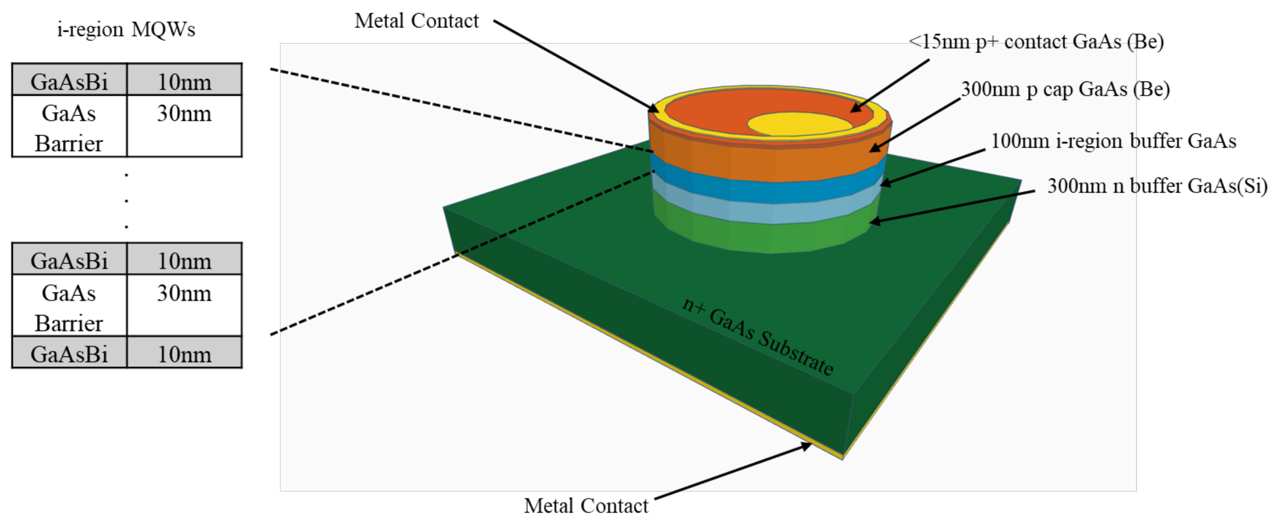


Figure 6.1 Schematic diagram of cross-section of the MQW p-i-n device structures

Table 6.1 Design of MQW samples

Layer	Number of periods, N	Barrier thickness (nm), L_W	Barrier thickness (nm), L_B	i-region thickness (nm), w	MQW $_{Bi(\%)}$
02QW	2	10	30	255	5
05QW	5	10	30	375	5
15QW	15	10	30	775	5
30QW	30	10	30	1375	5
40QW	40	10	30	1775	5
80QW	80	10	30	3375	5
120QW	120	10	30	4975	5

Each structure comprises GaAsBi wells of 10nm thickness (with $\sim 5\%$ *Bi) and 30,nm GaAs barriers, stacked to form the i-region. As* N *increases,* w *grows accordingly.*

Figure 6.1 illustrates a schematic cross-section of the p-i-n MQW photodiode structures investigated in this work. All samples were epitaxially grown on an n^+ GaAs substrate and designed to form a p-i-n configuration. As shown in the figure, the top side includes a thin (<15nm) p $^+$ GaAs contact layer (Be-doped), followed by a 300nm p-cap GaAs layer, whereas the lower part has a 300nm n-buffer GaAs (Si-doped). Between these doped layers resides the intrinsic MQW region, which consists of multiple GaAsBi/GaAs repeats.

In contrast to the approach of using thick GaAsBi slabs, the design here adopts a uniform well thickness of 10nm and a 30nm GaAs barrier in each period. The bismuth (Bi) composition is consistently near 5% in the quantum wells, ensuring that the band structure is modified without causing excessive strain. By controlling the number of

repeated QW periods from 2 up to 120, the total i-region thickness (w) varies significantly, as summarized in Table 6.1. For instance, 2QW exhibits just 2 repeats and an i-region of 255nm, whereas 120QW features 120 repeats and an i-region exceeding 4.9µm.

Because Bi has a large atomic radius and strong electronegativity difference relative to As, even a moderate Bi fraction of $\sim 5\%$ can significantly perturb the valence band edges and spin-orbit splitting energy in GaAsBi layers [1]. By distributing these thin GaAsBi quantum wells among GaAs barriers, the design seeks to fine-tune the material's optical and electrical behavior while avoiding the potential pitfalls (such as high dislocation densities) associated with a single thick GaAsBi layer [19][20]. The p-i-n arrangement, with Be-doped p-layers and Si-doped n-layers, aims to confine the depletion region around the MQWs under reverse bias, thus enabling an effective photodiode structure for subsequent avalanche and photocurrent characterization.

6.3 XRD Characterization of MQW Diodes

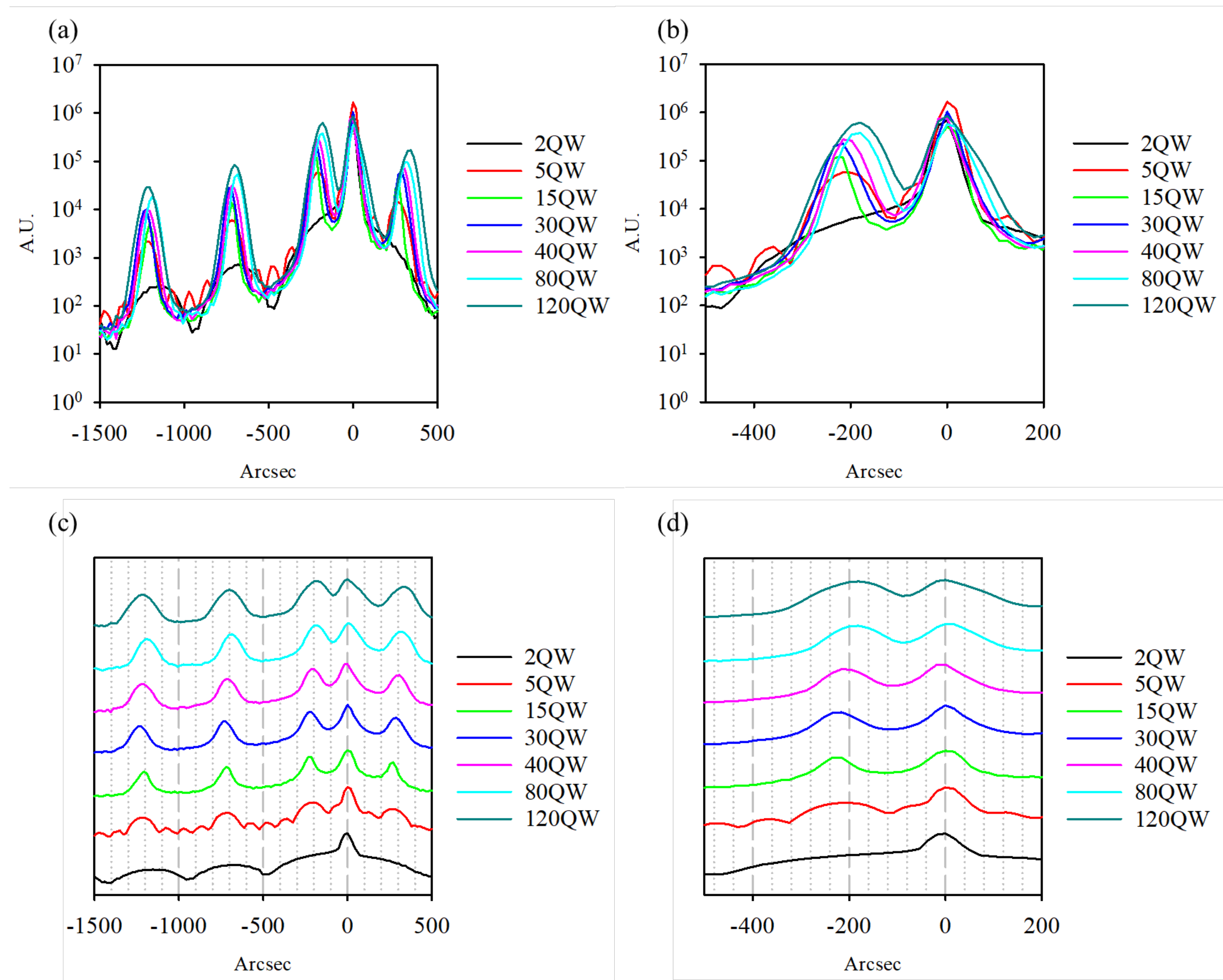


Figure 6.2 X-ray diffraction results of MQW diodes

Table 6.2 XRD fitting results of MQW diodes

Layer	Well number	Well thickness (nm)	Barrier thickness (nm)	QW period (nm)	Well Bi (%)	i-region thickness (nm)	i-region average Bi (%)
2QW	2	8	32.5	40.5	4.3	251	0.2741
5QW	5	6.6	30.9	37.5	4.5	357.5	0.41538
15QW	15	6.1	32.5	38.6	5.1	749	0.62303
30QW	30	6.5	30	36.5	4.6	1265	0.70909
40QW	40	5.7	31.1	36.8	5.2	1642	0.72205
80QW	80	5.4	32	37.4	5.5	3162	0.75142
120QW	120	5.4	31.2	36.6	5.5	4562	0.78124

Figure 6.2 compares the measured diffraction patterns for these MQW diodes. Panels (a) and (c) present broad-range scans (e.g., -1500 to +500 arcsec), where the main GaAs substrate peak occurs around 0 arcsec and multiple satellite fringes arise from the periodic well–barrier structure. Panels (b) and (d) zoom in on -400 to +200 arcsec to highlight subtle changes in fringe spacing and intensity.

Fringe Number and Separation

As the period count (N) increases, the total i-region thickness also grows, and the interference fringes become more closely spaced. 2QW, for instance, exhibits fewer satellite peaks than 120QW, which shows additional, more closely spaced oscillations. These satellites indicate both the overall thickness and the repeated well–barrier intervals.

Well Thickness and Bi Composition

Curve fitting suggests that each GaAsBi well remains close to the nominal 10 nm, though measured thicknesses can shift slightly (e.g., 5.4–6.6 nm in some samples). The Bi fraction in the wells is near 5 %, varying from approximately 4.3 % to 5.5 %. Such differences can result from local growth temperature variations, flux ratio fluctuations, or strain effects [21] [26] [27].

Average Bi in the i-region

As shown in Table 6.2, the average Bi content across the entire i-region is noticeably lower than 5 %, since GaAs barriers occupy most of the thickness. 2QW, for example, has an i-region average of about 0.27 % Bi, whereas 120QW—with more well material—reaches around 0.78 % Bi overall.

Increasing the MQW period count extends the i-region, leading to more pronounced and shifted satellite fringes in XRD scans [23] [28]. Minor deviations in well or barrier thickness can further alter fringe amplitude and spacing. The measured well Bi contents, from roughly 4.3 % to 5.5 %, align with the design target of 5 %, though local composition gradients may appear [22] [27]. In high-period structures (e.g., 80QW, 120QW), the cumulative lattice mismatch grows, potentially raising the risk of extended defects if the strain budget is exceeded [23] [26] [29]. Nevertheless, the strong satellite fringes visible in every sample confirm relatively uniform well–barrier interfaces.

These XRD findings form the basis for subsequent optical and electrical analyses. By correlating layer structures with photoluminescence or photoresponse data, one can examine how differences in well/barrier thicknesses or slight Bi-content variations affect device characteristics. Potential approaches such as strain compensation or reduced-temperature growth might further optimize interface quality for larger-period MQW designs [21] [26].

6.4 Electrical characterisation of MQW Diodes

6.4.1 IV measurements

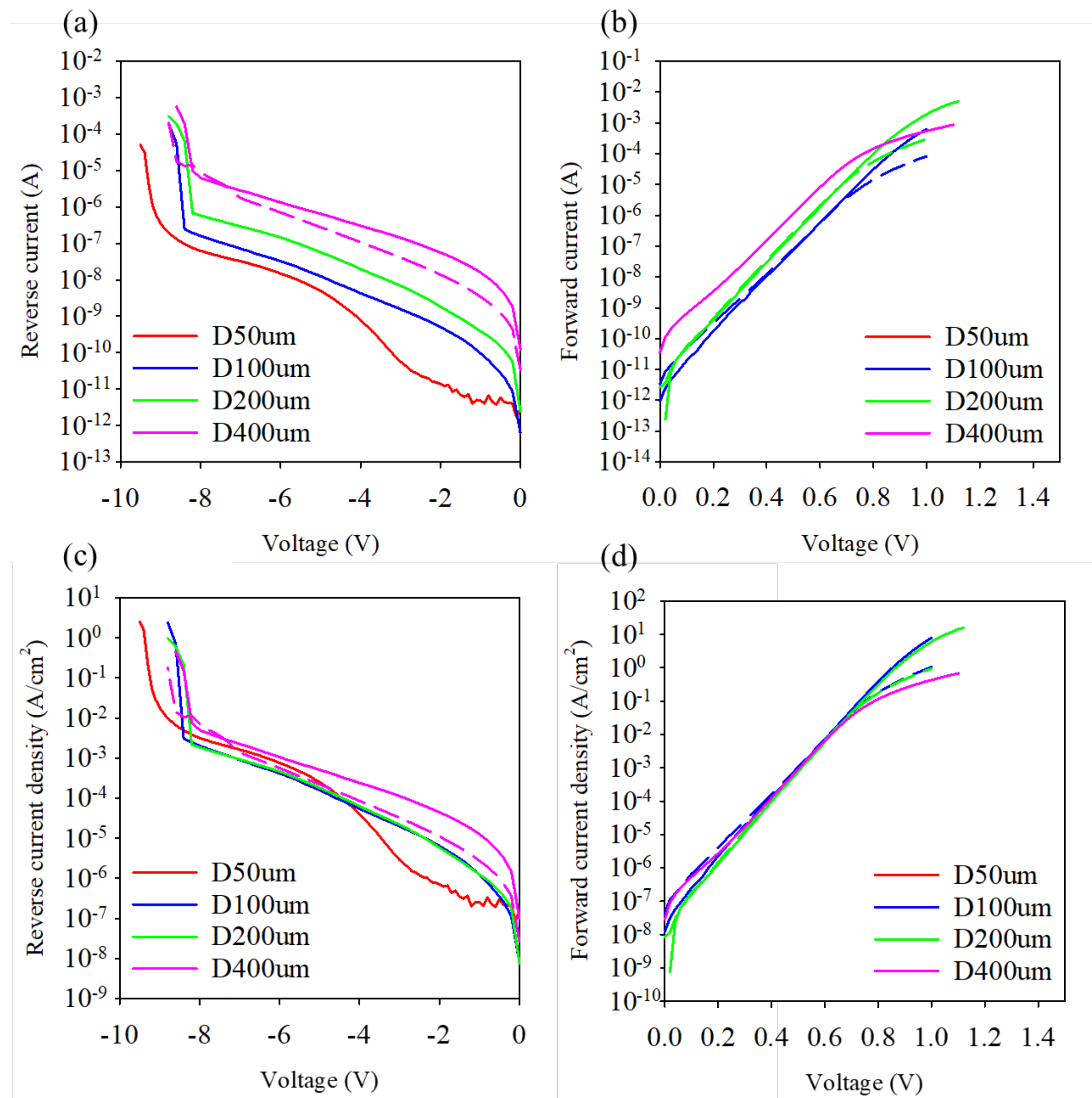


Figure 6.3: (a) Reverse dark current, (b) forward dark current, (c) reverse dark current density, and (d) forward dark current density of the 2QW photodiodes measured for multiple nominal mesa diameters (D = 50, 100, 200, and 400 μm). In this and the following diameter-dependent plots in Chapter 6, traces with the same colour correspond to the same nominal diameter, while solid and dashed curves represent measurements from different nominally identical devices of that same diameter, thereby visualising device-to-device spread at fixed geometry.

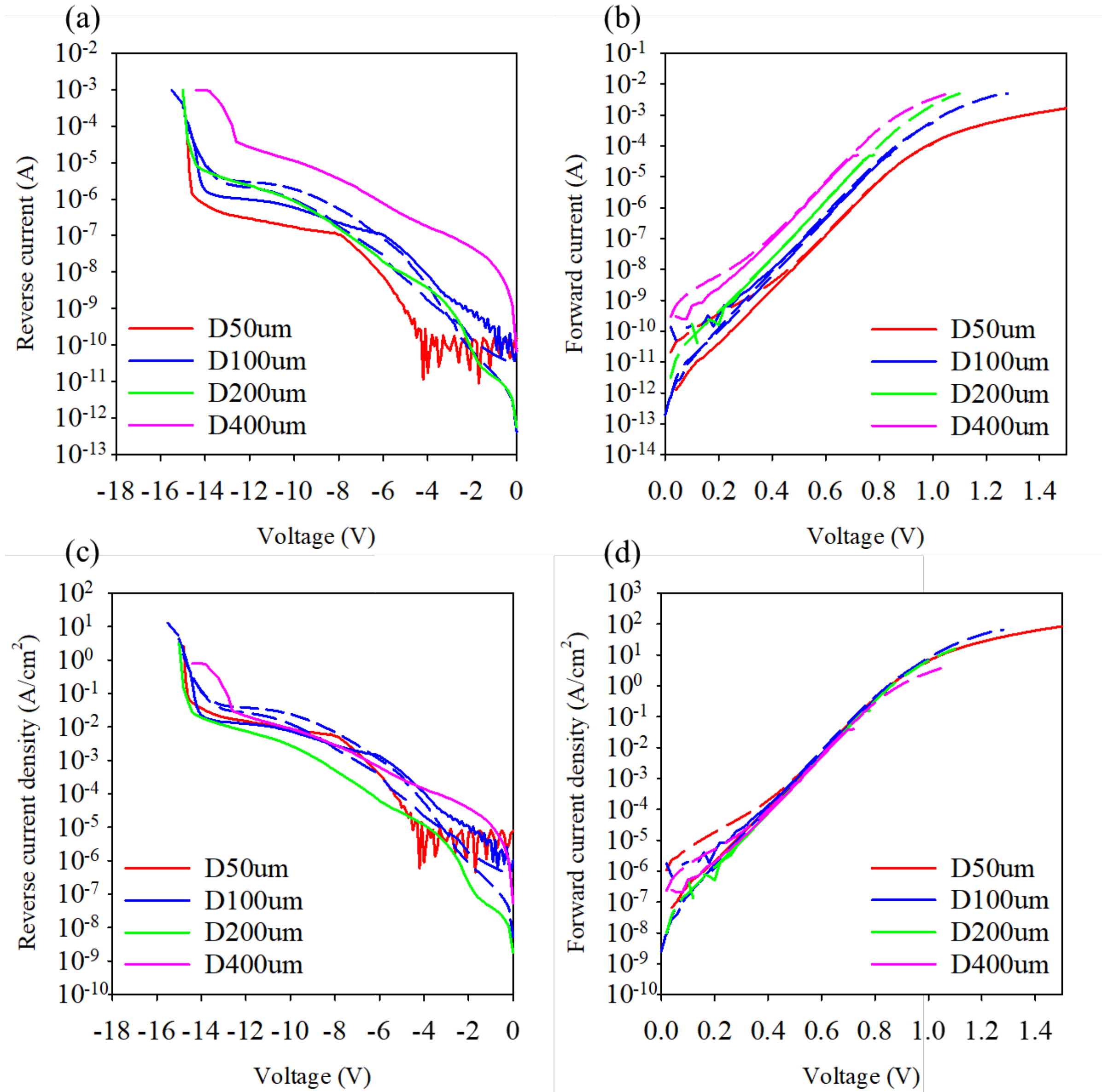


Figure 6.4 (a) Reverse dark current, (b) forward dark current, (c) reverse dark current density, (d) forward dark current density of 5QW

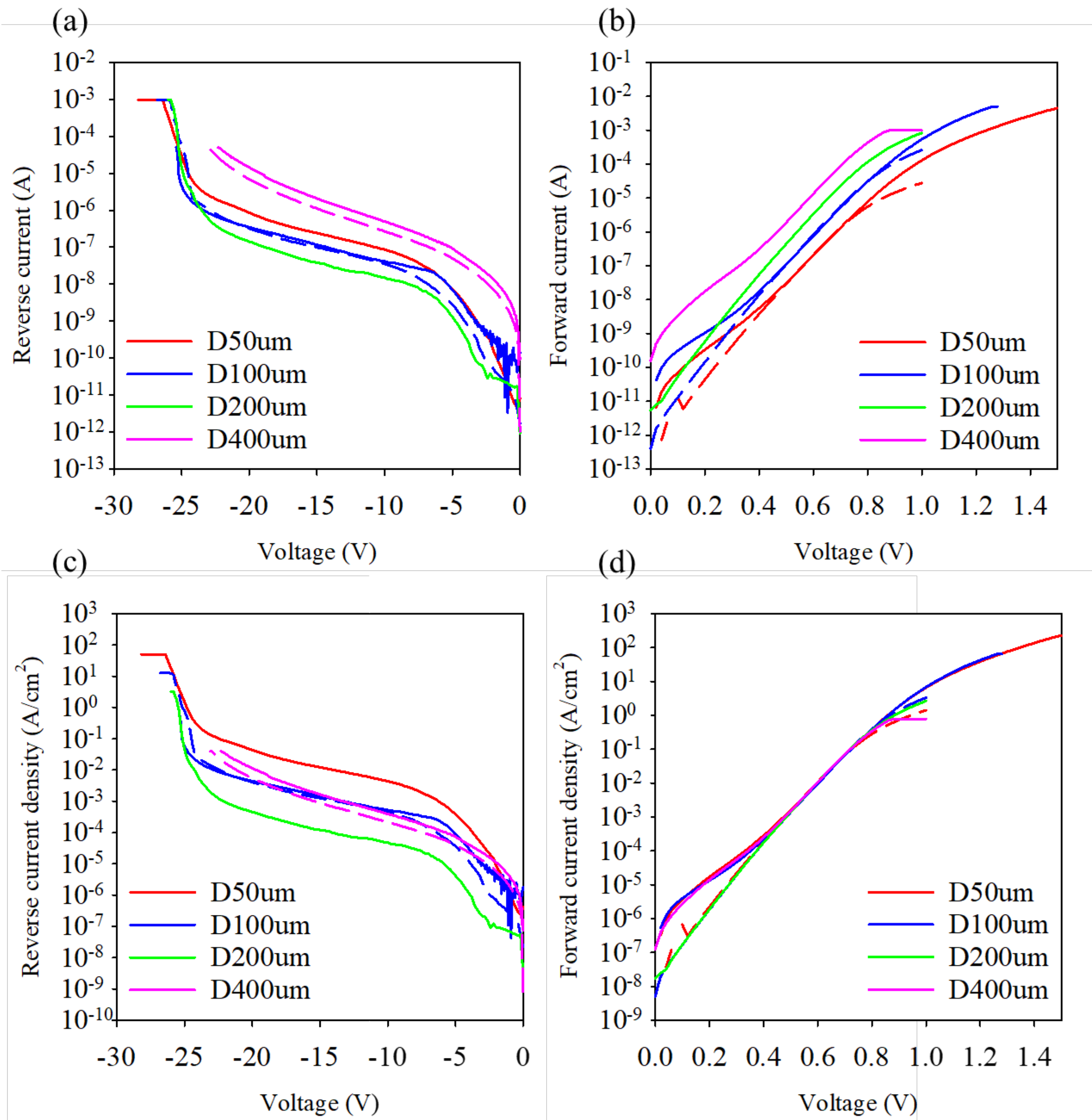


Figure 6.5 (a) Reverse dark current, (b) forward dark current, (c) reverse dark current density, (d) forward dark current density of 15QW

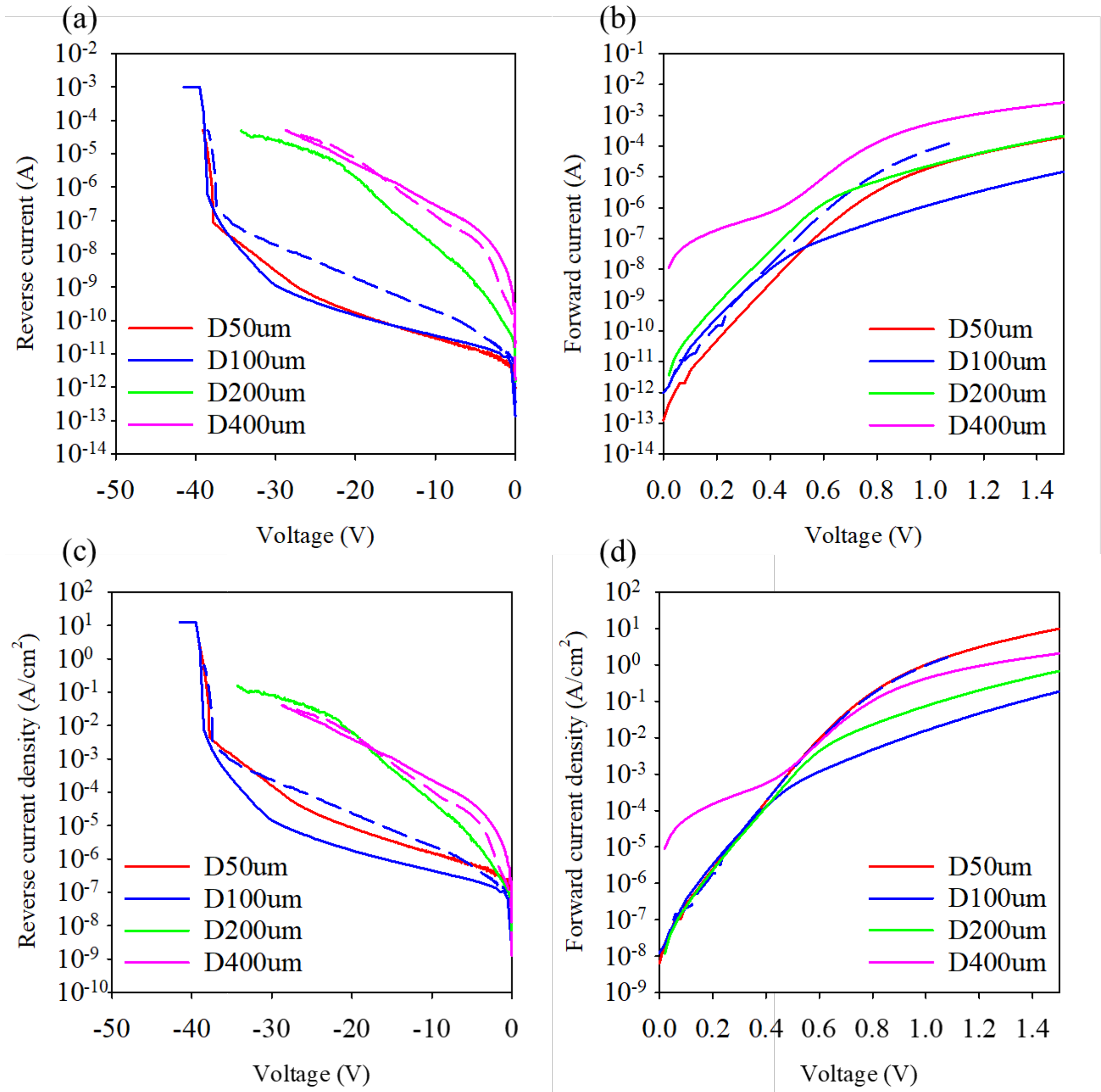


Figure 6.6 (a) Reverse dark current, (b) forward dark current, (c) reverse dark current density, (d) forward dark current density of 30QW

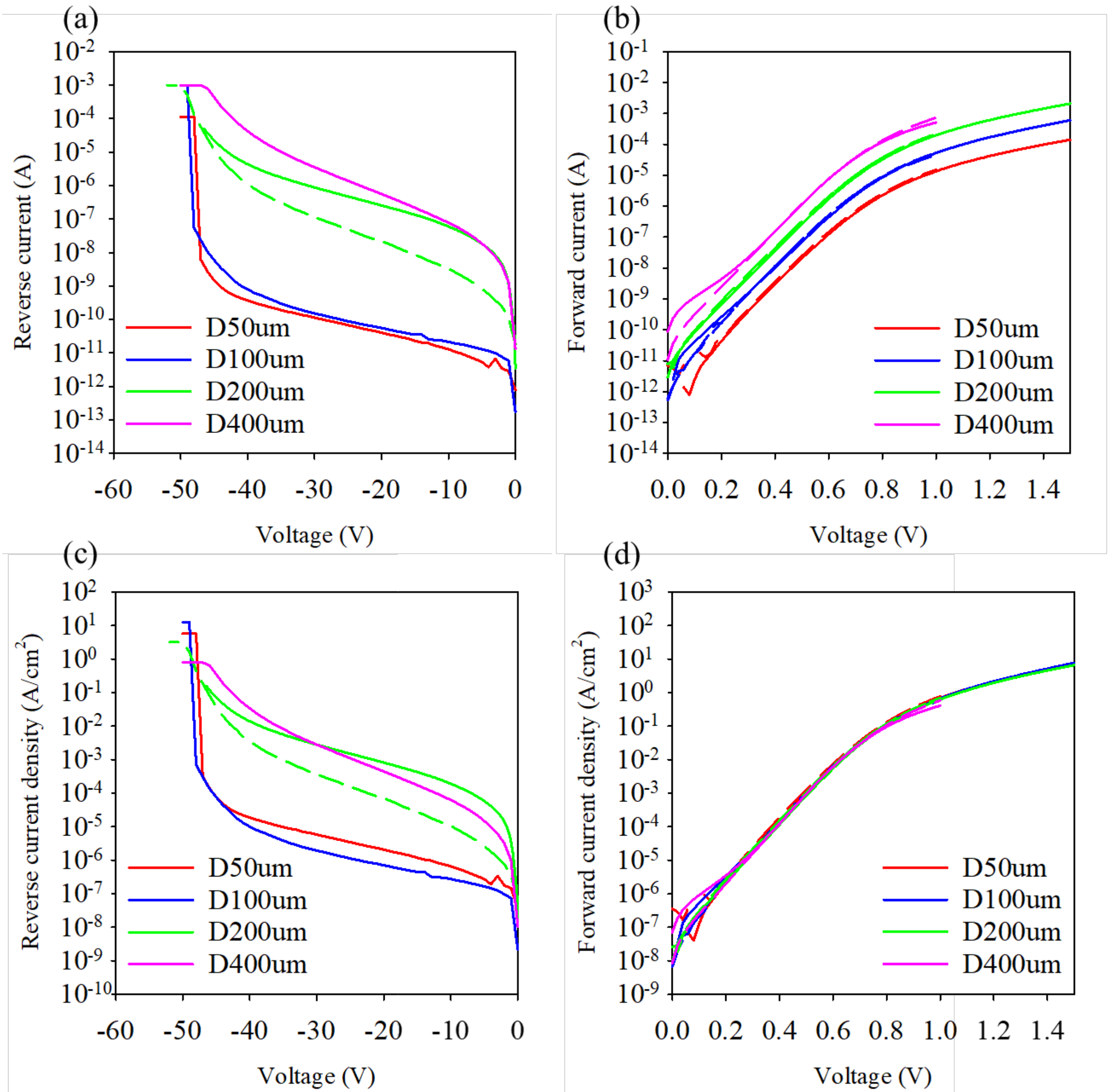


Figure 6.7 (a) Reverse dark current, (b) forward dark current, (c) reverse dark current density, (d) forward dark current density of 40QW

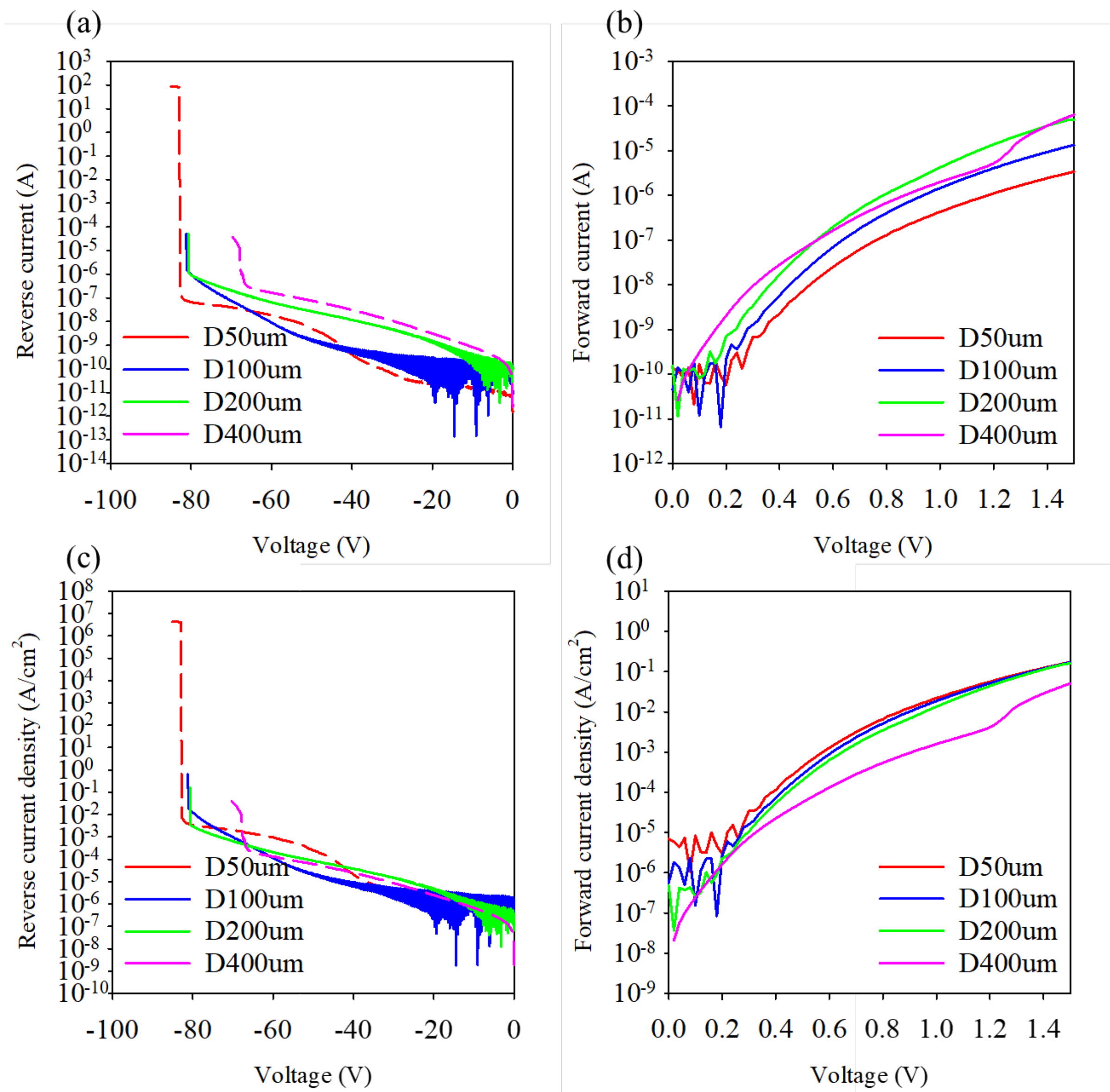


Figure 6.8 (a) Reverse dark current, (b) forward dark current, (c) reverse dark current density, (d) forward dark current density of 80QW

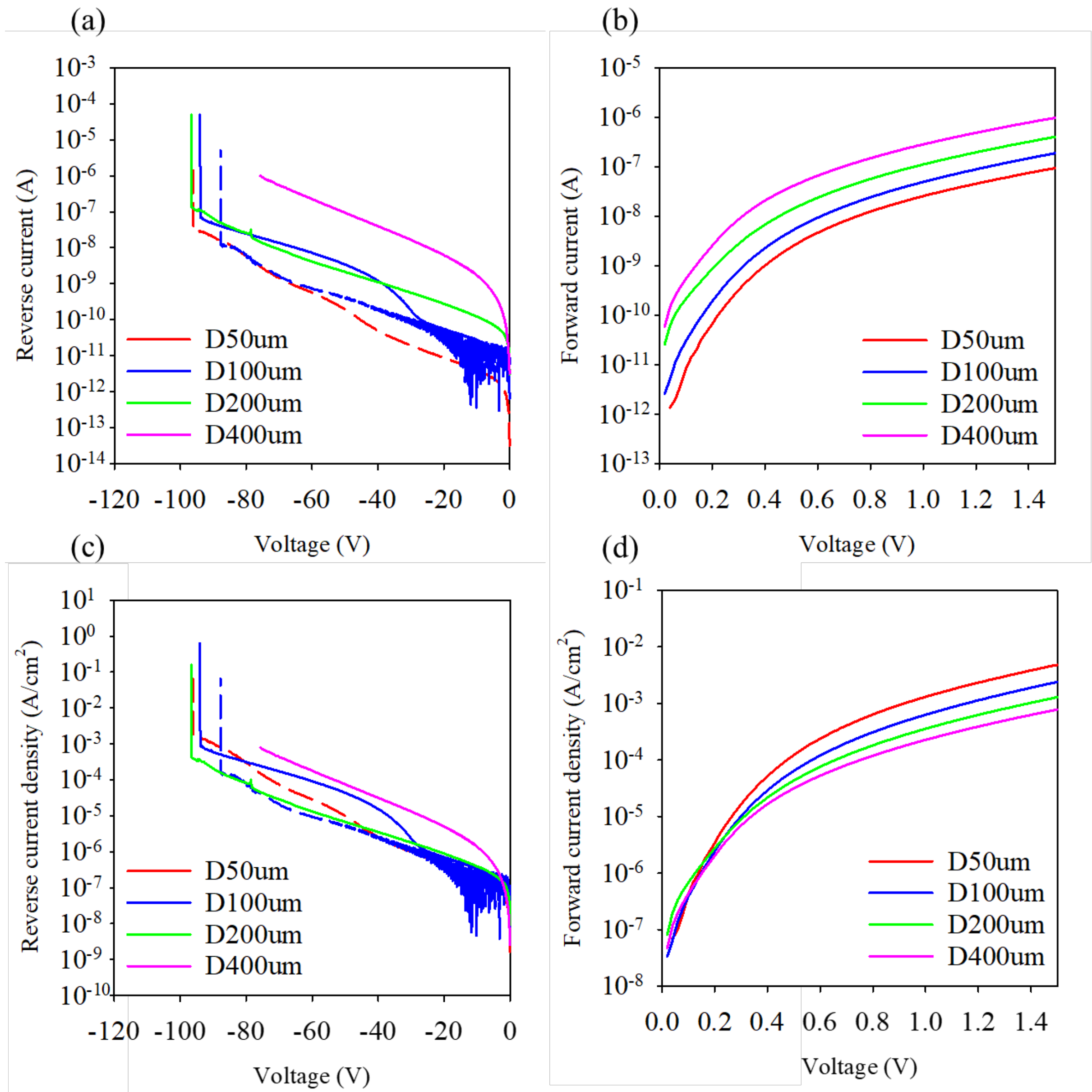


Figure 6.9 (a) Reverse dark current, (b) forward dark current, (c) reverse dark current density, (d) forward dark current density of 120QW

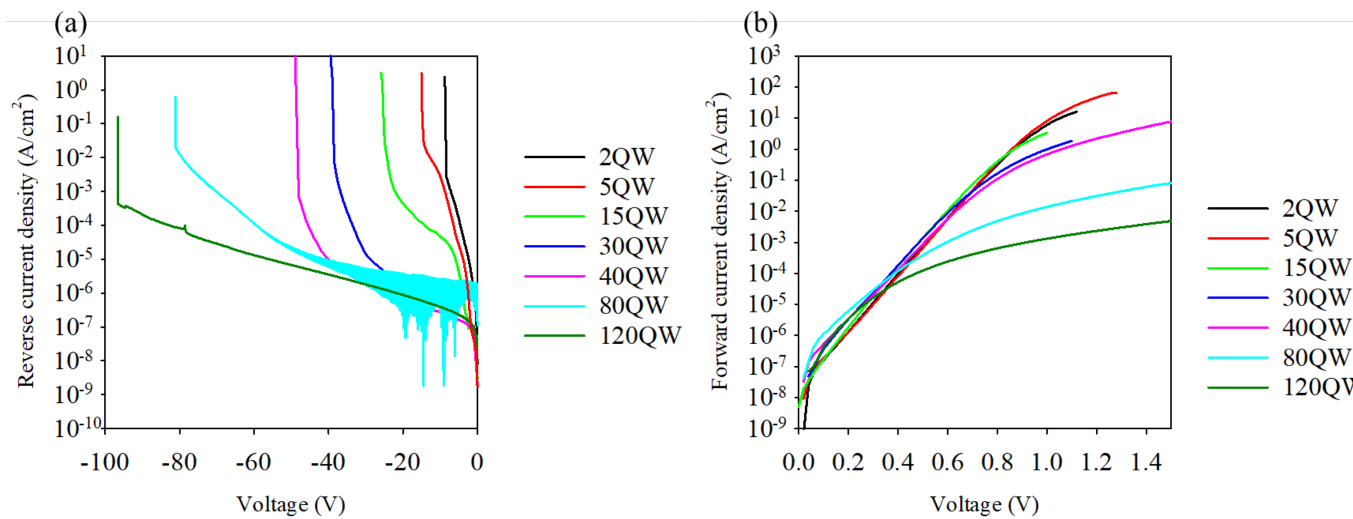


Figure 6.10 (a) Reverse dark current densities and (b) Forward dark current densities of the MQW samples

Table 6.3 Dark current density and ideality factor of each MQW sample

Samples	J0 (A/cm2)	Ideality Factor n
2QW	1.6×10^{-8}	1.85
5QW	1.5×10^{-8}	1.89
15QW	1.8×10^{-8}	1.78
30QW	3.8×10^{-8}	1.9
40QW	3.1×10^{-8}	1.88
80QW	2.5×10^{-7}	2.4
120QW	1.5×10^{-7}	2.6

This section presents the dark current and current–voltage (I–V) characteristics of the newly processed p-i-n GaAsBi/GaAs MQW photodiodes under both forward and reverse bias. The measurement data are displayed in Figure 6.7(a)–(d) and Figure 6.8(a)–(d), showing forward current, reverse current, and their respective current densities for several device diameters (for instance, 50 µm, 100 µm, 200 µm, and 400 µm).

Forward Bias Characteristics

From Figure 6.7(b) and Figure 6.8(b), one observes that the forward dark current remains bulk-dominated at relatively low voltages (around 0 V to 0.6 V). Beyond this region, the series resistance introduces an additional slope, causing a noticeable deviation from the near-exponential trend typical of diode conduction. As the mesa size increases, contact resistance or minor variations in doping profiles can lead to differences in the absolute forward current level. Nonetheless, all devices exhibit reasonably low forward leakage at small biases, consistent with earlier GaAsBi/GaAs MQW diodes [19].

Reverse Bias and Breakdown

In Figure 6.7(a) and Figure 6.8(a), the reverse current gradually increases with more negative voltage, eventually approaching a pronounced rise—suggestive of breakdown phenomena. Diodes having more MQW periods often show breakdown at higher MQW reverse voltages, reflecting an increased depletion width and altered electric field distribution. However, the reverse current does not strictly scale with device area in all cases, indicating sidewall effects or localized defects that can dominate edge leakage. Such behavior frequently appears in epitaxial structures where strain relaxation or dislocation networks create additional current paths [19][20].

Current Density and Device Comparisons

Normalizing the raw current by the diode area yields the reverse and forward current

densities seen in Figure 6.7(c), (d) and Figure 6.8(c), (d). While the forward current density among different device diameters tends to collapse onto a similar trend once series resistance is taken into account, the reverse current density may deviate if perimeter leakage channels become significant. Devices with fewer MQW periods typically exhibit lower dark current at higher reverse voltages—likely because fewer total layers reduce cumulative strain and defect formation. In contrast, structures with more periods risk a larger dislocation density, which can introduce enhanced leakage under reverse bias [20].

Possible Causes of Variation

Mesa Edge Contributions: Imperfections at the mesa perimeter (e.g., etch damage or insufficient passivation) often heighten reverse leakage.

Series Resistance in Forward Bias: Differences in doping profiles, metallization, or substrate resistivity can shift the voltage range where the series-resistance effect becomes prominent.

Defect-Related Leakage: As the MQW stack increases in thickness, the risk of strain-induced dislocations grows, potentially raising both forward and reverse leakage paths.

Device Geometry: Non-ideal scaling with area frequently arises from local inhomogeneities or from geometric factors like contact overlap and sidewall shape.

Overall, Figure 6.7 and Figure 6.8 demonstrate that, while forward currents remain within typical p-i-n diode limits up to moderate voltages, reverse breakdown behaviors differ significantly depending on the MQW period count and device size. Further analysis—such as temperature-dependent I–V or microstructural examination—would be beneficial to clarify how strain and defect densities influence these diodes' leakage and breakdown characteristics.

6.4.2 CV measurements

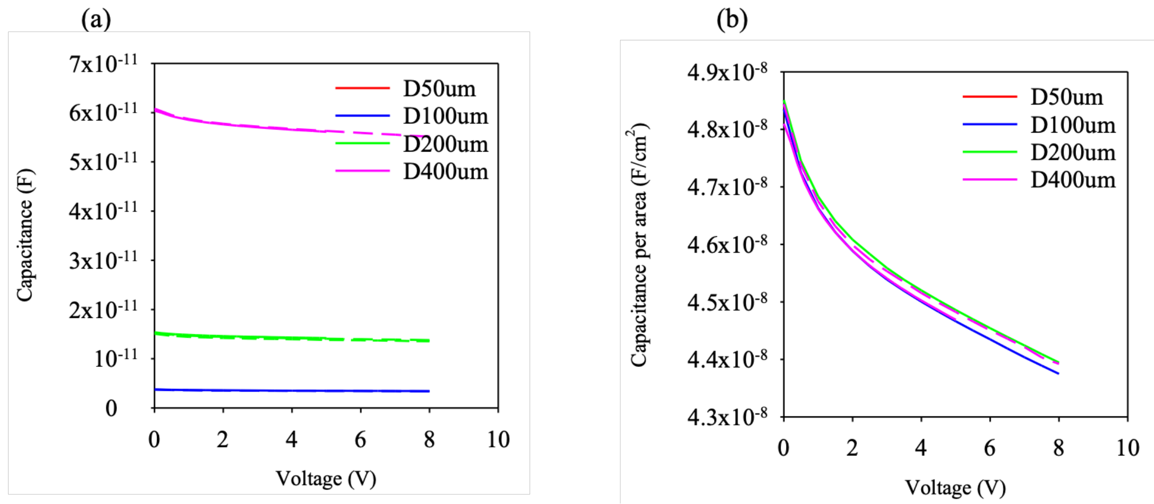


Figure 6.11: (a) C–V and (b) extracted capacitance per unit area of 2QW photodiodes for multiple nominal mesa diameters. Traces with the same colour correspond to the same nominal diameter, while solid and dashed curves represent different nominally identical devices measured at that same diameter.

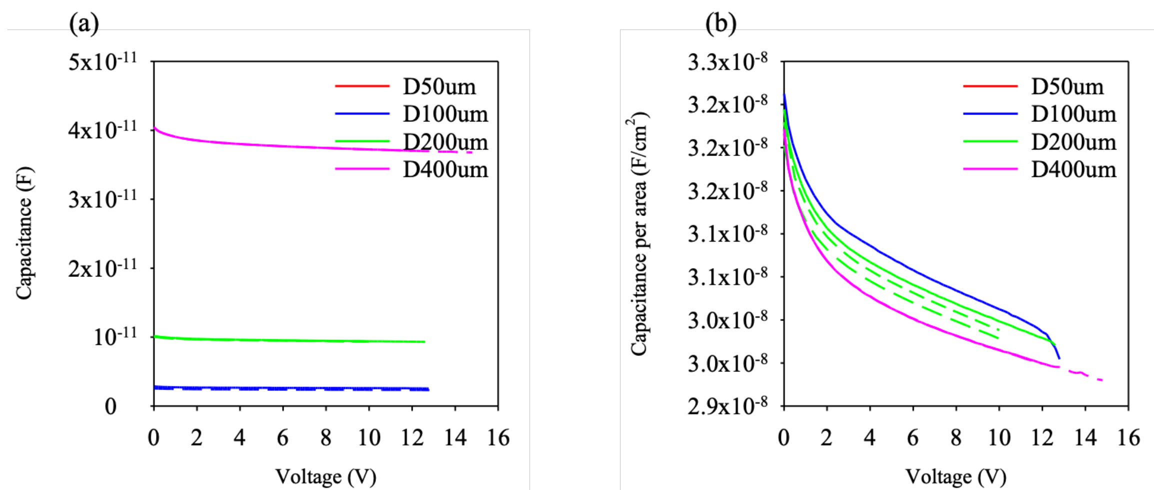
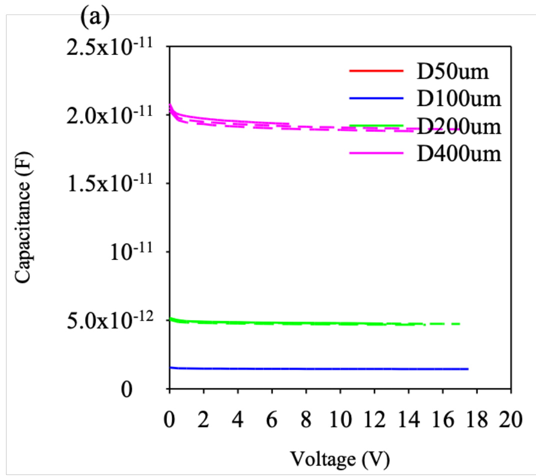


Figure 6.12 (a) CV data of 5QW (b) Capacitance per area of 5QW



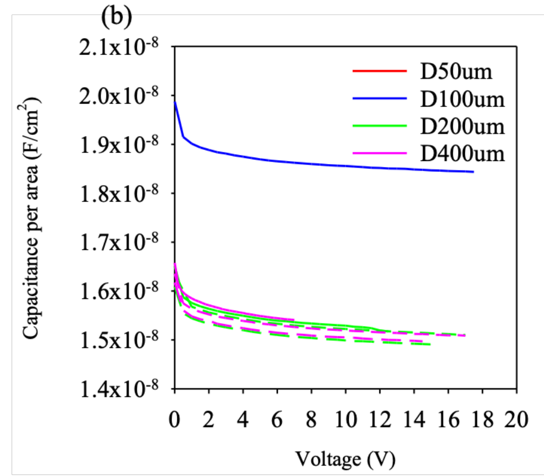
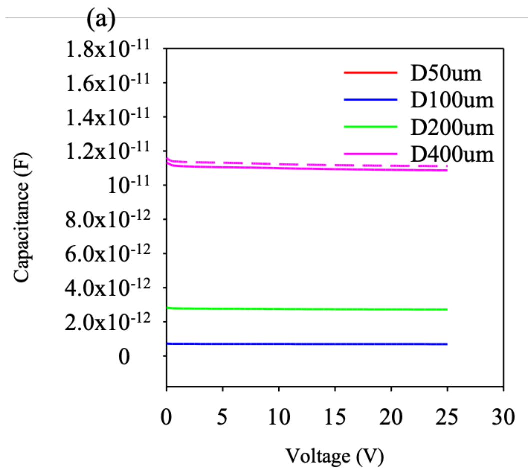


Figure 6.13 (a) CV data of 15QW (b) Capacitance per area of 15QW



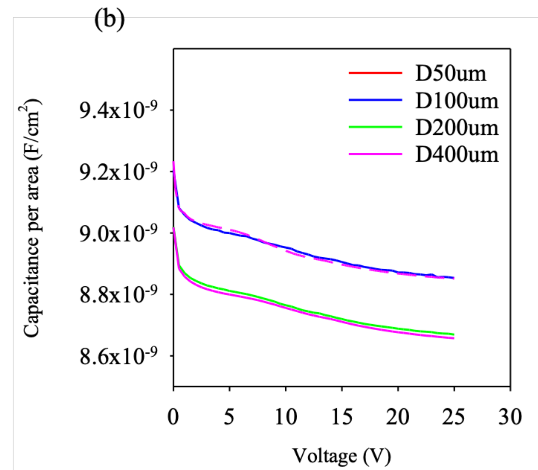


Figure 6.14 (a) CV data of 30QW (b) Capacitance per area of 30QW

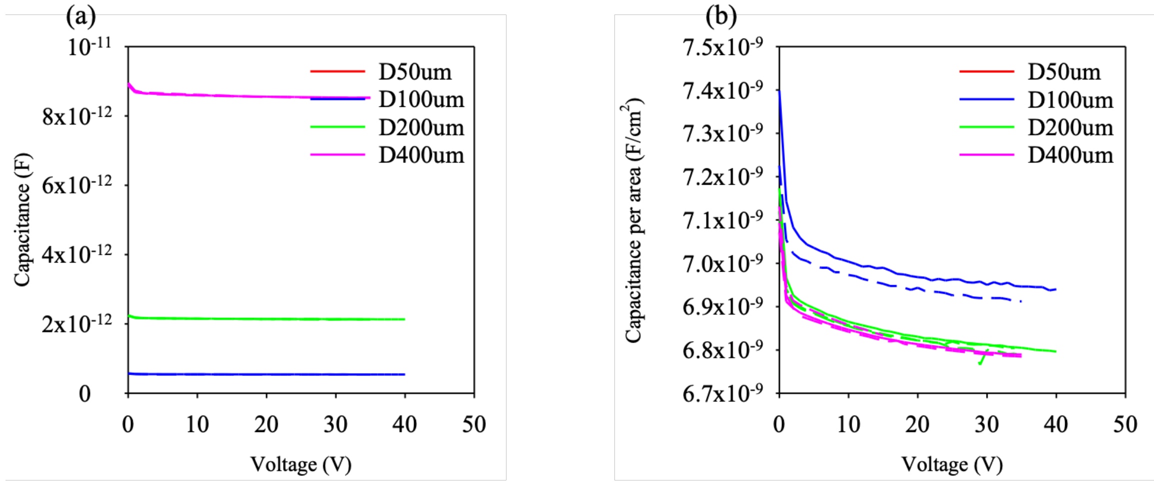


Figure 6.15 (a) CV data of 40QW (b) Capacitance per area of 40QW

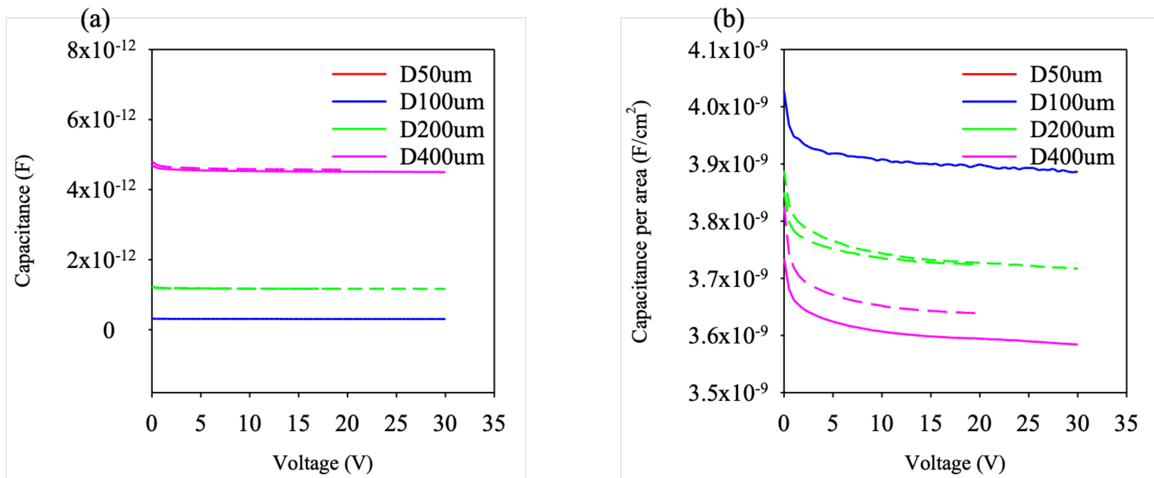


Figure 6.16 (a) CV data of 80QW (b) Capacitance per area of 80QW

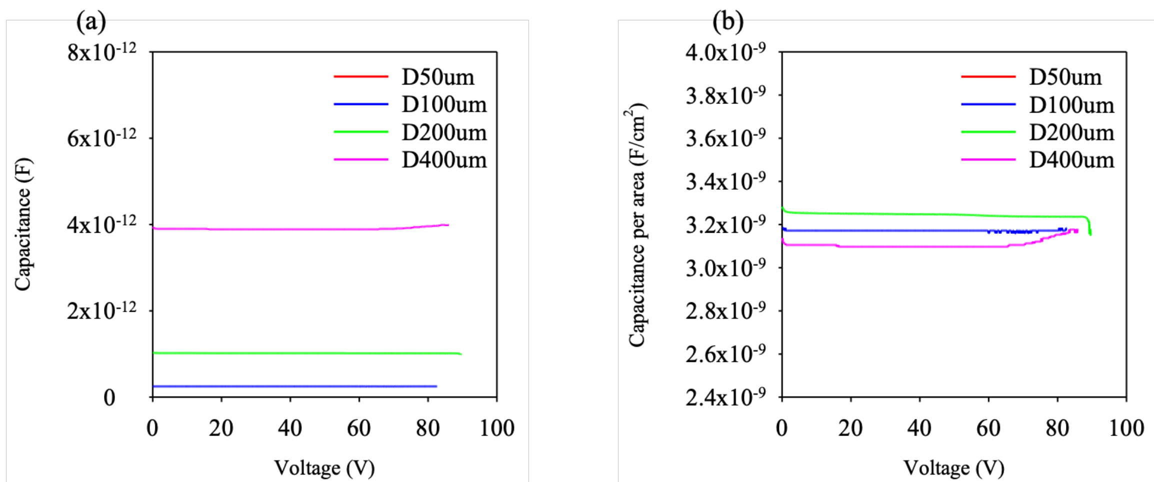


Figure 6.17 (a) CV data of 120QW (b) Capacitance per area of 120QW

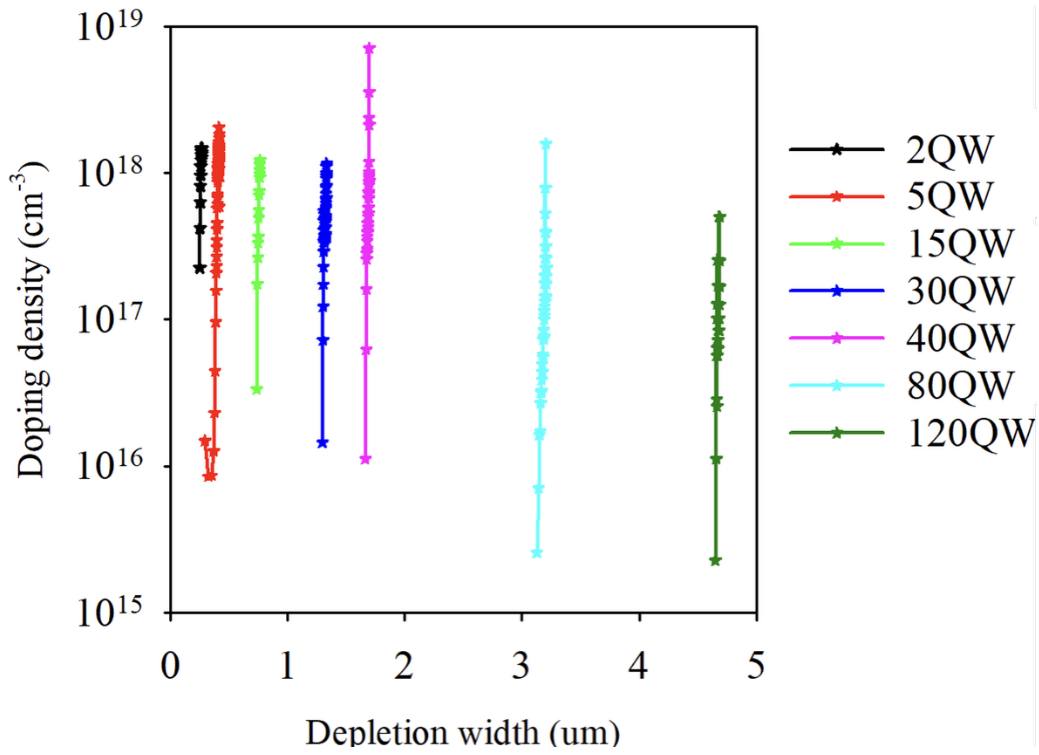


Figure 6.18 Doping profile of MQW samples

Table 6.4 CV model parameters compared with design thickness

Samples	Nominal i-region thickness (nm)	i-region thickness from CV (nm)	Error (%)
2QW	255	258	1.2
5QW	375	410	9.3
15QW	775	760	-1.9
30QW	1375	1254	-8.8
40QW	1775	1692	-4.7
80QW	3375	3241	-4.0
120QW	4975	4683	-5.9

This section describes the capacitance–voltage (C–V) behavior of the devices under various reverse biases, as illustrated by Figure 6.11(a)–(b) through Figure 6.17(a)–(b). Each figure compares the raw capacitance [Figs. (a)] and area-normalized capacitance [Figs. (b)] for different diode diameters (such as 50 µm, 100 µm, 200 µm, and 400 µm). These measurements help elucidate how the depletion region evolves with bias and whether the device capacitance scales predictably with area.

Capacitance vs. Voltage and Area Dependence

2QW and 5QW Samples

In Figure 6.11(a)–(b) (2QW) and Figure 6.12(a)–(b) (5QW), the measured capacitance remains nearly constant at low voltages, then gradually decreases as the reverse bias increases. When normalized by device area, most curves converge into a narrow band, indicating that the bulk depletion region expands similarly in each diode. However, minor deviations appear among the largest and smallest mesa diameters, suggesting that edge effects or doping non-uniformities can impact the total capacitance [19].

15QW and 30QW Samples

Figure 6.13(a)–(b) (15QW) and Figure 6.14(a)–(b) (30QW) show a more pronounced voltage dependence, especially for smaller-diameter devices. As the reverse bias goes beyond 10–15 V, the capacitance stabilizes at slightly lower levels, implying a near-complete depletion of the intrinsic region. Differences in the area-normalized plots underscore that certain devices might have partial doping fluctuations or sidewall contributions, leading to slight mismatches in the absolute capacitance between different diameters [19][20].

40QW and 80QW Samples

Moving to Figure 6.15(a)–(b) (40QW) and Figure 6.16(a)–(b) (80QW), the depletion appears to extend into a thicker region, requiring higher voltages (over 30 V) before the capacitance saturates. The area-normalized curves reveal that while most diodes show a consistent downward trend, the largest device diameters can exhibit marginally lower

capacitance-per-area, possibly due to less-pervasive fringing fields or a more effective depletion profile at the device center compared with the edges.

120QW Sample

Figure 6.17(a)–(b) (120QW) presents data up to ~100 V. In this high-period structure, a larger depletion width is expected, and the measured capacitance remains small across a wide voltage span. Some devices experience only a minor change in area-normalized capacitance with increasing bias, indicating that the i-region is nearly or fully depleted much earlier than the maximum tested voltage. Any residual differences can emerge from subtle doping gradients or from incomplete depletion of certain quantum wells if the doping or conduction band offsets vary locally.

Model Parameters and Thickness Estimation

Table 6.4 compares the nominal i-region thickness to values extracted from C–V fitting. The results show a small to moderate deviation (on the order of a few percent) between measured and designed thicknesses for most samples. For instance, 2QW and 5QW deviate by around 1–9%, while 15QW and 30QW show negative deviations of roughly 2–9%. This spread points to variations in actual doping profile, interface abruptness, or possible strain relaxation in the thicker structures. Nonetheless, all measured thicknesses remain sufficiently close to design targets for reliable device operation [19].

Doping Profile

In Figure 6.18, the doping density extracted from the C–V data is plotted against depletion width for selected devices. While the doping levels broadly align with the nominal doping in the buffer or cap regions, slight fluctuations arise—particularly in thicker MQW samples, which are more prone to strain-induced defects or localized doping anomalies. These inhomogeneities can manifest as local variations in the C–V curves and produce the small mismatches in area-scaling observed in earlier figures

[20].

Summary

Overall, the C–V data in Figure 6.11 through Figure 6.17 demonstrate that as the number of quantum well periods increases, the device tends to exhibit lower capacitance at any given bias, reflecting the thicker depletion region. Normalization by device area indicates that most diodes follow a broadly consistent trend, though certain mesa sizes deviate at higher bias due to edge effects or doping non-uniformities. Fitting the C–V curves yields i-region thicknesses reasonably close to their design values. In addition, doping profiles derived from these measurements help identify minor concentration gradients or interface changes, especially in high-period devices. These findings lay the groundwork for subsequent avalanche and photocurrent analyses, where a well-characterized depletion region is crucial for interpreting multiplication behaviors.

6.5 Photo spectrum of MQW Diodes

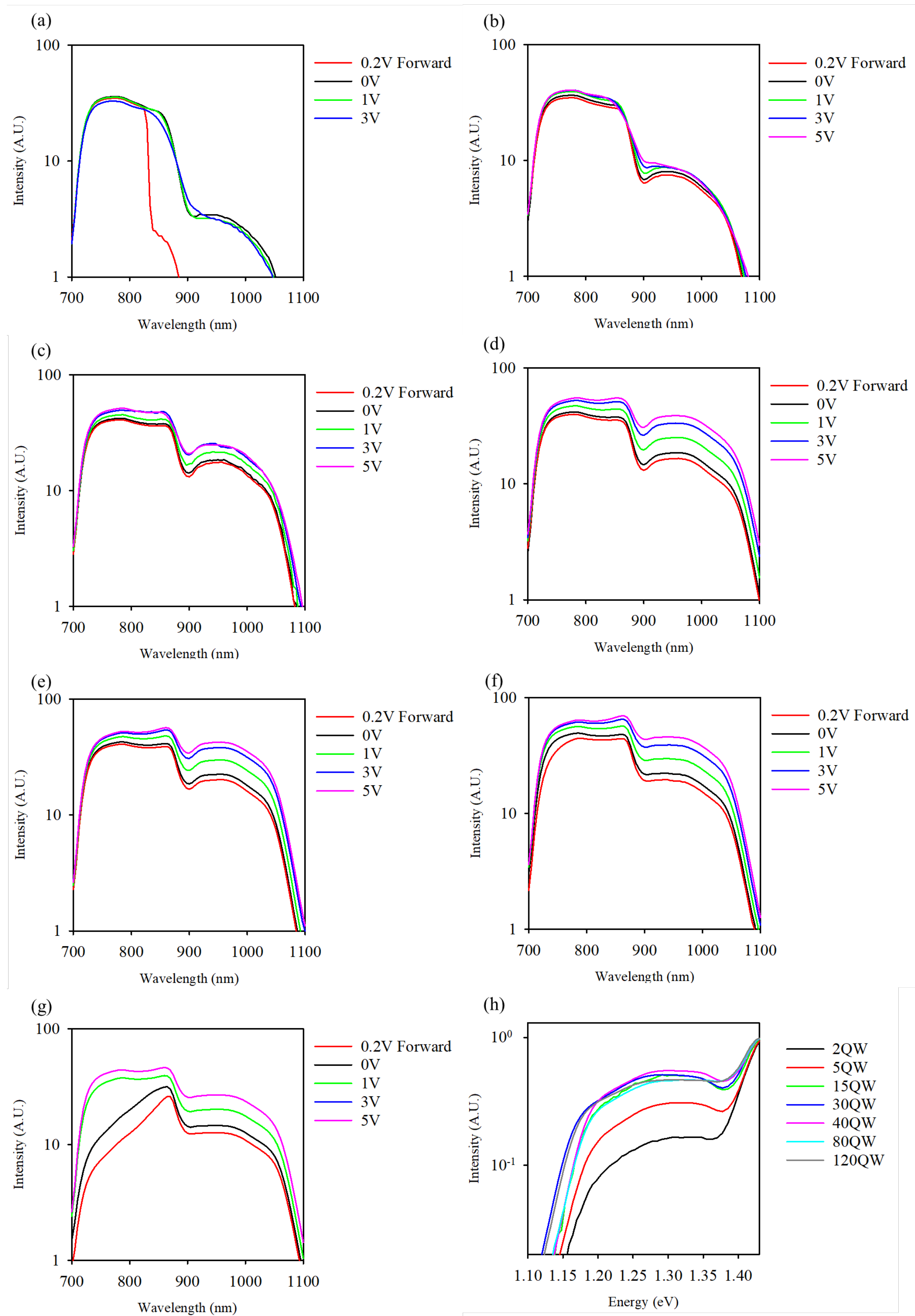


Figure 6.19 (a)-(g) Bias dependent photocurrent intensity of MQW devices; (h) Photocurrent intensity of MQW devices (@0V)

This section presents the photocurrent spectra obtained from a set of diodes under various applied voltages in the 700–1100 nm range, as depicted in Figure 6.19(a)–(g). Each subfigure shows how the optical response changes with bias from a slight forward level (0.2 V) to moderate reverse voltages (1–5 V). The devices include multiple quantum well (MQW) structures with different numbers of periods, allowing for a direct comparison of how the well count affects spectral absorption.

Bias-Dependent Spectra

(a) 2QW

In Figure 6.19(a), the 2QW device displays a main absorption feature near 800–900 nm, corresponding to the GaAs band edge, and a weaker tail around 1000 nm. As the bias changes from 0.2 V forward to 1–5 V reverse, the intensity at wavelengths above 900 nm increases. This indicates that the extended depletion region under higher electric fields improves carrier collection from the MQW layers [19].

(b) 5QW

Figure 6.19(b) illustrates the spectra of the 5QW sample over the 700–1100 nm range. The photocurrent in the 900–1000 nm region rises as the bias moves to higher reverse levels, suggesting that slightly more extended quantum wells or multiple wells can enhance the near-infrared response.

(c) 15QW

In Figure 6.19(c), the 15QW sample shows a clear boost in signal at 1 V, 3 V, and 5 V relative to 0 V, particularly beyond 900 nm. This improvement highlights the impact of deeper depletion on the optical transitions within the GaAsBi layer, which can extend absorption to around 1050 nm [20].

(d) 30QW

Figure 6.19(d) corresponds to the 30QW device, where the absorption edge shifts slightly further into the infrared regime compared to devices with fewer wells. At 5 V reverse bias, the photocurrent from 1000 nm to 1100 nm grows noticeably, indicating that increasing the Bi-doped quantum well stack effectively narrows the bandgap region and broadens the IR absorption window.

(e) 40QW

Figure 6.19(e) (40QW) shows relatively high baseline photocurrent even at 0.2 V forward bias, reflecting an enhanced absorption around 800–900 nm due to the increased number of wells. As the bias is shifted to 1–5 V in reverse, the photocurrent amplitude continues to rise, underscoring how the extended depletion region aids carrier extraction [19].

(f) 80QW

In Figure 6.19(f), the 80QW device exhibits a prominent GaAs-related peak near 800 nm, while maintaining stronger, flatter absorption above 900 nm. This behavior suggests that with higher well counts, the effective band structure is more strongly influenced by Bi doping and that a larger depleted volume also enhances the collection of carriers generated at longer wavelengths [20].

(g) 120QW

Finally, Figure 6.19(g) for the 120QW sample reveals a pronounced rise in photocurrent under 5 V reverse bias, extending up to around 1100 nm. The thick MQW i-region benefits from a more substantial carrier collection path, potentially pushing the cutoff slightly beyond 1.1 μm. Although a higher number of periods often leads to increased strain or defects (and thus higher dark current), the spectral data presented here demonstrate effective absorption from visible to near-infrared wavelengths.

Comparison Across Multiple Devices

Figure 6.19(h) contrasts the spectral responses of 2QW, 5QW, 15QW, 30QW, 40QW, 80QW, and 120QW devices at 0 V, plotted against energy. As the well count increases, the onset of absorption shifts further into lower energies (longer wavelengths), indicating a stronger bandgap reduction. Differences in the spectral magnitude near 1.1 eV highlight how the fixed Bi content, combined with varying well counts, affects local strain relaxation and effective band-edge tuning.

In conclusion, these photo-spectrum measurements illustrate how devices ranging from 2QW to 120QW respond to light between 700 nm and 1100 nm. With increasing numbers of wells and higher reverse biases, quantum well effects become more pronounced, strengthening absorption at longer wavelengths beyond 1 μm. Designers must still balance any strain-induced defects that arise with higher well counts, as these may negatively impact dark current levels. Future work involving more detailed structural analysis and avalanche-gain characterization could provide further insight into the trade-offs between extended infrared absorption and potential device reliability issues.

6.6 Photomultiplication characterisation of MQW Diodes

To distinguish electron-initiated and mixed-carrier avalanche multiplication in a single p-i-n polarity, wavelength-selective optical excitation was employed to control the carrier-generation profile. For short-wavelength illumination (e.g., 455 nm in this work), absorption in the GaAs cap/cladding is strong and photogeneration is confined close to the illuminated surface; under reverse bias the resulting electrons drift across the depleted multiplication region while the corresponding holes are collected locally, thereby approximating predominantly electron injection. The multiplication measured under this condition is denoted M_e. In contrast, near-infrared excitation at wavelengths below the GaAs bandgap (e.g., 940-980 nm) penetrates the GaAs layers and is absorbed predominantly in the GaAsBi quantum wells, i.e., within the depleted MQW region. In this case photocarriers are generated directly in the high-field region and both electrons and holes can contribute to impact ionisation, producing a mixed-injection multiplication factor, M_{mix}. This wavelength-selective injection approach is widely used in APD characterisation when separate p-i-n and n-i-p test structures are not available [30].

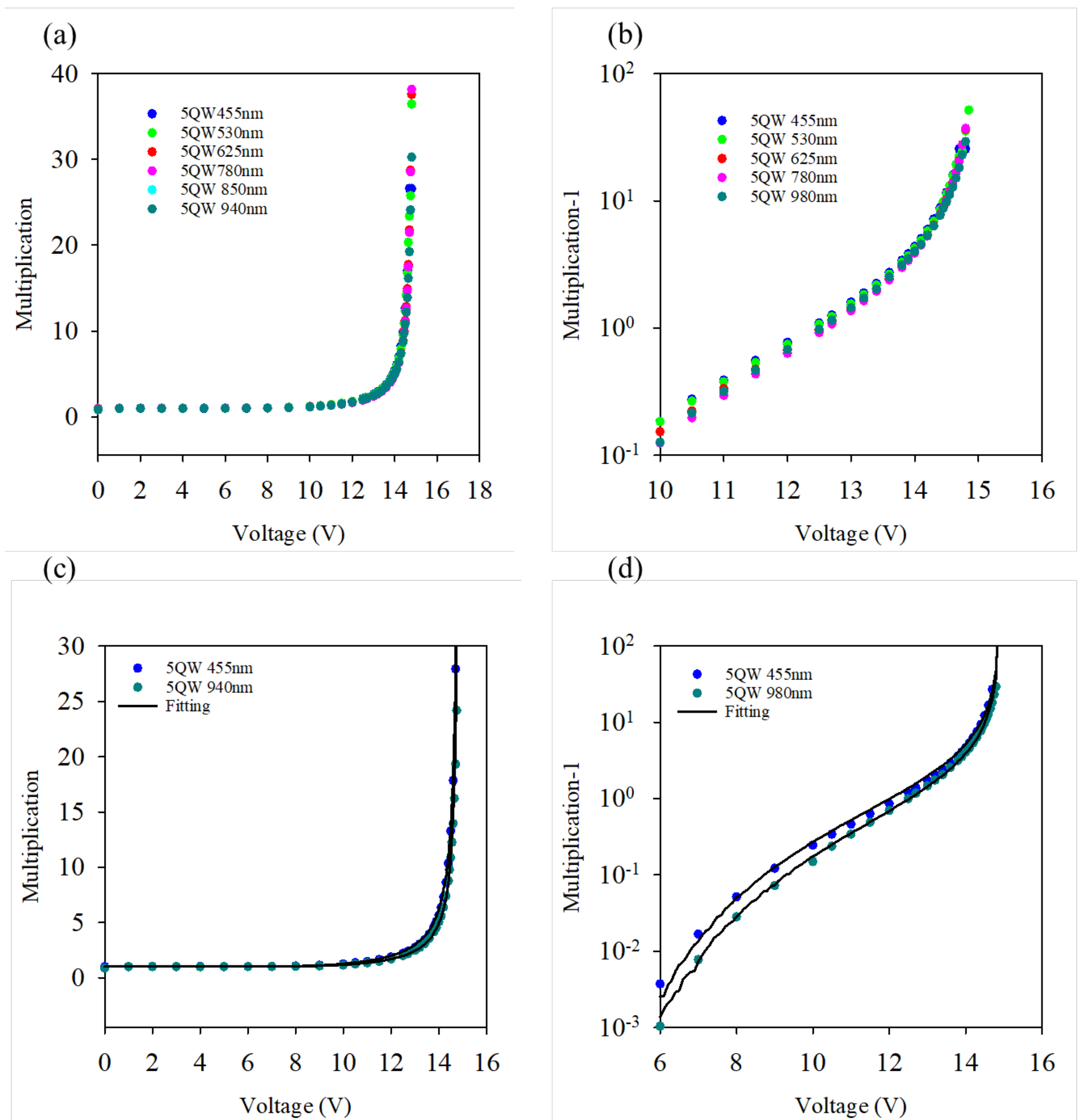


Figure 6.20 (a) M_e (blue dots), M_{530} (green dots), M_{625} (red dots), M_{780} (cyan dots), M_{850} (purple dots), M_{mix} (dark green dots) of 5QW at different voltage; (b) M_e - 1(blue dots), M_{530} - 1(green dots) , M_{625} - 1(red dots), M_{780} - 1(cyan dots), M_{mix} - 1(dark green dots) of 5QW; (c) M_e (blue dots), M_{mix} (dark green dots) of 5QW with RPL fitting (solid lines) at different voltage; (d) M_e - 1(blue dots), M_{mix} - 1(dark green dots) of 5QW with RPL fitting (solid lines) at different voltage.

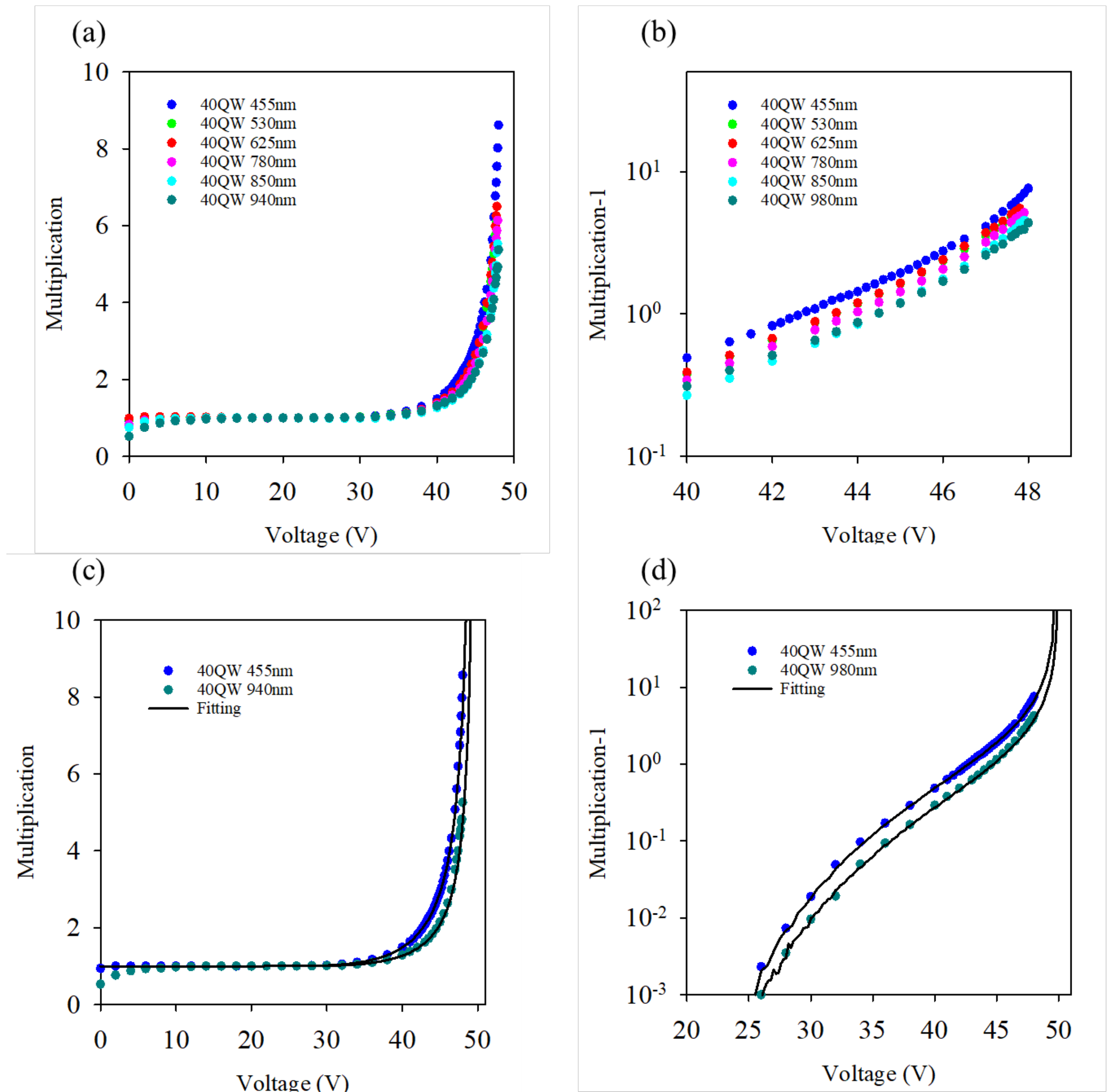


Figure 6.21 (a) M_e (blue dots), M_{530} (green dots), M_{625} (red dots), M_{780} (cyan dots), M_{850} (purple dots), M_{mix} (dark green dots) of 40QW at different voltage; (b) M_e - 1(blue dots), M_{530} - 1(green dots) , M_{625} - 1(red dots), M_{780} - 1(cyan dots), M_{850} - 1(purple dots), M_{mix} - 1(dark green dots) of 40QW; (c) M_e (blue dots), M_{mix} (dark green dots) of 40QW with RPL fitting (solid lines) at different voltage; (d) M_e - 1(blue dots), M_{mix} - 1(dark green dots) of 40QW with RPL fitting (solid lines) at different voltage.

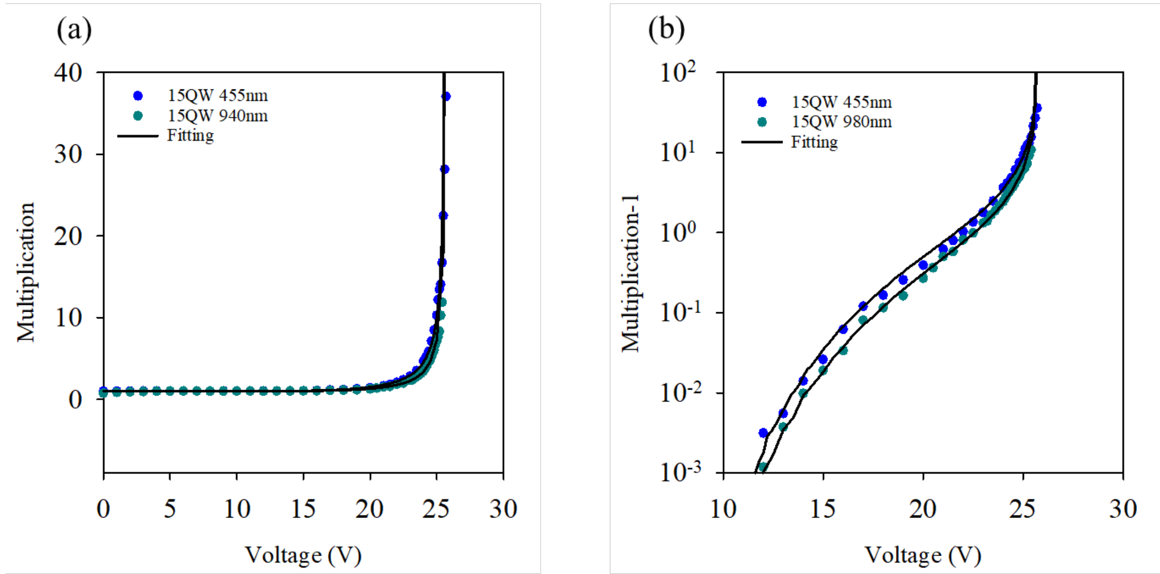


Figure 6.22 (a) M_e (blue dots), M_{mix} (dark green dots) of 15QW with RPL fitting (solid lines) at different voltage; (b) M_e - 1(blue dots), M_{mix} - 1(dark green dots) of 15QW with RPL fitting (solid lines) at different voltage.

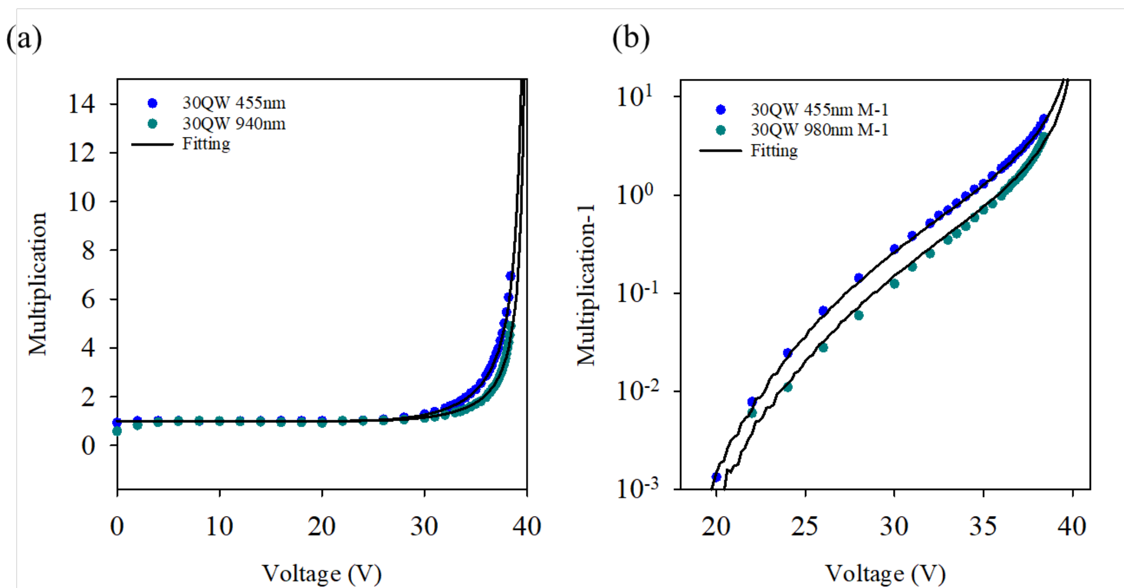


Figure 6.23 (a) M_e (blue dots), M_{mix} (dark green dots) of 30QW with RPL fitting (solid lines) at different voltage; (b) M_e - 1(blue dots), M_{mix} - 1(dark green dots) of 30QW with RPL fitting (solid lines) at different voltage.

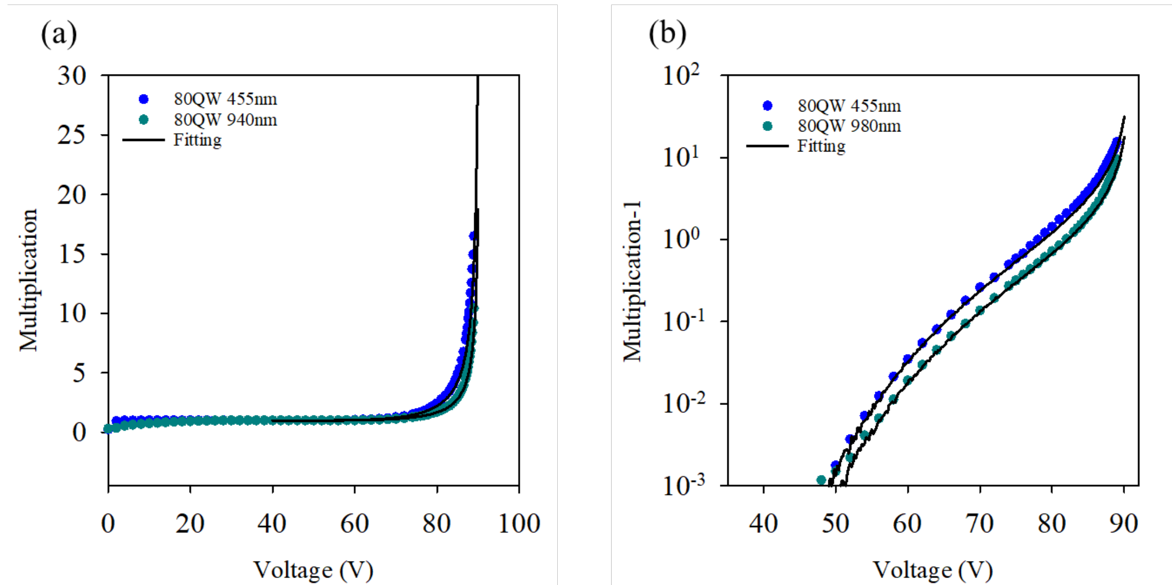


Figure 6.24 (a) M_e (blue dots), M_{mix} (dark green dots) of 80QW with RPL fitting (solid lines) at different voltage; (b) M_e - 1(blue dots), M_{mix} - 1(dark green dots) of 80QW with RPL fitting (solid lines) at different voltage.

M_e denotes multiplication extracted under short-wavelength excitation (approximating electron injection from the illuminated side), whereas M_{mix} (e.g., under long-wavelength excitation) corresponds to mixed electron-hole injection when absorption occurs predominantly within the depleted MQW region.

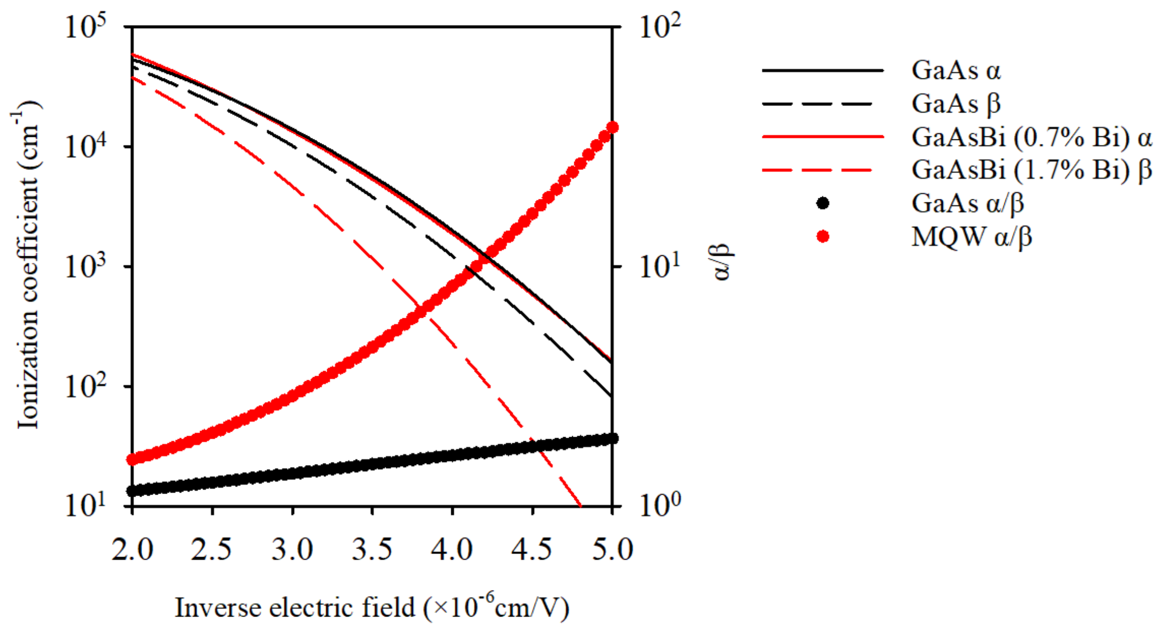


Figure 6.25 α of GaAs (black line), β of GaAs (black dash line), α of GaAsBi with 0.7% Bi(red line), β of GaAsBi with 1.7% Bi (red dash line), α/ β ratio of GaAs (black dots), α/ β ratio of MQW(red dots), vs inverse electric field

Table 6.5 Average period Bi concentration of MQW diodes by XRD

Well number	i-region thickness(nm)	Average period Bi%
2	258	1.058462
5	385	0.961165
15	745	0.957231
30	1250	0.996667
40	1605	0.953055
80	3162	0.928125
120	4062	0.951923

6.6 Photomultiplication Characterisation of MQW Diodes

This section summarizes the internal gain (avalanche multiplication) measurements performed on p-i-n photodiodes containing multiple quantum wells (MQWs). The corresponding data are presented in Figures 6.20–6.24. Compared to traditional single-layer GaAsBi, these MQW devices—of identical total thickness—were designed to selectively tailor the impact-ionization coefficients of electrons and holes. The black solid lines in the figures represent Random Path Length (RPL) simulations based on bulk GaAsBi with different Bi compositions (0.7% or 1.7%) for reference [19][20].

6.6.1 Setup Summary

Photomultiplication measurements were carried out under laser illumination at multiple wavelengths in order to control the spatial profile of photogeneration relative to the high-field region. The reverse bias was increased incrementally from low field to the onset of avalanche and beyond, and the multiplication factor was evaluated using the gain definition introduced in Chapter 4 [Eq. (4.11)], where the low-bias photocurrent (well below avalanche onset) provides the primary current reference, and the photocurrent at higher reverse bias corresponds to the multiplied current.

The use of different laser wavelengths is central to separating near-electron-initiated multiplication from mixed-carrier multiplication. For illumination from the p-side of a p-i-n mesa, short-wavelength excitation (e.g., 532 nm) is absorbed within a shallow depth near the surface, such that most electron–hole pairs are generated within the upper p+ cap/cladding rather than throughout the depleted i-region. Under this condition, the minority electrons drift into the high-field region and initiate impact ionisation, and the measured multiplication is denoted as M_e. In contrast, for longer-wavelength excitation (e.g., 980 nm) with a much larger absorption length, a substantial fraction of carriers is generated within the depleted region, so that both electrons and holes can contribute to the avalanche process. The multiplication extracted in this regime is therefore denoted as M_{mix}, reflecting mixed injection. Intermediate wavelengths (e.g., 530 nm, 625 nm, 780 nm, and 850 nm) provide additional data points that trace the transition between these limiting cases; however, the subsequent discussion and RPL comparison focus primarily on M_e and M_{mix} as representative near-electron-initiated and mixed-carrier conditions, respectively.

To benchmark the MQW impact-ionisation behaviour against bulk GaAsBi, the black

solid curves in Figs. 6.20–6.24 show Random Path Length (RPL) simulations using bulk GaAsBi ionisation coefficients reported for different Bi compositions [19,20]. In these comparisons, M_e (and M_e^{-1}) is used as the most sensitive diagnostic of the electron ionisation behaviour, while M_{mix} (and M_{mix}^{-1}) captures the onset of mixed-carrier multiplication and thus provides practical sensitivity to the hole ionisation contribution.

6.6.2 Photomultiplication Performance of Different MQW Periods

Figures 6.18-6.21 show the wavelength-dependent multiplication for representative MQW devices (5QW and 40QW). As the reverse bias increases into the high-field regime, the multiplication rises rapidly, with the onset voltage and achievable gain depending on the MQW period number and the associated electric-field distribution. Importantly, the measured multiplication depends on excitation wavelength because the optical absorption profile controls whether carriers are injected predominantly as electrons from the illuminated side (M_e) or generated within the depleted MQW region with mixed electron-hole injection (M_{mix}). For the shortest-wavelength excitation (e.g., 455nm), photogeneration is confined near the surface and the measured multiplication is therefore largely electron-initiated; the resulting M_e trends can be compared directly with the electron-ionisation-dominated RPL reference simulation. For sub-bandgap near-infrared excitation (940-980nm), photons penetrate the GaAs layers and are absorbed mainly in the GaAsBi wells, so carriers are generated within the high-field region and both carrier types can contribute to avalanche initiation, yielding M_{mix}.

Figures 6.22-6.24 further show that higher-period MQW devices (e.g., 15QW, 30QW, and 80QW) typically require larger reverse biases to reach tens-to-hundreds multiplication. Across these devices, M_{mix} increasingly exceeds M_e at high bias, consistent with a non-negligible hole contribution under mixed injection. When benchmarked against the bulk GaAsBi reference simulations, the overall trends support the conclusion that MQW band-structure engineering can preserve strong electron-initiated multiplication while suppressing hole-initiated ionisation relative to an equivalent-thickness single-layer design, thereby lowering the effective $k = \beta/\alpha$ relevant to excess-noise performance [19,20].

6.6.3 Influence of Average Period Bi Content on Impact Ionization

Based on Table 6.5, each MQW sample exhibits an average period Bi content of about 1%. From the best-fit results and curve comparisons, one finds:

1. Despite having only ~1% average Bi in the MQW region, the electron impact-ionization coefficient (α) still matches that of ~0.7% Bi bulk GaAsBi. This indicates that electron-driven avalanche is preserved effectively.
2. In parallel, the hole ionization coefficient (β) aligns well with the 1.7% Bi bulk GaAsBi level, implying that the MQW band-structure engineering significantly curbs hole-initiated avalanche.
3. Consequently, these devices exhibit an unusual combination where electron ionization behaves like ~0.7% Bi and hole ionization like ~1.7% Bi, thereby offering higher gain potential with reduced noise.

Overall, the experimental data in Figures 6.20–6.24 and the RPL fits indicate that:

For MQW samples with an average period Bi fraction of ~1%, electron impact ionization closely matches bulk GaAsBi with ~0.7% Bi, while the hole ionization coefficient behaves similarly to ~1.7% Bi.

This asymmetric band-structure adjustment appreciably lowers hole avalanche, sustaining a manageable level of dark current.

By distributing Bi into multiple thin quantum wells rather than a single thick layer, the design leverages valence-band modifications to achieve suppressed hole ionization. With further stress and defect control, such structures could be promising for near-infrared APDs requiring high sensitivity and low noise.

6.7 Discussion

Throughout this chapter, we have systematically examined the performance of p-i-n multiple quantum well (MQW) photodiodes, encompassing their dark current (I–V), capacitance–voltage (C–V), spectral response, and avalanche multiplication characteristics. We have specifically compared devices with different MQW periods to assess how increasing the number of wells affects both impact ionisation suppression and overall device quality. Below we summarize the key findings and potential challenges.

6.7.1 Quantum Well Design and Hole Suppression

By embedding multiple GaAsBi-containing layers into a GaAs p-i-n structure, the devices can effectively maintain a relatively high electron-initiated gain (α) while substantially reducing the hole impact ionisation coefficient (β), thereby decreasing β/α (i.e., k) [19][20]. Our results indicate that even with an average Bi content of roughly 1%, the extended MQW design significantly curtails hole-triggered avalanche at higher bias voltages. Notably, this design strategy enables improved device performance at a given Bi composition—or, conversely, meeting a similar target performance with a lower Bi fraction—thus easing the semiconductor growth constraints on epitaxial quality and strain management.

6.7.2 Strain Management and Device Uniformity

As the MQW period count increases (e.g., beyond 40QW or up to 120QW), the accumulated strain poses a growing risk of defect formation, which can manifest as elevated dark current or localized breakdown paths. Although more periods can indeed boost breakdown voltage and potentially reduce noise under high avalanche gain, optimizing the growth process and managing strain remain crucial. Techniques such as segmental epitaxial growth or strain-compensation layers in the barriers may help mitigate defects and maintain more uniform device performance.

6.7.3 Overall Performance and Application Prospects

Dark Current and Breakdown Voltage

The data presented here show that more MQW periods generally correlate with higher breakdown voltages yet also tend to raise dark current. A balance must be struck in practical designs to achieve sufficiently high gain (low β/α) while controlling dislocation densities and leakage.

Future Optimization

Nonlocal Impact Ionisation Models: Given the relatively thin MQW avalanche region and the band offsets at QW barriers, incorporating dead-space effects and wavefunction coupling may yield more accurate modelling.

Temperature Studies: Conducting temperature-dependent tests would give a broader view of device stability and performance margins.

6.7.4 Conclusion

In summary, the results presented in this chapter confirm that GaAsBi/GaAs MQW structures can successfully suppress hole impact ionisation while maintaining robust electron gain, thus achieving relatively low noise at high avalanche multiplications. At the same time, the accumulation of strain and potential defect formation underscore the importance of refined epitaxial techniques. Looking ahead, improved stress compensation and nonlocal modelling approaches could unlock even greater potential for these multi-period MQW diodes in NIR high-sensitivity APDs.

References

[1] G. P. Agrawal, *Fiber - optic communication systems*, Jun. 2021. doi:10.1002/9781119737391

[2] R. J. McIntyre, “Multiplication noise in uniform avalanche diodes,” *IEEE Transactions on Electron Devices*, vol. ED-13, no. 1, pp. 164–168, Jan. 1966. doi:10.1109/t-ed.1966.15851

[3] J. Beck et al., “The HgCdTe Electron Avalanche photodiode,” *Journal of Electronic Materials*, vol. 35, no. 6, pp. 1166–1173, Jun. 2006. doi:10.1007/s11664-006-0237-3

[4] A. R. Marshall, J. P. David, and C. H. Tan, “Impact ionization in InAs electron avalanche photodiodes,” *IEEE Transactions on Electron Devices*, vol. 57, no. 10, pp. 2631–2638, Oct. 2010. doi:10.1109/ted.2010.2058330

[5] “Si APD S16453-10K: Hamamatsu Photonics,” Si APD S16453-10K | Hamamatsu Photonics, https://www.hamamatsu.com/jp/en/product/optical-sensors/apd/si-apd/S16453-10K.html (accessed Mar. 23, 2024).

[6] X. Jin et al., “Very low excess noise Very Low Excess Noise Al0.75Ga0.25As0.56Sb0.44 Avalanche Photodiode,” *Optics Express*, vol. 31, no. 20, p. 33141, Sep. 2023. doi:10.1364/oe.500169

[7] S. Lee et al., “High Gain, Low Noise, Room Temperature 1550 nm GaAsSb/AlGaAsSb Avalanche Photodiodes,” *Optica*, vol. 10, no. 2, p. 147, Jan. 2023. doi:10.1364/optica.476963

[8] F. Capasso, Won-Tien Tsang, and G. F. Williams, “Staircase solid-state photomultipliers and avalanche photodiodes with enhanced ionization rates ratio,” *IEEE Transactions on Electron Devices*, vol. 30, no. 4, pp. 381–390, Apr. 1983. doi:10.1109/t-ed.1983.21132

[9] M. Ren et al., “AlInAsSb/GaSb staircase avalanche photodiode,” *Applied Physics Letters*, vol. 108, no. 8, Feb. 2016. doi:10.1063/1.4942370

[10] Y. Ma et al., “Enhanced carrier multiplication in InAs quantum dots for bulk avalanche photodetector applications,” *Advanced Optical Materials*, vol. 5, no. 9, Mar. 2017. doi:10.1002/adom.201601023

[11] A. C. Farrell et al., “Plasmonic field confinement for separate absorption-multiplication in InGaAs nanopillar avalanche photodiodes,” *Scientific Reports*, vol. 5, no. 1, Dec. 2015. doi:10.1038/srep17580

[12] Y. Liu et al., “Valence band engineering of GaAsBi for low noise avalanche photodiodes,” *Nature Communications*, vol. 12, no. 1, Aug. 2021. doi:10.1038/s41467-021-24966-0

[13] M. Usman, C. A. Broderick, A. Lindsay, and E. P. O’Reilly, “Tight-binding analysis of the electronic structure of dilute bismide alloys of GaP and GaAs,” *Physical Review B*, vol. 84, no. 24, Dec. 2011. doi:10.1103/physrevb.84.245202

[14] I. H. Öguzman, Y. Wang, J. Kolnik, and K. F. Brennan, “Theoretical study of hole initiated impact ionization in bulk silicon and GaAs using a wave-vector-dependent numerical transition rate formulation within an ensemble Monte Carlo Calculation,” *Journal of Applied Physics*, vol. 77, no. 1, pp. 225–232, Jan. 1995. doi:10.1063/1.359374

[15] J. W. Matthews and A. E. Blakeslee, “Defects in epitaxial multilayers: III. Preparation of almost perfect multilayers,” *Journal of Crystal Growth*, vol. 32, no. 2, pp. 265–273, Feb. 1976. doi:10.1016/0022-0248(76)90041-5

[16] R. France, C.-S. Jiang, and A. J. Ptak, “In situ strain relaxation comparison between GaAsBi and GaInAs grown by molecular-beam epitaxy,” *Applied Physics Letters*, vol. 98, no. 10, Mar. 2011. doi:10.1063/1.3562952

[17] R. D. Richards et al., “Photovoltaic characterisation of GaAsBi/GaAs multiple quantum well devices,” *Solar Energy Materials and Solar Cells*, vol. 172, pp. 238–243, Dec. 2017. doi:10.1016/j.solmat.2017.07.029

[18] K. Alberi et al., “Valence band anticrossing in GaBixAs1−x,” *Applied Physics Letters*, vol. 91, no. 5, Jul. 2007. doi:10.1063/1.2768312

[19] R. D. Richards, F. Bastiman, D. Walker, R. Beanland, and J. P. David, “Growth and Structural Characterization of GaAsBi/GaAs Multiple Quantum Wells,” *Semiconductor Science and Technology*, vol. 30, no. 9, p. 094013, Jul. 2015. doi:10.1088/0268-1242/30/9/094013

[20] D. Fan et al., “MBE grown GaAsBi/GaAs double quantum well separate confinement heterostructures,” *Journal of Vacuum Science & Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena*, vol. 31, no. 3, Feb. 2013. doi:10.1116/1.4792518

[21] S. Tixier *et al.*, “Molecular beam epitaxy growth of GaAs1−xBix,” *Appl. Phys. Lett.*, vol. 82, no. 14, pp. 2245–2247, Apr. 2003, doi: 10.1063/1.1565499.

[22] M. Henini *et al.*, “Molecular beam epitaxy of GaBiAs on (311)B GaAs substrates,” *Appl. Phys. Lett.*, vol. 91, no. 25, p. 251909, Dec. 2007, doi: 10.1063/1.2827181.

[23] J. Puustinen *et al.*, “Variation of lattice constant and cluster formation in GaAsBi,” *J. Appl. Phys.*, vol. 114, no. 24, p. 243504, Dec. 2013, doi: 10.1063/1.4851036.

[26] C. Ryang Wie, “High resolution x-ray diffraction characterization of semiconductor structures,” *Mater. Sci. Eng. R Rep.*, vol. 13, no. 1, pp. 1–56, Sep. 1994, doi: 10.1016/0927-796X(94)90008-6.

[27] H. Achour, S. Louhibi, B. Amrani, A. Tebboune, and N. Sekkal, “Structural and electronic properties of GaAsBi,” *Superlattices Microstruct.*, vol. 44, no. 2, pp. 223–229, Aug. 2008, doi: 10.1016/j.spmi.2008.05.004.

[28] D. F. Reyes *et al.*, “Bismuth incorporation and the role of ordering in GaAsBi/GaAs structures,” *Nanoscale Res. Lett.*, vol. 9, no. 1, p. 23, Dec. 2014, doi: 10.1186/1556-276X-9-23.

[29] T. B. O. Rockett *et al.*, “Influence of growth conditions on the structural and opto-electronic quality of GaAsBi,” *J. Cryst. Growth*, vol. 477, pp. 139–143, Nov. 2017, doi: 10.1016/j.jcrysgro.2017.02.004.

[30] J. S. Ng, S. L. Tan, Y. L. Goh, C. H. Tan, J. P. R. David, J. Allam, S. J. Sweeney, and A. R. Adams, “InGaAsN as absorber in APDs for 1.3 micron wavelength applications,” in 22nd International Conference on Indium Phosphide and Related

Materials (ICIPRM), 2010, pp. 187–190, doi: 10.1109/ICIPRM.2010.5516060.

Chapter 7 Conclusion and Future Work

7.1 Summary of Research

This dissertation systematically investigated the epitaxial growth, fundamental mechanisms, and device characteristics of avalanche photodiodes (APDs) based on GaAs/GaAsBi. Over the course of six chapters, both single-layer heterojunction and multiple-quantum-well (MQW) configurations were analyzed and experimentally validated. The main conclusions can be summarized as follows:

Accurate Control of GaAsBi Epitaxy

By incorporating a small amount of Bi into GaAs using molecular beam epitaxy (MBE), it is possible to significantly reduce the bandgap while enhancing the spin–orbit splitting. Bi compositions and thicknesses were established primarily from XRD (supported by optical consistency checks where applicable), while SIMS was used to verify the dopant profiles and the axial placement of the GaAsBi insertion layers, and to provide a consistent measure of relative Bi incorporation under identical SIMS conditions.

Single-Layer Heterojunctions for Avalanche Detectors

The studies in Chapter 5 demonstrated that inserting a 200 nm GaAsBi layer into the i-region of a p–i–n or n–i–p diode can substantially suppress the hole impact ionization coefficient (β). As Bi content increased from around 1.22% to 3.84%, the excess noise under high gain was reduced, while the location of the Bi layer—near the top or bottom of the i-region—affected the device breakdown characteristics at lower and moderate reverse biases. However, due to time constraints and growth schedules, higher-Bi n–i–p structures were not realized here, preventing a more complete comparison with their p–i–n counterparts. Further investigations on such additional n–i–p samples with higher Bi content would offer a more comprehensive experimental dataset.

Further Hole Ionization Suppression via MQW Structures

Chapter 6 examined GaAs/GaAsBi MQW structures. By distributing Bi-containing quantum wells over multiple periods (2–120), the devices not only extended the absorption band to around 1 μm but also substantially reduced hole-initiated impact

ionization at higher reverse bias, achieving a lower k = β/α. While more periods lead to higher breakdown voltages and improved gain–noise performance, strain accumulation can increase defect densities and dark current. Hence, incorporating thin GaAsBi layers periodically within the GaAs matrix can approximate the hole-suppression effect of a thick single-layer heterojunction, balancing material quality with device performance. Likewise, most of the MQW devices tested here were in p–i–n form; growing corresponding n–i–p MQW structures for direct comparison would further strengthen the conclusions.

Comprehensive Evaluation of Optical and Avalanche Performance

Across both single-layer heterojunction and MQW devices, incorporating Bi at levels above ~1% shifts the absorption edge into the near-infrared and enables measurable photoresponse in the 1.0-1.1 μm region. Within the spectral window investigated in this thesis, the longest-wavelength response is observed at approximately 1100 nm, most clearly in the high-period MQW devices where the enlarged absorption volume strengthens the sub-bandgap absorption tail. For the MQW photo-spectrum measurements, 1100 nm is also the upper limit of the available setup; therefore, extended-wavelength measurements are required to determine whether the response continues beyond 1.1 μm.

Under avalanche operation at high gain, the excess noise is notably reduced, confirming GaAsBi's potential for near-infrared APDs. Nonetheless, further extension to longer wavelengths via higher Bi contents and/or increased MQW period numbers tightens the material-quality budget: strain accumulation and defect formation can elevate dark current, and any residual surface leakage becomes increasingly important as the mesa size is reduced. These results therefore delineate a practical design space for balancing high gain, low noise, and material stability in next-generation GaAsBi-based APDs.

7.2 Future Work

The work presented in this thesis demonstrates GaAs/GaAsBi APDs with near-infrared photoresponse reaching ~1.1 μm and with clear evidence of hole-ionisation suppression under avalanche conditions. Building on this platform, future work should pursue two parallel objectives: (i) a more complete and quantitative characterisation of gain and noise over wavelength and temperature, and (ii) epitaxial and structural optimisation

aimed at extending the detection edge while maintaining low dark current, good uniformity, and robust manufacturability.

Maximum wavelength achieved and key limitations

In the present dataset, the strongest long-wavelength photoresponse is obtained in the highest-period MQW devices, with the 120QW structure exhibiting photocurrent extending to around 1100 nm under reverse bias. For the MQW photo-spectrum measurements in particular, 1100 nm coincides with the upper bound of the measurement window; therefore, extending the experimental setup to longer wavelengths (e.g., by using suitable optics and an InGaAs detector) is an immediate priority to establish the true cutoff beyond 1.1 μm.

Beyond instrumentation, pushing the absorption edge to longer wavelengths requires further bandgap reduction, most directly by increasing the Bi fraction in the absorbing region. However, higher Bi fractions and larger total GaAsBi thickness exacerbate strain accumulation and can promote defect formation and partial relaxation, which tend to elevate dark current and can compromise avalanche stability and noise performance. In addition, as the absorption edge is pushed deeper into the near-infrared, the devices become increasingly sensitive to process-dependent leakage pathways (including mesa sidewall and surface leakage), which must be controlled through improved isolation and passivation.

Temperature-dependent characterisation and detailed noise measurements

A focused measurement campaign should be carried out on a representative subset of devices (both heterojunction and MQW) to quantify avalanche gain, excess noise factor F(M), and gain-bandwidth characteristics as functions of temperature and excitation wavelength. Low-temperature multiplication and photo-spectrum measurements would help separate intrinsic impact-ionisation behaviour from thermally activated leakage mechanisms, while elevated-temperature measurements and bias-stress tests would probe stability and reliability under conditions relevant to practical deployment. Since the ultimate value proposition of GaAsBi APDs is low-noise gain, these extended measurements should prioritise noise-oriented metrics (including excess noise and instability/afterpulsing-related effects) rather than gain alone.

Growth optimisation toward higher-Bi n-i-p and higher-Bi MQWs

On the epitaxy side, the immediate growth target is to raise the Bi fraction in n-i-p heterojunction structures to match the 3-4% level already achieved in the p-i-n series,

thereby enabling a symmetric comparison and a more complete evaluation of Bi-layer placement effects. Once this milestone is reached, further growth trials can aim to increase Bi beyond the current range while preserving abrupt interfaces and low defect density.

In parallel, MQW growth should explore higher Bi fractions in the quantum wells, which would strengthen near-infrared absorption at a given period number and reduce reliance on very high period counts. Both routes must be coupled to explicit strain-management strategies, since higher Bi and higher effective GaAsBi thickness jointly reduce the available critical-thickness margin and can increase the propensity for defect generation.

Strain management and advanced device architectures

To mitigate the strain/defect penalty associated with higher Bi incorporation, strain-compensated designs, segmental epitaxy, and barrier optimisation should be explored to suppress dislocation formation while retaining the desired valence-band engineering. From a device-design viewpoint, implementing separate-absorption-multiplication (SAM) architectures would decouple the long-wavelength absorption region from the high-field multiplication region, offering an additional degree of freedom to suppress leakage and improve speed without sacrificing the intrinsic GaAsBi advantage of reduced hole ionisation.

Technological wavelength targets and integration considerations

In terms of application pull, operation around 1.06 μm is an immediately relevant target that is already approached by the present devices and is widely used in laser-based sensing systems. Extending performance toward the 1.3 μm telecommunications window would require substantially greater bandgap reduction than achieved in the current structures and therefore motivates the higher-Bi growth programme and more aggressive band-structure engineering. Finally, translation from proof-of-concept devices to deployable technology will require statistically robust large-area fabrication, improved mesa isolation and passivation, and wafer-level uniformity studies, together with consideration of integration with photonic platforms.

In conclusion, GaAsBi holds distinct advantages for avalanche detection, particularly in suppressing hole impact ionization and providing effective near-infrared coverage. By carefully controlling Bi content and quantum-well period numbers, it is possible to

maintain low noise under high gain while preserving material quality and device reliability. With continued refinement of epitaxial processes, strain management, and device architectures, GaAsBi APDs have the potential to enable high-sensitivity photodetection in near-infrared applications spanning optical communication, lidar, and related sensing systems.