Title
Non-Geiger mode single photon detector with multiple amplification and gain control mechanisms /

Permalink
https://escholarship.org/uc/item/1c91x4w3

Author
Rahman, Samia Nawar

Publication Date
2014

Peer reviewed|Thesis/dissertation
Non-Geiger mode single photon detector with multiple amplification and gain control mechanisms

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Electrical Engineering (Nanoscale Devices and Systems) by Samia Nawar Rahman

Committee in charge:
Professor Yu-Hwa Lo, Chair
Professor Peter Asbeck
Professor Prabhakar Bandaru
Professor Chung-Kuan Cheng
Professor Yuan Taur

2014
The dissertation of Samia Nawar Rahman is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____________________________________________

_____________________________________________

_____________________________________________

_____________________________________________

_____________________________________________

Chair

University of California, San Diego

2014
DEDICATION

This dissertation is dedicated to my loving parents, dear husband and dear brother.
# TABLE OF CONTENTS

Signature Page ............................................................................................................ iii
Dedication ................................................................................................................ iv
Table of Contents ...................................................................................................... v
List of Figures ........................................................................................................... viii
List of Tables .......................................................................................................... xii
Acknowledgements ............................................................................................... xiii
Vita ............................................................................................................................ xvi
Abstract of the Dissertation ................................................................................. xvii

Chapter 1 Avalanche Photo Detectors ........................................................................ 1
1.1 Introduction to Avalanche Multiplication ......................................................... 1
   1.1.1 Impact Ionization Mechanism ................................................................. 1
   1.1.2 Statistical Properties of Avalanche Multiplication ............................... 3
1.2 Important Figures of Merits & Definitions ....................................................... 6
   1.2.1 Sensitivity .............................................................................................. 6
   1.2.2 Dynamic Range ..................................................................................... 6
   1.2.3 Single Photon Detection Efficiency (SPDE) ......................................... 7
   1.2.4 Bit Error Rate (BER) ........................................................................... 7
   1.2.5 Reliability ............................................................................................ 9
1.3 Conventional Detectors .................................................................................... 10
1.4 Geiger Mode Detectors ................................................................................... 11
1.5 Motivations & Challenges .............................................................................. 15
1.6 Dissertation Outline ....................................................................................... 17
Chapter 2 Integrated Gain & Negative Feedback Mechanism ........................................ 18
  2.1 Design Motivation .......................................................................................... 18
  2.2 Operating Principle of MAGIC Detector ...................................................... 19
  2.3 Internal Gain Mechanisms ........................................................................... 25
    2.3.1 Coupled Avalanche – Bipolar Gain ...................................................... 25
    2.3.2 Field Effect Transistor (FET) Gain ..................................................... 26
  2.4 Concept of Negative Feedback .................................................................... 27
  2.5 Analytical Modeling ..................................................................................... 28
    2.5.1 Model Setup ......................................................................................... 29
    2.5.2 Stability Analysis .................................................................................. 34
    2.5.3 Steady State Gain Analysis .................................................................. 35
    2.5.4 Model Results ....................................................................................... 36
  2.6 Conclusion .................................................................................................... 38

Chapter 3 MAGIC Detector Design, Fabrication & Characterization .................. 39
  3.1 Epitaxial Design ............................................................................................ 39
    3.1.1 Design-1 ............................................................................................... 40
    3.1.2 Design-2 ............................................................................................... 42
  3.2 Device fabrication .......................................................................................... 44
    3.2.1 Iteration 1 ............................................................................................. 44
    3.2.2 Iteration 2 & 3 ..................................................................................... 53
  3.3 Device Characterization ................................................................................. 64
    3.3.1 Experimental Setup .............................................................................. 64
    3.3.2 Sensitivity Measurement ....................................................................... 66
    3.3.3 Result Analyses for iterations ................................................................ 69
      3.3.3.1 Iteration 1 .................................................................................... 69
      3.3.3.2 Iteration 2 .................................................................................... 74
      3.3.3.3 Iteration 3 .................................................................................... 79
  3.4 Conclusion .................................................................................................... 84
Chapter 4 Unified Device Modeling ........................................................................... 86
  4.1 Motivation ........................................................................................................... 87
  4.2 Model Algorithm ............................................................................................... 87
  4.3 Simulation Results ............................................................................................. 93
  4.4 Conclusion ......................................................................................................... 98

Chapter 5 Conclusions ............................................................................................... 99
  5.1 Thesis Summary ................................................................................................. 99
  5.2 Outlook .............................................................................................................. 100

References .................................................................................................................. 102
Appendix A: Matlab Code for Unified Device Model ............................................... 108
LIST OF FIGURES

Figure 1.1: The probability $P(M)$ that 1, 5 and 20 carriers result in gain $M$ with average gain of 100 (thick lines) and 500 (thin lines) ................................................................. 4

Figure 1.2: Excess Noise Factor vs Multiplication Gain for different $k$ values..............5

Figure 1.3: (a) An example of signal fluctuation. (b) Gaussian probability distribution of ‘0’ and ‘1’. The dashed region expressing the error in detection.........................8

Figure 1.4: Typical current-voltage characteristics of an APD.................................11

Figure 1.5: (a) Schematic diagram for passive quenching circuit. (b) Typical output response from passive quenched APD.................................................................13

Figure 1.6: Gated mode for single photon detection.............................................15

Figure 2.1: Schematic illustration of MAGIC detector structure with electrical connections for biasing. The drawing is not to scale....................................................... 19

Figure 2.2: Energy band diagram at different instances while a photo signal is being detected ........................................................................................................... 22

Figure 2.3: Emitter-base-collector illustration for the bipolar gain process..........26

Figure 2.4: FET channel and gate illustration.......................................................27

Figure 2.5: Schematic diagram of the structure used in the analytical model........29

Figure 2.6: Stability chart of the device. In the “unstable regime” the DC gain approaches infinity and the device operates similar to a Geiger-mode detector although at lower bias voltage. Effective device area, $A=200\mu m^2$, Multiplication region width, $W=0.8\mu m$, Barrier width, $L=0.4\mu m$, $\tau=100\text{ns}$.........................................................37

Figure 2.7: Coupled avalanche-bipolar gain vs bipolar gain for different input photons per pulse. Effective device area, $A=200\mu m^2$, Multiplication region width, $W=0.8\mu m$, Barrier width, $L=0.4\mu m$, $\tau=100\text{ns}$.........................................................37

Figure 3.1: Band positions of III-V materials lattice matched to InP .....................40
Figure 3.2: Simulated band diagram of Design-1 under zero bias.........................42

Figure 3.3: Simulated band diagram of Design-2 under zero bias.........................44

Figure 3.4: Fabrication process steps for iteration 1...........................................45

Figure 3.5: Oxide cracking in high stress points from continuous heating from Plasma causing unwanted PI etching.................................................................51

Figure 3.6: Fringes on contact pads indicating PI thickness is very small in that region and etching is almost complete. Clean pad (bottom left corner) have no PI left .......52

Figure 3.7: Fabrication steps Iteration 2 &3...............................................................55

Figure 3.8: Device at 100X magnification after second metal layer deposition..........61

Figure 3.9: Packaged device ready for testing.....................................................63

Figure 3.10: Experimental setup for sensitivity measurement..............................64

Figure 3.11: Correlation between trigger signal, laser signal and output response.....66

Figure 3.12: Current-Voltage characteristics of an SAM structure depicting the punch through condition .................................................................67

Figure 3.13: Bottom Electrode I-V. Measurement temperature,$T = 180K$, $V_{center} = 0.9V$ ...........................................................................................................69

Figure 3.14: Center Electrode I-V. Measurement temperature,$T = 180K$, $V_{center} = 0.9V$ ...........................................................................................................70

Figure 3.15: Coupled Avalanche-Bipolar gain. Measurement temperature,$T = 180K$, $V_{center} = 0.9V$ ...........................................................................................................71

Figure 3.16: Total gain for different photon number inputs. Measurement temperature,$T = 180K$, $V_{center} = 0.9V$ ...........................................................................................................72
Figure 3.17: Simulated structure for examining the effect of interface traps in the channel current

Figure 3.18: Conduction band diagram along the channel with and without interface states at a low gate or bottom electrode bias. Acceptor trap density, \( Na = 1e14 cm^{-2} \)

Figure 3.19: npn test structures for testing leakage current by applying a voltage between the two electrodes

Figure 3.20: Leakage current from n-p-n test structures with different channel interfaces

Figure 3.21: Coupled Avalanche-Bipolar gain for 6250 photons from iteration 2

Figure 3.22: Total gain for 6250 photons from iteration 2

Figure 3.23: FET gain for 6250 photons from iteration 2

Figure 3.24: Analog response for 10,000 photons per pulse at \( = 155K \)

Figure 3.25: Oscillatory behavior at high applied bias for iteration 2 devices

Figure 3.26: DC current-voltage characteristics of bottom electrode

Figure 3.27: DC current-voltage characteristics of center electrode

Figure 3.28: Peak Current distribution for different number of photons per pulse

Figure 3.29: Bit Error Rate for different photon numbers

Figure 3.30: Coupled Avalanche Bipolar gain from iteration 3 device

Figure 3.31: Coupled Avalanche Bipolar gain from iteration 3 device
Figure 4.1: MAGIC detector schematic structure with InGaAs/InP material. $\Delta E_c$ and $\Delta E_v$ are the conduction band and valence band offsets at the material interfaces. $V_b$ is the applied voltage bias. The coordinates in the Figure correspond to the coordinates in the device model...............................88

Figure 4.2: Passive quenched SPAD. A series resistor $R$, is connected with the APD to quench the single carrier response from avalanche multiplication..................91

Figure 4.3: Monte Carlo simulation flow chart for passively quenched APDs, self-quenched APDs, and the MAGIC detector.................................................................92

Figure 4.4: Distributions of output currents to single photons for different device structures ............................................................95

Figure 4.5: Mean current amplitude for different photon numbers.......................98
LIST OF TABLES

Table 3.1: Epitaxial design detail of Design 1 structure…………………………..41

Table 3.2: Epitaxial design detail of Design 2 structure…………………………..43

Table 4.1: SPDE and BER of different technologies…………………………………96
ACKNOWLEDGEMENTS

First of all, I would like to thank my Advisor, Prof. Yu-Hwa Lo, for his guidance throughout my entire Ph.D. career. His vast experience and thorough knowledge was the key to my development as a successful researcher. He would always provide advice and creative solutions to any problems I faced and always kept me encouraged whenever, I had difficulties throughout my Ph.D. career. I would also like to thank the rest of my committee members: Prof. Peter Asbeck, Prof. Prabhakar Bandaru, Prof. Chung-Kuan Cheng, and Prof. Yuan Taur for the constructive suggestions and guidance they provided.

I would like to thank all my past and present lab members for the support they have provided over the many years. I would take the chance to thank Dr. James Cheng for mentoring me and teaching me the experimental techniques and for all the useful discussions we had. I thank Yuchun Zhou, Yu-Hsin Liu, David Hall and Zhe Mei for their participation and the insightful discussions on various projects. I also thank the rest of the group for their constant friendship and support: Dr. Arthur Zhang, Dr. HongKwon Kim, Dr. Sifang You, Dr. Sung Hwan Cho, Dr. Tsung-Feng Wu, Dr. Wen Qiao, Roger Chiu, Ifitikhar Ahmad Niaz and Tiantian Zhang.

I would also like to thank Dr. Wei Lu, Dr. Winnie Chen, and Dr. Katherine Baker for the tricks and companionship they provided regarding assistance from other research groups. I also need to thank the Nano3 staff for the excellent support they provided which made my research possible: Larry Grissom, Sean Parks, Ivan Harris, Dr. Xuekun Lu, Dr. Ahmet Erten, Ryan Anderson, Dr. Maribel Montero, and Dr. Bernd Fruhberger.
Most importantly I would like to thank my family for their constant encouragement during my academic career. My Parents, Mrs. Mariam Rahman, Prof. Mujibur Rahman, and my only brother, Muntasir Raihan Rahman who are my inspirations, were always there whenever I felt discouraged or had any kind of problems. I would take the chance to convey my very special thanks to my dear husband, Hasan Faraby who keeps encouraging me to reach my dream and supported me in all possible ways.

Portions of Chapter 2 is a reprint of material as it appears in the following publication: Samia Nawar Rahman, David Hall, Zhe Mei and Yu-Hwa Lo, “Integrated 1550nm photoreceiver with built-in amplification and feedback mechanisms”, Optics Letters 38, 4166 (2013). The dissertation author was the primary investigator and author of this material.

Portions of Chapter 3 is a reprint of materials as it appears in the following publications: Samia Nawar Rahman, David Hall, Zhe Mei and Yu-Hwa Lo, “Integrated 1550nm photoreceiver with built-in amplification and feedback mechanisms”, Optics Letters 38, 4166 (2013); Samia Nawar Rahman, David Hall, Zhe Mei and Yu-Hwa Lo, “Negative feedback and multiple gain mechanisms for sensitivity improvement in 1550nm optical detection”, IEEE Photonics Conference (IPC), 364 (2013). The dissertation author was the primary investigator and author of these materials.

Portions of Chapter 4 is a reprint of material as it is accepted in the following Journal: Samia Nawar Rahman, David Hall and Yu-Hwa Lo, “Non-Geiger mode single photon detector with multiple amplification and gain control mechanisms”, Journal of
Applied Physics. The dissertation author was the primary investigator and author of this material.
VITA

2008 Bachelor of Science, Bangladesh University of Engineering & Technology
Electrical and Electronic Engineering

2011 Master of Science, University of California, San Diego
Electrical Engineering (Nanoscale Devices and Engineering)

2014 Doctor of Philosophy, University of California, San Diego
Electrical Engineering (Nanoscale Devices and Engineering)

PUBLICATIONS


ABSTRACT OF THE DISSERTATION

Non-Geiger mode single photon detector with multiple amplification and gain control mechanisms

by

Samia Nawar Rahman

Doctor of Philosophy in Electrical Engineering (Nanoscale Devices and Systems)

University of California, San Diego, 2014

Professor Yu-Hwa Lo, Chair

In this thesis, we present a new device principle, Multiple Amplification Gain with Internal Control (MAGIC), as a new class of detectors that promise high single-photon detection efficiency and a much wider dynamic range than any single photon detectors reported to date. A preliminary version of such device that contains some key features of the MAGIC detector with multiple gain mechanisms and gain control characteristics has been implemented experimentally which has shown a sensitivity of 10 photons per bit and a much wider dynamic range than Geiger-mode SPADs (Single Photon Avalanche Detectors). The device has three gain mechanisms monolithically integrated into it along with a negative feedback mechanism. The integrated gain provides a strong output signal while being operated below breakdown voltage and the
negative feedback mechanism controls the gain fluctuation. Therefore the signal to noise ratio is improved which is important for high sensitivity. This new device concept holds promise for single photon sensitivity, high detection efficiency, and wide dynamic range; and the device can achieve the above desirable characteristics at bias below the breakdown voltage to assure high device reliability. A Unified Device Model, based on physical model of the invented device has been developed to analyze its fundamental characteristics and also compare the results with conventional SPADs with different quenching mechanisms. The model takes into account all three stochastic processes that control the overall gain and fluctuations of photocurrent of the proposed device, and the model for different types of SPADs may be considered as a special case of the comprehensive model. By comparing the proposed device with passively quenched and self-quenched SPADs, the invented device shows a significant increase in detection efficiency, orders of magnitude lower error rates for single photon detection, and much greater dynamic range, attributed to the integration of avalanche multiplication and bipolar amplification as two coupled gain mechanisms in conjunction with the self-quenching feedback to regulate the output signal.
Chapter 1

Avalanche Photo Detectors

Avalanche based photodiodes (APD) are widely used due to their internal gain mechanism resulting in superior sensitivity. APDs are designed for two different modes of operations - linear mode and Geiger mode, classified in terms of different figures of merits. This chapter presents a review of these two existing APD technologies and the important figures of merits of near infrared photon detectors.

1.1 Introduction to Avalanche Multiplication

1.1.1 Impact Ionization Mechanism

Avalanche multiplication [1]–[6] is a carrier generation process where under certain conditions e.g. high electric field an electron or hole can acquire sufficient kinetic energy to knock off another e-h pair from an atom of a suitable band gap material. The newly generated electrons and holes are also energetic and can create more electron and hole pairs in the same way as the primary carrier. In this way a single carrier can generate numerous carriers through carrier multiplication by impact ionization process.
The impact ionization rate $\alpha(E)$ of electrons and $\beta(E)$ of holes are defined as the number of electron and hole pairs generated by that electron or hole in a unit distance in the electric field. $\alpha(E)$ and $\beta(E)$ depend on the electric field strength and the semiconductor band structure [7].

$$\alpha(E) = \alpha_0 \exp \left( - \frac{c_n}{E} \right)^{m_n}$$  \hspace{1cm} (1.1)

$$\beta(E) = \beta_0 \exp \left( - \frac{c_p}{E} \right)^{m_p}$$  \hspace{1cm} (1.2)

The rates in eqn. (1.1) and (1.2) can be derived by first principle calculations, but they are usually fitted to experimental data for different material over different electric field range. The ionization rates are dependent on the electric field across the material and also the energy band gap of the material [7]. Intuitively, it is harder to knock off an electron from an atom of a material with a higher band gap. The material dependent effects are included in parameters $c_n$ and $c_p$ of eqn. (1.1) and eqn. (1.2). Note that the rates are also an implicit function of temperature; as the temperature is reduced, energy loss to phonon scattering is less, which increases the ionization rate [8].

The multiplication gain of a p-i-n, with a multiplication width of W is defined by [6],

$$M = \left\{ 1 - \int_0^W \alpha \exp \left( - \int_x^W (\alpha - \beta) dx' \right) dx \right\}^{-1}$$  \hspace{1cm} (1.3)

Here, $\alpha$ and $\beta$ are spatially dependent due to non-uniformities of electric field or change of material through the space. The gain equation simplifies to the following eqn. (1.4) for uniform electric field and therefore uniform ionization rates.

$$M = \frac{1-k}{\exp(-\alpha W(1-\beta/\alpha))-\beta/\alpha}$$  \hspace{1cm} (1.4)
When the denominator of eqn. (1.4) becomes zero, the gain $M$ approaches infinity, and the voltage that gives rise to infinite gain is called the breakdown voltage.

1.1.2 Statistical Properties of Avalanche Multiplication

Avalanche multiplication is a statistical process. The number of electron hole pairs generated at each instance is different every time even though the operating conditions, such as the applied voltage and temperature, are the same for a given structure. Also, the time at which the first electron hole pair is generated is not the same at each instance. This fluctuating behavior gives rise to varying electric current and the noise from this phenomenon is known as excess noise.

Back in 1972, McIntyre developed a model to derive the gain distribution function $P_{n,n+r}$ [9]. It predicts the probability of generating $r$ electron hole pairs from $n$ number of initial carriers for a multiplication gain of $M$, where $k$ is the ratio between impact ionization rates of two types of carriers.

$$P_{n,n+r}(M) = \frac{n(1-k)^r}{(n+kr)^r} \frac{(n+r)}{(1-k)^r} \times \left( \frac{1+k(M-1)}{M} \right)^{(n+kr)} \left( \frac{M-1}{M} \right)^r$$  \hspace{1cm} (1.5)

The probability function $P(M)$ for Si is plotted for different number of carriers in Figure 1.1. For a single carrier triggered avalanche event, in most cases, there is only a few ionization events. However, there is a small but certain probability that a high gain event can be triggered. For larger number of carrier, there is always a moderate number of carriers generated by the avalanche process depending on the average gain.
Figure 1.1: The probability $P(M)$ that 1, 5 and 20 carriers result in gain $M$ with average gain of 100 (thick lines) and 500 (thin lines) [7].

Back in 1966, McIntyre also developed a model [10] to analyze the statistical gain fluctuation, otherwise known as excess noise factor. In general, excess noise factor is defined as,

$$F(M) = \frac{\langle M^2 \rangle}{\langle M \rangle^2} = 1 + \frac{\langle \Delta M^2 \rangle}{\langle M \rangle^2}$$  \hspace{1cm} (1.6)

Based on McIntyre’s model, $F(M)$ is given by,

$$F(M) = M \left(1 - (1 - k) \left(\frac{M-1}{M}\right)^2\right)$$  \hspace{1cm} (1.7)

From eqn. (1.6), it can be seen that the noise factor is a strong function of the multiplication gain, $M$ and the $k$ parameter of the material. Intuitively, higher multiplication gain means there are more carriers in the multiplication region, which in turn makes the impact ionization process more violent and the statistical fluctuation becomes higher.
A high $k$ value means both types of carriers participate in the avalanche process, and therefore the process becomes a positive feedback process, since both the newly generated electron and holes are knocking of electrons from the valence band of the atoms and generating new electron hole pairs. In this case, the noise is intuitively higher and this is also derived by McIntyre [10]. On the other hand if only one type of carrier, either electron or hole participates in impact ionization, then the process becomes more predictable and the noise factor becomes lower. Figure 1.2 shows the noise factor against the multiplication gain for various materials. We can see that Si, which has a relatively low $k$ value, has a much lower excess noise factor compared to InP which has a higher $k$ value. The $k$ value of a material is dependent on the semiconductor band structure [7].

![Figure 1.2: Excess Noise Factor vs Multiplication Gain for different $k$ values [11].](image)
1.2 Important Figures of Merits & Definitions

1.2.1 Sensitivity

Sensitivity of an optical detector is defined as the minimum power of the received light that is required to produce a certain signal to noise ratio [12]. The minimum power level or sensitivity is usually expressed in dBm or number of photons per pulse. Sensitivity is an important property for certain applications like free space communications, remote sensing, because the resolution of the data or image is directly related to the smallest detectable signal.

The sensitivity of a photo detector is dependent on the signal strength as well as the noise sources present in the system [13]. As the gain of an APD is increased, the signal strength becomes higher. However, the excess noise, which is a component of the shot noise will also increase with the gain [14] and degrade the signal to noise ratio. Besides the excess noise, thermal noise and photon statistics noise [15] also limit the signal to noise ratio. Therefore, while designing a detector, attention needs to be paid that the overall signal to noise ratio is maximized as the signal is amplified.

1.2.2 Dynamic Range

An ideal photo detector should be able to detect a large range of optical power. Dynamic range is defined as the ratio of the maximum detectable power and the minimum detectable power and is expressed in the units of decibels. The maximum detectable power can be limited by detector damage and saturation effect, whereas the minimum detectable power is sensitivity and is determined by the signal to noise ratio.
1.2.3 Single Photon Detection Efficiency (SPDE)

Single photon sensitive detectors are characterized by the single photon detection efficiency. This is the probability that the detector will generate a meaningful output signal in response to a single photon in the input. High SPDE is required for achieving high fidelity data transmission. The detection efficiency is limited by several issues, including dark counts [16] from thermal generation [17] and afterpulsing caused by trapped carriers [18].

To detect a single photon over a time interval, two conditions must be satisfied. Firstly, the incident photon must generate an electron-hole pair, which is determined by the internal quantum efficiency of the detector. Secondly, the photo generated carriers must trigger a detectable avalanche pulse. Therefore the SPDE is the product of the quantum efficiency and the avalanche probability [7].

For avalanche based single photon detectors, as the applied voltage or electric field is increased, the avalanche probability increases and the SPDE improves, but the bias voltage is eventually limited by the dark count and afterpulsing as their frequencies also increase with the bias voltage. Lowering the operating temperature can often yield higher SPDE due to the reduced dark count rate.

1.2.4 Bit Error Rate (BER)

In the context of a photo detector, BER [14] is defined as the probability of detecting light in dark or dark in light. Mathematically, if ‘light on’ condition is expressed as ‘1’ and dark condition is expressed as ‘0’, then the BER is,
\[ BER = \frac{1}{2} (P(0/1) + P(1/0)) \]  

Here, \( P(0/1) \) is the probability of deciding ‘0’, when ‘1’ is received and \( P(1/0) \) is the probability of deciding ‘1’, when ‘0’ is received.

The origin of error is the current fluctuation from different noise sources, including shot noise, thermal noise which are approximately explained with Gaussian statistics [12].

Figure 1.3 shows an example of fluctuating signal received by a detector and the Gaussian probability density [14]. The overlapped part is the region where the reception is erroneous. So intuitively, to minimize the error probability, one must focus on increasing the signal strength which will shift the ‘light on’ distribution farther from the dark distribution as well as minimizing the spread of each distribution so that the overlap is minimum.

Figure 1.3: (a) An example of signal fluctuation. (b) Gaussian probability distribution of ‘0’ and ‘1’. The dashed region expressing the error in detection [14].
1.2.5 Reliability

In addition to having high performance from a photodetector, it is also important that these detectors can withstand normal and extreme operating conditions, e.g. high applied voltage, high and low temperatures which might be necessary to push the performance limit.

Tunneling is one major reliability issue which is dominant under high electric field conditions [19]. As mentioned earlier, avalanche photodiodes require sufficient electric field for impact ionization process and depending on the required gain, the electric field required will vary.

For a reverse biased p-n diode or p-i-n diode, band to band tunneling [5], [19] occurs under high electric field. Under strong reverse bias, electrons can acquire substantial probabilities to travel from the valence band to the conduction band, producing tunneling current. This tunneling current, if comparable to or higher than the photocurrent of the detector, can saturate the overall current and make photon detection very difficult.

Tunneling mechanism is severe for low band gap materials and also when the doping of p or n region is very high [5]. Therefore, while designing APDs [20], material must be chosen in a way that keeps the tunneling current under control. For certain wavelengths e.g. 1550nm wavelength absorption, the absorption region band gap is low, about 0.75eV and, therefore, is more prone to tunneling at the required voltage for avalanche multiplication. Therefore, often a separate absorption-multiplication region structure [21] is chosen, where the absorption region is suitable for the specific
wavelength absorption, while the multiplication is designed to occur at a higher band gap material to minimize the tunneling current.

1.3 Conventional Detectors

Conventional APD detectors are reverse biased PIN diodes operating below the breakdown voltage. Figure 1.4 shows the typical current-voltage characteristics of an APD. In this specific case, the breakdown voltage is around 30 V. In conventional mode, the multiplication gain is a moderately low, finite value gain depending on the operating voltage or more accurately the electric field across the multiplication region. Since the multiplication gain is finite, the current generated by different number of photon inputs is a function of the input photon numbers and therefore, from the output current level, the input conditions can be extracted. Also, due to low operating voltage the device can withstand higher optical power without possible damage and this allows a higher dynamic range.

The state-of-the-art conventional detectors have a poor sensitivity due to the low operating gain and therefore the minimum number of detectable photons within a pulse signal is a few hundreds [22], [23]. In linear mode, the gain is usually kept low, so that the excess noise factor [14], which is a strong function of gain, stays low.
Conventional detector are reliable because the required electric field is usually much low and the tunneling current is low. Also, at low gain mode, the total current is low, which results in low power consumption and avoids any possible electro migration [24] that is usually seen in high gain APDs.

1.4 Geiger Mode Detectors

As mentioned earlier, the gain of an APD is a strong function of the electric field applied across it. Geiger mode APDs are a class of APDs that operate above breakdown voltage, determined by the breakdown electric field of the material. In this mode a small increase in voltage causes orders of magnitudes increase in the current level and the gain is very high, approaching infinite. Since the gain is extremely high (~10^5-10^6) [20], an
avalanche event triggered by a single photon can generate a strong signal, and ideally, if there is no false count or non-ideal background signal, then the avalanche signal can be easily detected. This is why Geiger mode APDs [25]–[33] are a good choice for single photon detection. In terms of sensitivity, therefore, Geiger mode APDs are superior to conventional mode detectors.

However, the dynamic range of these devices is limited because at higher optical power the devices can be damaged by high current density. Since the gain approaches infinite and the device has to be quenched by an external circuit, the photo signal generated from one, ten or hundred photons generates roughly the same strength of output signal. Therefore from the output current of these detectors, no information regarding the input photon numbers can be extracted.

Besides, practically there are many non-ideal issues that practically prevent efficient detection of single photons at every trial. False counts triggered by thermally generated carriers or trapped carriers can prevent true signal detection. While the detector is working on the false trigger, any incoming signal will be missed, and this puts a limit on the detection efficiency of these detectors. As operating bias of the device is increased, the avalanche gain increases, but the dark current or thermal generation also increases, which makes it difficult to improve the detection efficiency effectively.

Above breakdown voltage, the current increases very rapidly and there has to be some mechanisms to quench or stop the current increase to avoid device degradation from thermal runaway [20]. To do this, external [34] and internal [35], [36] quenching mechanisms have been employed in different technologies.
There are mainly three external quenching mechanisms used for quenching the avalanche pulse. Passive Quenched APDs use a large external resistor (100k-10Mohms) in series with the p-i-n diode as shown in Figure 1.5(a). When there is no avalanche current, the drop across the resistor is negligible, and the total applied voltage is dropped across the APD and it is ready to detect any external input. When avalanche pulse is triggered, the current through the circuit rapidly increases, and that causes a large voltage drop across the resistor. Since the applied voltage is constant, the voltage drop across the APD is now lower. This weakens the impact ionization process and as soon as all the carriers leave the multiplication region, the avalanche pulse is quenched.

As soon as the avalanche pulse is quenched, the voltage drop across the multiplication region starts increasing, and the APD is recovered. Due to the large RC time constant from the external resistor, the recovery charging is slow [37] and the detector cannot be operated at high frequency.

![Schematic diagram for passive quenching circuit](image1)

**Figure 1.5:** (a) Schematic diagram for passive quenching circuit. (b) Typical output response from passive quenched APD [37].
In an active quenching [33], [37], [38] circuit, a feedback loop is used to force the bias across the SPAD to a lower value as soon as the current rises above a preset level. Usually, a comparator circuit is used to monitor the current level and a controlled bias voltage source is used to control the voltage across the SPAD. This procedure is much faster than the passive quenching procedure, but requires much complicated circuitry.

Usually, for III-V based SPADs, gated mode operation [37], [39] is used, where the device bias is periodically turned high and low. When the photon pulse arrives, the device bias is high and ready to trigger an avalanche pulse in response to the incoming pulse. The ‘on’ state is usually very narrow, a few nanoseconds or just enough to trigger a pulse. In the ‘off state’, any incoming photon or dark inputs will not be detected. This mode is useful in avoiding afterpulsing [37] which is a serious problem for high defect density material. The limitation of this mode is that, the photon arrival time has to be very predictable.
Figure 1.6: Gated mode for single photon detection.

1.5 Motivations & Challenges

Near infrared single photon avalanche diodes (SPADs) find applications in free space optical communications [40], Light Detection and Ranging (LIDAR) [41], remote sensing, medical imaging, quantum computing and communications, etc. In Free Space Communications, the optically modulated signal is transmitted through the air and is collected at the receiver system where the optical signal is converted back to an electrical signal or to the original form. Signals transmitting through free space can be degraded or attenuated by different mechanisms like absorption, scattering etc. By the time the transmitted signal reaches the receiver, the signal can be very weak. The signal, if comparable to the background noise level, will make the detection process very difficult. In LIDAR technology, an optical signal is sent to a distant object and the
position or the height of that object is determined by measuring and analyzing the back scattered light from it. Again, these signals travel through air and the intensity attenuate over distance. Besides, the light transmitted back from a tiny object can be very weak. For these reasons, the sensitivity of an optical receiver is very important. Now, the optically backscattered light from a near range object or a large object can be much stronger and therefore the receiver should have the capability of detecting the strong optical signal as well. So the applications mentioned earlier require ultrahigh sensitivity to the level of single or a few photons, high photon detection efficiency, high frequency response, and significant dynamic range. However, due to the nature of the avalanche multiplication, it is not possible to attain all these required properties within a single device under the same operating conditions. In spite of decent dynamic range and good reliability, conventional APD receivers are only sensitive to at best hundreds of photons per bit [22], [23]. On the other hand, Geiger mode avalanche detectors suffer from poor single-photon detection efficiency due to false counts from thermal generation and after pulsing and poor dynamic range. Passively quenched Geiger mode detectors have poor single-photon detection efficiency due to high excess noise at high voltage and have slow recovery time. Implementing a self-quenching mechanism internally to the devices have shown improvements in minimizing the gain fluctuation of Geiger mode devices but still does not show significant improvements in the detection efficiency [20], [35], [36]. To summarize, there is no current technology which can offer ultra-high sensitivity, large dynamic range, and high detection efficiency from the same device at the same time. The goal of this work was to develop a device technology that can have all the necessary properties within the same device.
1.6 Dissertation Outline

Following the first chapter of introduction, chapter 2 will present the concept of integrated gain mechanism into a single three terminal device, and support the concept with an analytical model based on drift diffusion model and local field approximation [2], [9], [10], [42] for impact ionization. Chapter 3 will discuss the design, fabrication detail of the two iterations, explain the experimental setup and present the critical results and findings. Chapter 4 will present a Monte Carlo Model for analyzing the statistics of the device with all the necessary physical models included in it. The model has the capability to turn on and off each mechanism and examine the effects independently. Finally, Chapter 5 will summarize this dissertation.
Chapter 2

Integrated Gain & Negative Feedback Mechanism

Chapter 2 provides the physical mechanisms of the MAGIC detector device operation. After presenting the device structure and explaining the device operation, the different mechanisms will be discussed with further detail.

2.1 Design Motivation

As mentioned in chapter 1, the Geiger mode detectors, being high avalanche gain devices, can detect single photons, whereas conventional detectors can only detect at best 500-1000 photons per pulse. The reason for the low sensitivity is that these detectors can’t be operated at high avalanche gain due to the high excess noise. In one hand avalanche gain is the most important mechanism for high sensitivity, but on the other hand increasing avalanche gain to a high level, can ruin the sensitivity by increasing the current or gain fluctuation known as excess noise. Therefore, some additional gain mechanisms are needed internal to the device, if we want to detect low number of photons having the device biased below its breakdown voltage. The avalanche gain will be chosen at a level such that the excess noise does not limit the sensitivity. In addition to the multiple gain mechanisms, a negative feedback mechanism will be employed to take care of the gain fluctuation. The negative feedback mechanism
was proved to be effective in minimizing the excess noise in Geiger mode [20], [35], [36], and the theory suggests it should also be useful in Non-Geiger mode devices.

2.2 Operating Principle of MAGIC Detector

The design presented here [43], [44] integrates three amplification mechanisms: avalanche multiplication, bipolar transistor gain, and field-effect-transistor gain, with a built-in negative feedback control loop monolithically via band gap engineering of III-V compound semiconductors lattice matched to InP. The device has three terminals, as illustrated in Figure 2.1. The detailed fabrication process flow will be shown in Chapter 3.

Figure 2.1: Schematic illustration of MAGIC detector structure with electrical connections for biasing. (The drawing is not to scale).
To explain the device operating principle we will use the example of one of our epitaxial structures, which will be referred to as Design-2. Other than the heterostructure barrier and emitter design, everything else of the designs is nearly identical to Design-1 in the previous section.

In our structure, we used a separate absorption and multiplication region structure [8], [36], [45]. We had a low band gap material InGaAs to absorb the 1550nm light. The thickness of the absorption region was chosen based on Beer Lambert’s law [46], such that 100% of the incident power incident would be absorbed in that region. We used a high band gap material InP as the multiplication region. The reason we used a high band gap material for multiplication region is to minimize tunneling current [5] at high electric fields which was necessary to obtain the required avalanche gain. The photo generated carrier from the absorption region will travel to the multiplication region and initiate avalanche multiplication. On top of the p-i-n region, we had a heterostructure layer to create a potential barrier for electrons and holes. For the device to operate in a way that will be explained below, it is necessary to have a valence band offset between the p-i-n region and the heterostructure region (also known as Transient Carrier Buffer or TCB region) to provide a barrier for holes and a conduction band offset between the n-p-n region and the TCB region to put a control on any electron emission from the n+ InP region. Later, we will see, how this two band offset levels control the device performance in terms of detector sensitivity and dynamic range.

For biasing the device, there was a center electrode and an outer electrode in the lateral direction to control the current through a channel that would form during the detection process. The bottom electrode in the vertical direction was to control the
avalanche multiplication in the p-i-n region. After taking a cutline in vertical direction Figure 2.2 band diagrams are obtained. For better understanding, the instantaneous band diagrams are explained next.

As soon as photons are absorbed to produce electron-hole pairs in the InGaAs layer as shown in Figure 2.2(b), holes enter the avalanche area and experience avalanche multiplication as the first mechanism of signal amplification (Figure 2.2(c)). Holes produced by avalanche multiplication are temporarily stopped by an energy barrier produced by a heterostructure barrier (Figure 2.2(d)). This lowers the electric field in the multiplication region due to field screening effect [47] (Figure 2.2(e)). The extent of electric field screening depends on the instantaneous concentration of the accumulated holes in the barrier. At the same time, these accumulated holes reduce the energy barrier for electron injection from the emitter to the multiplication region (Figure 2.2(f)). With electron injection (Figure 2.2(g)) from the emitter under center electrode, the device experiences increased impact ionization which resulted in continued increase in the hole accumulation (Figure 2.2(h)). These accumulated holes then create a bias to the gate of the heterojunction FET (Field-Effect Transistor) with the center electrode as the source and the outer ring electrode as the drain. This lateral FET produces additional gain to the output signal (current of the center electrode). With increased hole concentration in the barrier, the electric field screening gets stronger and the field across the multiplication region gets weaker. Therefore, there are hardly any new electron hole pair generation and as soon as the accumulated holes escape the barrier by thermionic emission or tunneling, the field screening effect disappears, and the detector reverts back to its original state (Figure 2.2(a)) where it is ready to detect another photon signal.
Figure 2.2: (a)-(j) Energy band diagram at different instances during single photon detection.
Figure 2.2: (a)-(j) Energy band diagram at different instances during single photon detection (continued).
Figure 2.2: (a)-(j) Energy band diagram at different instances during single photon detection (continued).
2.3 Internal Gain Mechanisms

2.3.1 Coupled Avalanche – Bipolar Gain

The device operation consists of a coupled avalanche-bipolar gain mechanism. The electron emission is analogous to a bipolar gain mechanism, because the accumulated holes which can be thought of as the base current is causing the highly doped n⁺ InP to emit electrons and be collected at the multiplication region. Therefore, there is an emitter, base and a collector in the structure (Figure 2.3). The primary multiplied carriers trigger electron emission from the emitter, the emitted electrons trigger more impact ionization events, the enhanced avalanche event creates more multiplied holes and triggers further electron emission. These two mechanisms are so involved, that it is very tricky to decouple these two components. Avalanche gain which was introduced in Chapter 1 is the primary gain mechanism of this device. This is the first gain mechanism that comes into action when a signal is being detected by the
device. So, this gain value has to be of a certain value so that the other gain mechanisms can pick up the signal from there. If the primary avalanche gain is low, then hole accumulation is low and that might not be sufficient to trigger any electron emission or create any electron channel for a small number of photon input. If the photon signal consists of many photons, then even though the avalanche gain is low, that can trigger emission due to large number of total holes accumulated. But that’s not our targeted mode of operation. To have a highly sensitive detector, therefore, the avalanche gain has to be sufficient, but not too high to be ruined by the large excess noise or reliability issues from high voltage operation. An analytical model will be presented in section 2.5 for better understanding of the device stability and the coupled gain mechanisms.

![Emitter-base-collector illustration for the bipolar gain process.](image)

**Figure 2.3:** Emitter-base-collector illustration for the bipolar gain process.

### 2.3.2 Field Effect Transistor (FET) Gain

We have seen that the holes generated by the coupled avalanche-bipolar gain mechanism are blocked in the barrier. These holes attract electrons that are blocked by a conduction band offset between the p-InP and TCB region interface. After a certain concentration of electrons are accumulated, the p type InP starts to invert to n-type and
an electron channel is formed between the two lateral electrodes. This is analogous to a Field Effect Transistor (FET) because, the accumulated holes which can be thought of as the gate, are controlling the current through the channel for a given structure. The FET gain of the detector depends on the lifetime of the electrons in the channel which in other words depends on the lifetime of holes in the barrier or the hole escape time. The hole escape time is expressed as a parameter $\tau_p$, which will be used in the analytical model in section 2.5. The FET gain also depends on the electron transit time, ($t = \frac{l}{v}$) through the channel. Here $l$ is the channel length and $v$ is the carrier velocity. A smaller channel length and a larger lateral bias would give a higher FET gain for a given channel material. However, the channel length is limited by the process resolution. And since at high lateral bias, the current saturates, the FET gain will also saturate with increasing bias.

![Figure 2.4: FET channel and gate illustration.](image)

### 2.4 Concept of Negative Feedback

The negative feedback mechanism [48] has been employed for III-V Geiger mode detectors by our group [20], [35], [36] in the last few years. The core factor that
controls this mechanism is the presence of the heterostructure barrier or the transient carrier buffer (TCB) region design, more precisely the band offset between the TCB layer material and the nearby layer material. In section 2.2 we have seen that holes collected by the valence band offset screen the electric field across the multiplication region, thereby reducing the impact ionization rate which is a strong function of electric field [5], [6], [49]. This mechanism not only quenches the output pulse of a Geiger mode detector, but also works to minimize the gain or current fluctuation by controlling the impact ionization rate at each instance.

The built-in negative feedback suppresses the noise of avalanche process by producing instantaneous regulations of the impact ionization. As soon as the device “senses” that the number of carriers generated in the multiplication region is greater than the mean value, the rate of avalanche multiplication is reduced because more holes are accumulated and cause stronger field screening to bring down the number of carriers [50]. Conversely, if the device senses a lower impact ionization rate than the mean value, the field-screening effect is weakened to keep the impact ionization process going on until the amount of accumulated charge (holes) approaches the mean value. The experimental results presented in Chapter 3 and the Unified Device Model results presented in Chapter 4 is a manifestation of this mechanism.

2.5 Analytical Modeling

In this section, we present a physical model to describe the physics of the APDFET structure. Although a lateral FET is integrated with the detector, this FET can be treated separated as an amplifier. What is included in our treatment is the ‘electron
emitter” at the center electrode. Without including the lateral FET, we can reduce the problem into a 1-D problem while retaining all the key features of the device (i.e. self-pulsation, self-quenching).

Our treatment is based on the rate equations [6] with empirical phenomenological constants such as hole escape time, $\tau_p$. This is, of course, an oversimplified model but within a certain range of operation conditions, our treatment is expected to produce results consistent with experimental observations.

2.5.1 Model Setup

The simplified device structure (Figure 2.5) for the purpose of device modeling is shown below. Essentially the device consists of 3 regions: absorption region, multiplication region, and Transient-Carrier-Buffer (TCB) region.

The ionization coefficients for electrons and holes [5] are:

$$\alpha_n = \alpha_{oe} e^{-c_n/E} \quad \beta_p = \beta_{oe} e^{-c_p/E} \quad (2.1)$$

The continuity equations are:
\[ e \frac{\partial p(x,t)}{\partial t} = -\frac{d}{dx}J_p(x) + eG_p = -\frac{d}{dx}J_p(x) + \beta_p J_p + \alpha_n J_n \] (2.2)

\[-e \frac{\partial n(x,t)}{\partial t} = -\frac{d}{dx}J_n(x) + (-e)G_n = -\frac{d}{dx}J_n(x) - \beta_p J_p - \alpha_n J_n \] (2.3)

If we neglect diffusion currents and assume that both electrons and holes travel at their saturation velocities then we get the following,

\[ \frac{1}{v_p} \frac{\partial J_p(x,t)}{\partial t} + \frac{\partial}{\partial x} J_p(x,t) = \beta_p J_p + \alpha_n J_n \] (2.4)

\[ \frac{1}{v_n} \frac{\partial J_n(x,t)}{\partial t} - \frac{\partial}{\partial x} J_n(x,t) = \beta_p J_p + \alpha_n J_n \] (2.5)

The total current \( J \) is the sum of the electron current, hole current, and displacement current and is \( x \)-independent.

\[ J - \frac{\varepsilon}{x} \frac{dV(x,t)}{dt} = J_p(x) + J_n(x) = \text{indep. of } x \] (2.6)

Then from eqn. (2.4), (2.5) and (2.6) we get,

\[ \frac{1}{v_p} \frac{\partial J_p(x,t)}{\partial t} + \frac{\partial}{\partial x} J_p(x,t) = \beta_p J_p + \alpha_n \left[ J(t) - \frac{\varepsilon}{W} \frac{dV}{dt} - J_p(x) \right] \] (2.7)

\[ \frac{1}{v_n} \left[ \frac{dj}{dt} - \frac{\varepsilon}{W} \frac{d^2V_M}{dt^2} \right] - \frac{1}{v_n} \frac{\partial J_p(x,t)}{\partial t} + \frac{\partial}{\partial x} J_p(x,t) = \beta_p J_p + \alpha_n \left[ J(t) - \frac{\varepsilon}{W} \frac{dV}{dt} - J_p(x) \right] \] (2.8)

Performing \( v_p \times (2.7) + v_n \times (2.8) \) and rearranging the terms, we obtain,

\[ \frac{\partial}{\partial x} J_p(x,t) = (\beta_p - \alpha_n) J_p + \alpha_n \left[ J(t) - \frac{\varepsilon}{W} \frac{dV}{dt} \right] - \frac{1}{(v_p + v_n)} \left[ \frac{dj}{dt} - \frac{\varepsilon}{W} \frac{d^2V_M}{dt^2} \right] \] (2.9)

Solving eqn. (2.9) we have,

\[ e^{-AW} J_p(W_-) - J_p(0) = (1 - e^{-AW}) \frac{1}{A} \left[ \alpha_n J(t) - \frac{1}{(v_p + v_n)} \frac{dj}{dt} - \frac{\varepsilon}{W} \alpha_n \frac{d^2V_M}{dt} \right] \] (2.10)

Here, \( A \equiv \beta_p - \alpha_n \)
We use \( W_- \) to specify the boundary of the integration over the multiplication region. Since we have ignored the thickness of the accumulated holes at the interface between the multiplication region and the TCB region, we will use \( W_- \) to represent the boundary of the multiplication region and \( W_+ \) to represent the beginning of the TCB region. This distinction is important since the E-field is not continuous at \( W \) due to the accumulated holes. Now, for bipolar transistor gain \( \gamma \),

\[
J_p(W_-) = J - J_n(W + L) - \frac{e}{W} \frac{dV_M}{dt} = J - \gamma J_p(W_+) - \frac{e}{W} \frac{dV_M}{dt}
\]

(2.11)

\[
J_p(W_+) = J - J_n(W + L) - \frac{e}{L} \frac{d(V - V_M)}{dt} = J - \gamma J_p(W_+) - \frac{e}{L} \frac{d(V - V_M)}{dt} = \frac{eP_i}{\tau_p}
\]

(2.12)

\[
J_p(W_+) = \frac{1}{1+\gamma} \left[ J - \frac{e}{L} \frac{d(V - V_M)}{dt} \right] = \frac{eP_i}{\tau_p}
\]

(2.13)

Therefore, from eqn. (2.11)

\[
J_p(W_-) = J - \frac{\gamma}{1+\gamma} \left[ J - \frac{e}{L} \frac{d(V - V_M)}{dt} \right] - \frac{e}{W} \frac{dV_M}{dt} = J - \frac{\gamma}{1+\gamma} \left[ \frac{e}{L(1+\gamma)} + \frac{e}{W} \right] \frac{dV_M}{dt}
\]

(2.14)

Then we have from eqn. (2.13)

\[
P_i = \frac{\tau_p}{e(1+\gamma)} \left[ J + \frac{e}{L} \frac{dV_M}{dt} \right]
\]

(2.15)

Using eqn. (2.13) and (2.14) in the continuity equation and assuming the bipolar gain “\( \gamma \)” is a slow function of time or more or less constant, we have

\[
\frac{\partial P_i}{\partial t} = \frac{1}{e} \left( J_p(W_-) - J_p(W_+) \right) = \frac{-e}{e} \left[ \frac{1}{W} + \frac{1}{L} \right] \frac{dV_M}{dt} = \frac{\tau_p}{e(1+\gamma)} \left[ \frac{dJ}{dt} + \frac{e}{L} \frac{d^2V_M}{dt^2} \right]
\]

(2.16)

From (2.16), we obtain,

\[
\frac{dJ}{dt} = -\frac{e}{L} \frac{d^2V_M}{dt^2} - \frac{(1+\gamma)e}{\tau_p} \left[ \frac{1}{W} + \frac{1}{L} \right] \frac{dV_M}{dt}
\]

(2.17)

The relation between \( J(0_-) \) or \( V_M(0_-) \) is derived below to obtain proper initial conditions.
From eqn. (2.15), \( P_i(0_-) = \frac{\tau_p}{e(1+y)} J(0_-) \)

From eqn. (2.16), \( eP_i(0_-) = \epsilon \left[ \frac{1}{W} + \frac{1}{L} \right] [V - V_M(0_-)] \)

\[ V_M(0_-) = V - \frac{\tau_p}{e(1+y)} \left( \frac{W}{W+L} \right) J(0_-) \]

\[ J(t) = \frac{-\epsilon}{L} \left( \frac{dV_M}{dt} \right) + \frac{(1+y)e}{\tau_p} \left[ \frac{1}{W} + \frac{1}{L} \right] (V_M(0_-) - V_M(t)) + J(0_-) \]  \( (2.18) \)

From eqn. (2.14) and eqn. (2.18),

\[ J_p(W_-) = \frac{1}{1+y} \left[ -\epsilon \left( \frac{dV_M}{dt} \right) + \frac{(1+y)e}{\tau_p} \left( \frac{1}{W} + \frac{1}{L} \right) (V - V_M(t)) \right] - \left[ \frac{\epsilon y}{L(1+y)} + \frac{\epsilon}{W} \right] \left( \frac{dV_M}{dt} \right) = \]

\[ \frac{\epsilon}{\tau_p} \left( \frac{1}{W} + \frac{1}{L} \right) (V - V_M(t)) - \epsilon \left[ \frac{1}{L} + \frac{1}{W} \right] \left( \frac{dV_M}{dt} \right) \]

\( (2.19) \)

To model single photon response, we use eqn. (2.10) and the excitation optical pulse,

\[ J_p(x = 0, t) = \frac{\epsilon}{S} \delta(t) + J_p(x = 0, t = 0_-) = \frac{\epsilon}{S} \delta(t) + e^{-AW} J(0_-) - \frac{1}{1+y} \]

\[ e^{-AW} \frac{1}{A} [ \alpha_n J(0_-)] \]  \( (2.20) \)

Here, S: cross section area of the device. Now we substitute eqn. (2.17) to eqn. (2.20) into eqn. (2.10) and obtain,

\[ e^{-AW} \left\{ \frac{\epsilon}{\tau_p} \left( \frac{1}{W} + \frac{1}{L} \right) (V - V_M(t)) - \epsilon \left( \frac{1}{W} + \frac{1}{L} \right) \left( \frac{dV_M}{dt} \right) \right\} - \frac{\epsilon}{S} \delta(t) - e^{-AW} J(0_-) - \frac{1}{1+y} \]

\[ + (1 - e^{-AW}) \frac{1}{A} [ \alpha_n J(0_-)] \]

\[ = (1 - e^{-AW}) \frac{1}{A} \left\{ \alpha_n \left[ -\epsilon \left( \frac{dV_M}{dt} \right) + \frac{(1+y)e}{\tau_p} \left( \frac{1}{W} + \frac{1}{L} \right) (V - V_M(t)) \right] - \frac{1}{(v_p+v_n)} \left[ \frac{-\epsilon}{L} \frac{d^2V_M}{dt^2} \right. \]

\[ \left. - \frac{(1+y)e}{\tau_p} \left( \frac{1}{W} + \frac{1}{L} \right) \left( \frac{dV_M}{dt} \right) - \frac{\epsilon}{W} \alpha_n \frac{dV_M}{dt} + \frac{\epsilon}{W(v_p+v_n)} \frac{d^2V_M}{dt^2} \right\} \]

\[ \Rightarrow C_1 (V - V_M) + C_2 \frac{dV_M}{dt} + C_3 \frac{d^2V_M}{dt^2} \]
\[\frac{\dot{Y}(t)}{Y(0)} = \frac{\epsilon}{S} \delta(t) + J(0) \left[ e^{-AW} \left( \frac{1}{1+\gamma} + \frac{\alpha_n}{A} \right) - \frac{\alpha_n}{A} \right] \]  

(2.22)

\[C_1 = e^{-AW} \frac{\epsilon}{\tau_p} \left( \frac{1}{W} + \frac{1}{L} \right) - (1 - e^{-AW}) \frac{1}{A} \alpha_n \frac{(1 + \gamma)\epsilon}{\tau_p} \left( \frac{1}{W} + \frac{1}{L} \right) \]

\[= -\frac{1}{A} \alpha_n \frac{(1 + \gamma)\epsilon}{\tau_p} \left( \frac{1}{W} + \frac{1}{L} \right) + e^{-AW} \frac{\epsilon}{\tau_p} \left( \frac{1}{W} + \frac{1}{L} \right) \left[ 1 + \frac{\alpha_n}{A} (1 + \gamma) \right] \]

\[= \frac{\epsilon}{\tau_p} \left( \frac{1}{W} + \frac{1}{L} \right) e^{-AW} \left\{ 1 + \frac{\alpha_n}{A} (1 - e^{AW}) (1 + \gamma) \right\} \]

\[\therefore C_1 = \frac{\epsilon}{\tau_p} \left( \frac{1}{W} + \frac{1}{L} \right) e^{-AW} \left\{ 1 + \frac{\alpha_n}{A} (1 - e^{AW}) (1 + \gamma) \right\} \]  

(2.23)

\[C_2 = -e^{-AW} \frac{\epsilon}{\tau_p} \left( \frac{1}{W} + \frac{1}{L} \right) + (1 - e^{-AW}) \frac{1}{A} \alpha_n \frac{\epsilon}{L} \]

\[- (1 - e^{-AW}) \frac{1}{A} \frac{1}{\tau_p} \left( \frac{1}{W} + \frac{1}{L} \right) + (1 - e^{-AW}) \frac{1}{A W} \alpha_n \]

\[= \frac{\epsilon}{\tau_p} \left[ \frac{\alpha_n}{A} - \frac{(1 + \gamma)}{A(v_p + v_n) \tau_p} \right] \left( \frac{1}{W} + \frac{1}{L} \right) - e^{-AW} \left[ 1 + \frac{\alpha_n}{A} - \frac{(1 + \gamma)}{A(v_p + v_n) \tau_p} \right] \]

\[= \frac{\epsilon}{\tau_p} \left[ \frac{\alpha_n}{A} - \frac{(1 + \gamma)}{A(v_p + v_n) \tau_p} \right] \left( \frac{1}{W} + \frac{1}{L} \right) e^{-AW} \]

\[\therefore C_2 = \frac{\epsilon}{\tau_p} \left[ \frac{\alpha_n}{A} - \frac{(1 + \gamma)}{A(v_p + v_n) \tau_p} \right] \left( \frac{1}{W} + \frac{1}{L} \right) e^{-AW} \]  

(2.24)

\[C_3 = -e(1 - e^{-AW}) \frac{1}{A(v_p + v_n)} \left( \frac{1}{W} + \frac{1}{L} \right) \]

(2.25)

Eqn. (2.21) is a highly non-linear differential equation because all \(C\)'s are time dependent through \(A\) and \(\alpha_n\) in \(C\)'s being highly sensitive to \(V_M\).

We can solve the differential equation using the initial condition, \(V'_M(0^-) = 0\)  

(2.26)
We need to find $V_M(0^-)$ to solve the differential equation. $V_M(0^-)$ is a parameter that is to be chosen. It cannot be “found”. Physically, $V_M(0^-)$ is determined by the “history” of the device including any thermal generation, background illumination, etc.

### 2.5.2 Stability Analysis

In practice the device will reach a steady state condition after initial single photon (or thermal) excitation. On the other hand, if the device cannot find a steady state solution after initial excitation, the device is unstable and tends to be self-oscillating.

To find the steady-state solution and to study the device stability, we try to solve eqn. (2.22),

$$C_1(V - V_M) + C_2 \frac{dV_M}{dt} + C_3 \frac{d^2(V_M)}{dt^2} = \frac{e}{S} R u(t) + J(0_-) \left\{ e^{-AW} \left[ \frac{1}{1+y} + \frac{\alpha_n}{A} \right] - \frac{\alpha_n}{A} \right\}$$

The additional term “R” is the average rate of constant illumination applied to the device at $t=0$, and $u(t)$ is the unit step function. Assuming the steady-state exists, $d/dt=0$ as $t \to \infty$, we obtain

$$V_M(\infty) = V - \frac{e}{C_1 S} R - \frac{J(0_-)}{C_1} \left\{ e^{-AW} \left[ \frac{1}{1+y} + \frac{\alpha_n}{A} \right] - \frac{\alpha_n}{A} \right\}$$ (2.27)

$$C_1(t \to \infty) = \frac{e}{\tau_p} \left( \frac{A}{W} + \frac{1}{L} \right) e^{-AW} \left\{ 1 + \frac{\alpha_n}{A} (1 - e^{AW})(1 + y) \right\}$$ (2.28)

$$A(\infty) = \beta_{oe} e^{-c_p W/V_M} - \alpha_{oe} e^{-c_n W/V_M}$$

$$\alpha_n(\infty) = \alpha_{oe} e^{-c_n W/V_M}$$ (2.29)
To find a physically meaningful solution, we require, \( V_M \leq V \), or \( C_1(t \to \infty) > 0 \). Otherwise, the steady-state solution does not exist or \( d/dt \) cannot become zero as \( t \to \infty \) and the device will oscillate and produce self-pulsation. Therefore we achieve the necessary condition for a stable solution of the device: \( C_1(t \to \infty) > 0 \) or equivalently,

\[
1 + \frac{\alpha_n}{\alpha_n} (1 - e^{AW})(1 + \gamma) > 0 \quad \text{as} \ t \to \infty \tag{2.30}
\]

In the case of InP, \( A \equiv \beta_p - \alpha_n > 0 \) in all bias conditions. So, eqn. (2.30) can be written as

\[
\gamma < \frac{A(t=\infty)}{\alpha_n(e^{AW-1})} - 1 \tag{2.30-a}
\]

This is the necessary and sufficient condition for the device to be stable.

In the extreme case of \( \gamma = 0 \) (i.e. no bipolar effect or no electron injection), eqn. (30-a) gives rise to

\[
(e^{A(\infty)W} - 1)\alpha_n < A(\infty)
\]

\[\Rightarrow \beta_p - \alpha_n < \frac{1}{W} [\ln(\beta_p) - \ln(\alpha_n)] \tag{2.31}\]

### 2.5.3 Steady State Gain Analysis

From eqn. (2.18) and eqn. (2.27, 2.28):

\[
J(\infty) = \frac{(1+\gamma)e}{\tau_p} \left[ \frac{1}{W} + \frac{1}{L} \right] e^{\frac{1}{C_1}S}R + J(0-)
\]

So the steady-state gain becomes

\[
G = \frac{J(\infty) - J(0-)}{eR} S = \frac{(1+\gamma)e}{\tau_p C_1} \left[ \frac{1}{W} + \frac{1}{L} \right] e^{\frac{(1+\gamma)e^{AW}}{1+\alpha_n/A(1-e^{AW}(1+\gamma))}} \tag{2.32}
\]
Although the steady-state gain of the device appears to be independent of the initial condition since the expression of eqn. (2.32) does not show explicit dependence on \( J(0_-) \), in fact \( J(0_-) \) has strong effect on the magnitude of steady-state gain. By choosing a \( J(0_-) \) value, the value of \( V_m(0_-) \) is determined and subsequently the steady-state gain is determined.

2.5.4 Model Results

For a given negative feedback strength characterized by hole escape time \( \tau_p \), a stability chart can be obtained (Figure 2.6), which indicates the maximum operating voltage after which the device enters the unstable mode, manifested by its DC gain approaching infinity in a similar fashion as a Geiger-mode device. For \( \tau_p = 100\,\text{ns} \), the stability condition from eqn. (2.28) is derived as a function of the applied voltage. For a larger bipolar gain, the device reaches unstable mode at an earlier voltage. This is intuitively logical, since a larger bipolar gain means more electron emission and that can cause the device to become unstable at a smaller voltage.

From eqn. (2.32), the coupled gain is plotted (Figure 2.7) against the bipolar gain, \( \gamma \) for different number of photon inputs. It can be seen that the coupled gain is predominantly from avalanche multiplication at low \( \gamma \) value, but as \( \gamma \) increases, the avalanche and bipolar amplifications both play important roles and the coupled gain increases substantially. For larger number of photons per pulse, the gain is smaller due to the negative feedback mechanism which will be explained in Chapter 3 with the experimental results.
Figure 2.6: Stability chart of the device. In the “unstable regime” the DC gain approaches infinity and the device operates similar to a Geiger-mode detector although at lower bias voltage. Effective device area, $A=200\text{um}^2$, Multiplication region width, $W=0.8\text{um}$, Barrier width, $L=0.4\text{um}$, $\tau_p=100\text{ns}$ [43].

Figure 2.7: Coupled avalanche-bipolar gain vs bipolar gain for different input photons per pulse. Effective device area, $A=200\text{um}^2$, Multiplication region width, $W=0.8\text{um}$, Barrier width, $L=0.4\text{um}$, $\tau=100\text{ns}$ [43].
2.6 Conclusion

This chapter introduced the concept of multiple gain integration into a single three terminal device and discussed the origin and contribution for each gain mechanisms. A negative feedback mechanism was also presented which was used to minimize gain fluctuation of the device by using the history of the carriers. A deterministic model is formulated to explain the gain integration and stability conditions of the device operation.

This chapter, in part, is a reprint of material as it appears in the following publication:

Samia Nawar Rahman, David Hall, Zhe Mei and Yu-Hwa Lo, “Integrated 1550nm photoreceiver with built-in amplification and feedback mechanisms”, Optics Letters 38, 4166 (2013). The dissertation author was the primary investigator and author of this material.
Chapter 3

MAGIC Detector Design, Fabrication and Characterization

Chapter 3 provides a detailed description of the device design for different iterations, the fabrication steps with the critical details. The experimental setup is also provided with elaborate information about the instruments used, and finally the device characterization and results are presented with physical explanations.

3.1 Epitaxial Design

The following section will present and discuss the epitaxial design of the different iterations of the MAGIC detector which was implemented for 1550nm light detection. All the layers in the design are lattice matched to III-V InP material. Figure 3.1 can explain the band alignments of the layers for both designs. In general, both designs start with an InP substrate and InP buffer layer. On top of this sits the thick InGaAs absorption layer for 1550nm wavelength absorption. To ensure smooth transport of hole from absorption region to the multiplication region, a grading layer set of InGaAsP is introduced between the absorption region and the n-InP region. Then there is the p-i-n InP region for impact ionization mechanism. The TCB layer and the emitter layer design will be varied in the following two epitaxial designs.
3.1.1 Design-1

Design-1 had a quenching barrier equivalent to the valence band offset of $\Delta E_v = 137\text{meV}$ between the InAlAs and InP [20]. The quenching barrier was effectively much stronger than this due to five periods of this InP-InAlAs layers. The bipolar emission was designed with a conduction band offset of $\Delta E_c = 286\text{meV}$ between the InP and InAlAs the emitter region [20], [51].

Figure 3.1: Band positions of III-V materials lattice matched to InP [20].
Table 3.1: Epitaxial design detail of Design-1 structure.

<table>
<thead>
<tr>
<th>Layer</th>
<th>Material</th>
<th>Thickness (µm)</th>
<th>Doping (cm$^{-3}$)</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>InP</td>
<td></td>
<td>n-type</td>
<td>Substrate</td>
</tr>
<tr>
<td>1</td>
<td>InP</td>
<td>0.5</td>
<td>n-1×10$^{18}$</td>
<td>Buffer</td>
</tr>
<tr>
<td>2</td>
<td>InGaAs</td>
<td>1.5</td>
<td>-</td>
<td>Absorption</td>
</tr>
<tr>
<td>3</td>
<td>InGaAsP (1.55µm)</td>
<td>0.03</td>
<td>-</td>
<td>Grading</td>
</tr>
<tr>
<td>4</td>
<td>InGaAsP (1.3µm)</td>
<td>0.03</td>
<td>-</td>
<td>Grading</td>
</tr>
<tr>
<td>5</td>
<td>InGaAsP (1.1µm)</td>
<td>0.03</td>
<td>-</td>
<td>Grading</td>
</tr>
<tr>
<td>6</td>
<td>InP</td>
<td>0.2</td>
<td>n-1.15×10$^{17}$</td>
<td>Field Control</td>
</tr>
<tr>
<td>7</td>
<td>InP</td>
<td>0.8</td>
<td>-</td>
<td>Multiplication</td>
</tr>
<tr>
<td>8</td>
<td>InP</td>
<td>0.2</td>
<td>p-2.5×10$^{17}$</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>InAlAs</td>
<td>0.05</td>
<td>-</td>
<td>TCB</td>
</tr>
<tr>
<td>10</td>
<td>InP</td>
<td>0.05</td>
<td>-</td>
<td>TCB</td>
</tr>
<tr>
<td>11</td>
<td>InAlAs</td>
<td>0.05</td>
<td>-</td>
<td>TCB</td>
</tr>
<tr>
<td>12</td>
<td>InP</td>
<td>0.05</td>
<td>-</td>
<td>TCB</td>
</tr>
<tr>
<td>13</td>
<td>InAlAs</td>
<td>0.05</td>
<td>-</td>
<td>TCB</td>
</tr>
<tr>
<td>14</td>
<td>InP</td>
<td>0.05</td>
<td>-</td>
<td>TCB</td>
</tr>
<tr>
<td>15</td>
<td>InAlAs</td>
<td>0.05</td>
<td>-</td>
<td>TCB</td>
</tr>
<tr>
<td>16</td>
<td>InP</td>
<td>0.05</td>
<td>-</td>
<td>TCB</td>
</tr>
<tr>
<td>17</td>
<td>InAlAs</td>
<td>0.05</td>
<td>-</td>
<td>TCB</td>
</tr>
<tr>
<td>18</td>
<td>InP</td>
<td>0.03</td>
<td>5×10$^{16}$-1×10$^{17}$</td>
<td>Channel</td>
</tr>
<tr>
<td>19</td>
<td>InGaAsP (1.3µm)</td>
<td>0.05</td>
<td>5×10$^{17}$</td>
<td>Contact/Emitter</td>
</tr>
<tr>
<td>20</td>
<td>InP</td>
<td>0.15</td>
<td>2×10$^{18}$</td>
<td>Emitter</td>
</tr>
</tbody>
</table>
3.1.2 Design-2

After sensitivity measurement from Design-1 device, the sensitivity obtained was around 5000 photons per pulse, and the measurement results and the device physics supported the fact that the avalanche gain, which was supposed to be the primary gain mechanism, was too weak due to very strong quenching barrier and the bipolar emission was too strong due to the small conduction band offset in the emitter region. So, effectively, the device behaved like a conventional n-p-n bipolar junction phototransistor which is limited by thermal noise. Therefore, to adjust the balance of the two gain mechanisms, the avalanche gain needed to be increased. To achieve this the valence band offset for hole barrier was designed to be 100meV [52] between InAlAs-
InGaAlAs layer. To make bipolar injection much weaker another conduction band offset of 270meV [20], [51] was added in series with the previous design. This was implemented by inserting a thin InGaAs layer between InP Emitter layer and p-InP [Table 3.2]. Band diagram of this structure around the emitter and quenching barrier under zero bias is shown in Figure 3.2.

Table 3.2: Epitaxial design detail of Design-2 structure.

<table>
<thead>
<tr>
<th>Layer</th>
<th>Material</th>
<th>Thickness (µm)</th>
<th>Doping (cm$^{-3}$)</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>InP</td>
<td>n-type</td>
<td></td>
<td>Substrate</td>
</tr>
<tr>
<td>1</td>
<td>InP</td>
<td>0.5</td>
<td>n-1×10$^{18}$</td>
<td>Buffer</td>
</tr>
<tr>
<td>2</td>
<td>InGaAs</td>
<td>1.5</td>
<td>-</td>
<td>Absorption</td>
</tr>
<tr>
<td>3</td>
<td>InGaAsP (1.55µm)</td>
<td>0.03</td>
<td>-</td>
<td>Grading</td>
</tr>
<tr>
<td>4</td>
<td>InGaAsP (1.3µm)</td>
<td>0.03</td>
<td>-</td>
<td>Grading</td>
</tr>
<tr>
<td>5</td>
<td>InGaAsP (1.1µm)</td>
<td>0.03</td>
<td>-</td>
<td>Grading</td>
</tr>
<tr>
<td>6</td>
<td>InP</td>
<td>0.2</td>
<td>n-1.15×10$^{17}$</td>
<td>Field Control</td>
</tr>
<tr>
<td>7</td>
<td>InP</td>
<td>0.8</td>
<td>-</td>
<td>Multiplication</td>
</tr>
<tr>
<td>8</td>
<td>InP</td>
<td>0.2</td>
<td>p-2.5×10$^{17}$</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>InGaAlAs (1.1µm)</td>
<td>0.05</td>
<td>-</td>
<td>TCB</td>
</tr>
<tr>
<td>10</td>
<td>InAlAs</td>
<td>0.05</td>
<td>-</td>
<td>TCB</td>
</tr>
<tr>
<td>11</td>
<td>InP</td>
<td>0.03</td>
<td>1×10$^{17}$-2×10$^{17}$</td>
<td>Channel</td>
</tr>
<tr>
<td>12</td>
<td>InGaAs</td>
<td>0.03</td>
<td>5×10$^{17}$</td>
<td>Contact/Emitter</td>
</tr>
<tr>
<td>13</td>
<td>InGaAsP (1.3µm)</td>
<td>0.03</td>
<td>5×10$^{17}$</td>
<td>Contact/Emitter</td>
</tr>
<tr>
<td>14</td>
<td>InP</td>
<td>0.14</td>
<td>2×10$^{18}$</td>
<td>Emitter</td>
</tr>
</tbody>
</table>
3.2 Device Fabrication

3.2.1 Iteration 1

The device fabrication steps for iteration 1 are depicted in Figure 3.4. In this iteration, Design-1 epi-layer was used.
The first step was to define and deposit the center and outer metal electrode (Figure 3.4(a)). Negative photoresist NR9-1500PY was used to pattern the layer. The photoresist was spun at 3000rpm to get a nominal thickness of 1.5microns. The next steps were standard prebake for 80s at 150°C, UV exposure using Karl Suss MA6 mask aligner at 11mW/cm² for 9s. For negative resist, a post exposure bake is done at 100°C for 60s-80s, and then the resist is developed using commercial developer Futurrex RD6 for 10s. After patterning the resist, 30nm Ti and 100nm Au was deposited using Temescal BJD 1800 electron beam evaporator at a rate of about 1-1.5A°/s. The purpose

Figure 3.4: Fabrication process steps for iteration 1.
of the initial 30nm Ti was to provide adhesion of the Au metal layer with the InP substrate which is important during the wire bonding or probing steps. Then lift off was performed to remove the photoresist using Acetone, IPA and DI water. For efficient lift off, the resist thickness should be at least three times of the total metal thickness, so that the solvents can easily attack the photoresist beneath the metal. The reason why electron evaporation was chosen over sputtering or other deposition techniques was because, the directionality of this process is very good, which is critical to avoid severe sidewall deposition which could block the photoresist and hinder the lift off process.

After the metal electrodes are defined, the npn transistor was defined by selective etching of the top n+ layers (InP and InGaAsP) in Figure 3.4(b). First, positive resist Shipley S1805 was spun at 3000rpm (to get thickness of 500nm-700nm). After prebaking at 115° for 90s, the sample was exposed through the mask aligner for 9s. Developer Microposit MF321 was used to develop the resist for 40s, which was followed by a post bake at 115° for 180s. After development and before etching, the sample went through a 1 min O₂ descum at 200mTorr flow, 200W power to remove any residual photoresist. After patterning, the etching solution of 3:1 HCl: H₂O was prepared by adding 30mL of HCl into 10mL of H₂O and stirring for about 2 mins to stabilize the solution. The etch rate of InP in this solution is about 100nm/s. During this etching step, bubbles should appear on the reactive surface of the sample. InGaAsP acts as an etch stop layer for this process [53]. To etch these layers a diluted H₂SO₄ solution is prepared by adding 10mL 95%H₂SO₄ in 100mL H₂O and then adding 10mL H₂O₂ into it. For this step the p-InP layer acts as an etch stop layer. The etch rate in this solution is about 1nm/s. After each step, the sample is rinsed with DI water very thoroughly. The reason
wet etching is used to etch the n+ layers is because the good selectivity allows for some overetch to make sure there is no n+ layer residue which could create a leakage path in the channel and also to avoid any surface damage by a dry etch process. After etching is completed, the photoresist was rinsed in Acetone, IPA and water thoroughly. The height of the n+–p step was confirmed by Dektak 150 Stylus Profilometer or Veeco NT1100 Optical Profiling System.

After the wet etching step an oxide pad was patterned far away from the device to form a contact pad on it later (Figure 3.4(c)). In the same step, the oxide was also deposited on the channel to passivate the surface. SiO₂ was deposited using Oxford PECVD at 350°C. SiH₄ and N₂O mixture was used to form the SiO₂ on the sample. The total thickness of the insulator layer was 400nm with a refractive index of 1.45 at 635nm wavelength. The thickness was chosen such that the insulator does not breakdown at the bias of operation. The deposition rate was about 80nm/min. A blank substrate was placed in the chamber in addition to the actual sample. After deposition, the thickness was measured on that using Filmetrics F20 and it was also used to calibrate the etch rate in the etch solution. Positive photoresist Shipley S1805 was patterned on the real sample using the similar recipe used for the previous step of wet etching. Then, Buffered Oxide Etch (6:1 NH₄F:HF) was used to etch the oxide in the exposed area of the real sample. BOE etch cannot be indefinitely long for two reasons, first is the undercut due to the isotropic etch profile, second is that the Ti layer of the first metal layer will be attacked by the BOE. While designing the photo mask, the undercut due to the wet etch process was already taken into account. After etching, the photoresist was stripped using solvents and afterwards, an O₂ descum at 200mTorr flow, 200W power for a few
mintues. An overnight soak in acetone would help in case the resist is very hard to remove.

Mesa profiles were formed around the devices and the contact pads (Figure 3.4(d)). To eliminate the area that is not part of the device, especially attention was given such that the non-active area around the devices were etched all the way through the low band gap InGaAs area, so that the dark current from the area does not affect the measured current. First, Shipley S1818 positive photoresist was patterned to obtain a resist thickness of about 1.5 microns. The resist was spun at 4500rpm for 40s, prebaked at 105°C for 3min, exposed in the mask aligner for 18s, and developed in Microposit MF-321 for 40s. The etching solution was prepared by mixing 4:1:25 HBr:H₂O₂:H₂O vigorously for several minutes. This solution is non selective and will etch all the III-V layers of the epitaxial layer with a very negligible difference in etch rate. The solution was used to etch a few dummy patterned samples to calibrate the etch rate. The etch rate will increase with stirring and after a while the rate will be stable, and that etch rate is used to etch the exposed area. Care was taken so that the solution was used within few hours of preparing it, because the solution would expire and would not be useful after that. The samples were pasted in a microscope slide using photoresist for easier handling, and it was stirred vigorously in the etching solution to get a uniform and clean etching. For testing the etch rate, the step height for each sample was measured using Dektak 150 Stylus Profilometer. When the required step height from the real sample was achieved, the photoresist was stripped using solvents and O₂ descum.
After etching the III-V layers, lots of dangling bonds are exposed, and they had to be passivated immediately before the surface changes too much and permanent leakage path forms through the mesa sidewalls. Therefore, a Polyimide layer PI-2610 was spun on the sample as soon as the etching was done (Figure 3.4(e)). The PI-2610 was first diluted to obtain the desired thickness using the ration of 2:1 PI-2610:T9039. The solution after curing has to be thick enough, to cover the mesa sidewalls effectively without exposing the edge on the top of the island, but thin enough so that metal layer, which will be deposited later to form contact with the device, can run over the PI layer continuously without breaking. This is critical because the PI profile is vertical after etching. On the real sample and a blank Si sample, PI-2610 was spun at 500rpm for 20s and 2000rpm for 55s, then prebaked at 90°C for 90s and 150°C for 90s. Then the PI layer was cured at 350°C for 30mins, using a ramp up time of 80min and a ramp down time of 33 min in the Ulvac MILA-3000 Minilamp annealer. The slow ramp rate was chosen to avoid any thermal shock in the material. After this the PI thickness in the dummy sample was measured using Filmetrics 20 and the thickness range was found to be in the order of 1microns. RIE O₂ plasma was used to etch the PI-2610. Conventionally, CF₄ is also used for faster etching of the material, however, we had to avoid this Fluoride based components to avoid oxide etching and III-V layer etching. The selectivity between photoresist and PI-2610 for O₂ is very poor. For that reason a SiO₂ hard mask was used to pattern. First, 90nm-100nm SiO₂ was sputtered using AJA RF Sputter Deposition Tool at room temperature. The process parameters used were, 300W forward power, 10sccm Ar flow, 10mTorr pressure, 50% stage rotation for about 1 hr. Again a blank Si sample was used to measure the thickness and calibrate the etch
rate of the oxide layer. Negative photoresist NR9-3000PY was used to pattern the oxide hard mask. The resist was spun at 4500rpm for 40s, prebaked at 150°C for 2min. This bake time is longer than the suggested time in the datasheet to account for the fact that the curing is slower due to the thick insulator (PI-2610) layer. Then the resist is exposed through the mask aligner for 10s, and afterwards a post exposure baking is done for 80s at 100°C. Futurexx developer RD6 is used to develop the resist for 10s. All the baking time, exposure time and develop time might vary a little depending on the room temperature, humidity and lamp life of the mask aligner. After photoresist patterning, the oxide is etched using BOE or RIE etching. RIE etching is preferred due to the smaller feature size, but at the end a quick BOE etch is also done to clean the surface. The real sample and dummy sample with oxide is put in the Oxford Plasmalab 80 RIE chamber to etch SiO₂ using 25sccm Ar and 25sccm CHF₃ at 35mTorr, 200W. Once the oxide is etched and the PI is exposed, the sample was put back into the same chamber again. Due to absence of CF₄ in the etching of a 1micron thick layer, the total etch time required was very long. If the plasma was turned on continuously for a long time, what happened was that the PI would expand and break the oxide in the critical points like sharp corners or sharp edges. So, the oxide mask was damaged in those points and the PI was removed from those area (Figure 3.5).
Figure 3.5: Oxide cracking in high stress points from continuous heating from Plasma causing unwanted PI etching.

To avoid cracking, the plasma etching was performed for 1 min, and then a 1 min rest was given, when the gas flow was present, but the RIE power was 0W. During the etching steps, the conditions were 100W power, 100mTorr pressure, 40sccm $O_2$. After few cycles of etching, the samples were taken out and examined under the optical microscope. The PI is very non uniform (Figure 3.6), and therefore all the openings did not clear at the same time. When the thickness of the PI was very small, there would be fringe like patterns (Figure 3.6) in most of the openings and some opening were already clear by that time. That’s when the PI etching was almost complete. Calculating the etch rate till that point, the remaining PI was removed and a small overetch was also performed to make sure the surface of the opening was clean.
Figure 3.6: Fringes on contact pads indicating PI thickness is very small in that region and etching is almost complete. Clean pad (bottom left corner) have no PI left.

Using Dektak 150 the PI thickness on top of the contact pad and the device openings were measured. Since the electron beam evaporation is directional deposition, there is minimum side wall deposition. So to make continuous metal line from contact pad to the top of the wall, a thick metal comparable to the PI thickness needed to be deposited. If the PI sidewall was about 1um thick, at least 700nm metal thickness would be needed to get a high yield from the process in terms of connected devices. Sputtering would be an option to have good side wall coverage, but then lift off becomes practically impossible especially for small patterns. In electron beam evaporation, if 700nm was deposited in the same run, the chamber would become too hot and damage the resist. In that case, the lift off would be very difficult and sometimes no lift off was possible at all. For these reason the tedious path of patterning and depositing and lifting off multiple times was chosen. At each run, no more that 300nm was deposited. A thick PR NR9-3000 was spun at a slow speed to get a thicker resist for easy lift off. 100nmTi and 200nm Au was evaporated (Figure 3.4(f)) using Temescal BJD 1800 electron beam evaporator. The deposition rate was very slow about 1Å/s to allow a very directional
deposition by minimizing sputtering which is very important for a good lift off. Also
the base pressure before starting the deposition was chosen to be low (<2e-7 Torr) to
minimize scattering for minimum sidewall deposition. After evaporation, the sample
was rested in acetone for 5 mins, rinsed for 5 mins and then sprayed hard till most of
the metal came off. If the lift off seemed hard, the sample was soaked in acetone
overnight and lifted off next day. A very gentle rub of a cleanroom swab was used to
remove pieces of metal films that were attaching by the end and won’t go by other
means. To this point the devices were taken ready to measure the dc current-voltage
characteristics between the top contacts. The I-Vs of several devices were measured to
see if the second metal layer was connected to the devices, or still are open. If most of
them seemed disconnected, another metal layer would be added by repeating the same
procedure used above, and the thickness was determined by seeing the metal in the
Scanning Electron Microscope. If there were some devices that showed normal I-V, we
knew, only a thin metal layer would be sufficient. Attention was given while redoing
the metal layer to maintain minimum misalignment.

To make a backside contact Indium solder was pasted in the backside of the
substrate (Figure 3.4(g). A backside metallization could also be done as well for this
purpose.

3.2.2 Iteration 2 & 3

The device fabrication steps for iterations 2 and 3 are depicted in Figure 3.7.
Iteration 2 was implemented using Design-1 and iteration 3 was implemented using
Design-2 epilayer. In both runs, the first step was to define and deposit the center and
outer metal electrode (Figure 3.7(a)). Negative photoresist NR9-1500PY was used to pattern the layer. The photoresist was spun at 3000rpm to get a nominal thickness of 1.5microns. The next steps were standard prebake for 80s at 150°C, UV exposure using Karl Suss MA6 mask aligner at 11mW/cm² for 9s. For negative resist, a post exposure bake is done at 100°C for 60s-80s, and then the resist is developed using commercial developer Futurrex RD6 for 10s. After patterning the resist, 30nm Ti and 100nm Au was deposited using Temescal BJD 1800 electron beam evaporator at a rate of about 1-1.5A°/s. The purpose of the initial 30nm Ti was to provide adhesion of the Au metal layer with the InP substrate which is important during the wire bonding or probing steps. Then lift off was performed to remove the photoresist using solvents Acetone, IPA and DI water. For efficient lift off, the resist thickness should be at least three times of the total metal thickness, so that the solvents can easily attack the photoresist beneath the metal. The reason why electron evaporation was chosen over sputtering or other deposition techniques was because, the directionality of this process is very good, which is critical to avoid severe sidewall deposition which could block the photoresist and hinder the lift off process.

After the metal electrodes are defined, the npn transistor was defined by selective etching of the top n+ layers (InP and InGaAsP for Design-1) and (InP, InGaAsP and InGaAs for Design-2) in Figure 3.7(b). First, positive resist Shipley S1805 was spun at 3000rpm (to get thickness of 500nm-700nm). After prebaking at 115° for 90s, the sample was exposed through the mask aligner for 9s. Developer Microposit MF-321 was used to develop the resist for 40s, which was followed by a post bake at
115° for 180s. After development and before etching, the sample went through a 1 min 0₂ descum at 200mTorr flow, 200W power to remove any residual photoresist.

After patterning, the etching solution of 3:1 HCl: H₂O was prepared by adding 30mL of HCl into 10mL of H₂O and stirring for about 2 mins to stabilize the solution. The etch rate of InP in this solution is about 100nm/s. During this etching step, bubbles should appear on the reactive surface of the sample. InGaAsP and InGaAs acts as an

Figure 3.7: Fabrication steps for Iterations 2 & 3.
etch stop layer for this process [53]. To etch these layers a diluted H₂SO₄ solution is prepared by adding 10mL 95%H₂SO₄ in 100mL H₂O and then adding 10mL H₂O₂ into it. For this step the p-InP layer acts as an etch stop layer. The etch rate in this solution is about 1nm/s. After each step, the sample is rinsed with DI water very thoroughly. The reason wet etching is used to etch the n+ layers is because the good selectivity allows for some over etch to make sure there is no n+ layer residue which could create a leakage path in the channel and also to avoid any surface damage by a dry etch process. After etching is completed, the photoresist was stripped by solvents Acetone, IPA and water thoroughly. The height of the n+-p step was confirmed by Dektak 150 Stylus Profilometer or Veeco NT1100 Optical Profiling System.

After the wet etching step an oxide pad was patterned far away from the device to form a contact pad on it later (Figure 3.7(c)). SiO₂ was deposited using Oxford PECVD at 350°C. SiH₄ and N₂O mixture was used to form the SiO₂ on the sample. The total thickness of the insulator layer was 400nm with a refractive index of 1.45 at 635nm wavelength. The thickness was chosen such that the insulator does not breakdown at the bias of operation. The deposition rate was about 80nm/min. A blank substrate was placed in the chamber in addition to the actual sample. After deposition, the thickness was measured on that using Filmetrics F20 and it was also used to calibrate the etch rate in the etch solution. Positive photoresist Shipley S1805 was patterned on the real sample using the similar recipe used for the previous step of wet etching. Then, Buffered Oxide Etch (6:1 NH₄F:HF) was used to etch the oxide in the exposed area of the real sample. BOE etch cannot be indefinitely long for two reasons, first is the undercut due to the isotropic etch profile, second is that the Ti layer of the first metal layer will be attacked
by the BOE. While designing the photo mask, the undercut due to the wet etch process was already taken into account. After etching, the photoresist was stripped using solvents and afterwards, an O₂ descum at 200mTorr flow, 200W power was done for a few mintues. An overnight soak in acetone would help in cases, where the resist is very hard to remove.

Mesa Profiles were formed around the devices (Figure 3.7(d)) and the contact pads, to eliminate the area that is not part of the device, especially attention was given such that the non-active area around the devices were etched all the way through the low band gap InGaAs area, so that the dark current from the area does not affect the measured current. First, Shipley S1818 positive photoresist was patterned to obtain a resist thickness of about 1.5 microns. The resist was spun at 4500rpm for 40s, prebaked at 105°C for 3min, exposed in the mask aligner for 18s, and developed in Microposit MF-321 for 40s. The etching solution was prepared by mixing 4:1:25 HBr:H₂O₂:H₂O vigorously for several minutes. This solution is non selective and will etch all the III-V layers of the epitaxial layer with a very negligible difference in etch rate. The solution was used to etch a few dummy patterned samples to calibrate the etch rate. The etch rate will increase with stirring and after a while the rate will be stable, and that etch rate is used to etch the exposed area. Care was taken so that the solution was used within few hours of preparing it, because the solution expired and would not be useful after that. The samples were pasted in a microscope slide using photoresist for easier handling, and it was stirred vigorously in the etching solution to get a uniform and clean etching. For testing the etch rate, the step height for each dummy sample was measured using
Dektak 150 Stylus Profilometer. When the required step height from the real sample was achieved, the photoresist was stripped using solvents and O₂ descum.

After etching the III-V layers, lots of dangling bonds are exposed, and they had to be passivated before the surface changes too much and causes permanent leakage path through the mesa sidewalls. Therefore, a Polyimide layer PI-2610 was spun on the sample as soon as the etching was done. The PI-2610 was first diluted to obtain the desired thickness using the ration of 2:1 PI-2610:T9039. The solution after curing has to be thick enough, to cover the mesa sidewalls effectively without exposing the edge on the top of the island, but thin enough, so that, metal layer which will be deposited later to form contact with the device, can run over the PI layer continuously without breaking. This is critical because the PI profile is vertical after etching. On the real sample and a blank Si sample, PI-2610 was spun at 500rpm for 20s and 2000rpm for 55s, then prebaked at 90°C for 90s and 150°C for 90s. Then the PI layer was cured at 350°C for 30mins, using a ramp up time of 80min and a ramp down time of 33 min in the Ulvac MILA-3000 Minilamp annealer. The slow ramp rate was chosen to avoid any thermal shock in the material. After this the PI thickness in the dummy sample was measured using Filmetrics 20 and the thickness range was found to be in the order of 1microns. RIE O₂ plasma was used to etch the PI-2610. Conventionally, CF₄ is also used for faster etching of the material, however, we had to avoid this Fluoride based components to avoid oxide etching and III-V layer etching. The selectivity between photoresist and PI-2610 for O₂ is very poor. For that reason a SiO₂ hard mask was used to pattern. First, 90nm-100nm SiO₂ was sputtered using AJA RF Sputter Deposition Tool at room temperature. The process parameters used were, 300W forward power,
10sccm Ar flow, 10mTorr pressure, 50% stage rotation for about 1 hr. Again a blank Si sample was used to measure the thickness and calibrate the etch rate of the oxide layer. Negative photoresist NR9-3000PY was used to pattern the oxide hard mask. The resist was spun at 4500rpm for 40s, prebaked at 150°C for 2min. This bake time is longer than the suggested time in the datasheet to account for the fact that the curing is slower due to the thick insulator (PI-2610) layer. Then the resist is exposed through the mask aligner for 10s, and afterwards a post exposure baking is done for 80s at 100°C. Futurexx developer RD6 is used to develop the resist for 10s. All the baking time, exposure time and develop time might vary a little depending on the room temperature, humidity and lamp life of the mask aligner. After photoresist patterning, the oxide is etched using BOE or RIE etching. RIE etching is preferred due to the smaller feature size, but at the end a quick BOE etch is also done to clean the surface. The real sample and dummy sample with oxide is put in the Oxford Plasmalab 80 RIE chamber to etch SiO₂ using 25sccm Ar and 25sccm CHF₃ at 35mTorr, 200W. Once the oxide is etched and the PI is exposed, the sample was put back into the same chamber again. Due to absence of CF₄ in the etching of a 1micron thick layer, the total etch time required was very long. If the plasma was turned on continuously for a long time, what happened was that the PI would expand and break the oxide in the critical points like sharp corners or sharp edges. So, the oxide mask was damaged in those points and the PI was removed from those area. To avoid cracking, the plasma etching was performed for 1 min, and then a 1 min rest was give, when the gas flow was present, but the RIE power was 0W. During the etching steps, the conditions were 100W power, 100mTorr pressure, 40sccm O₂. After few cycles of etching, the samples were taken out under examined under the
optical microscope. The PI is very non uniform, and therefore all the openings did not clear at the same time. When the thickness of the PI was very small, there would be fringe like patterns in most of the openings and some opening were already clear by that time. That’s when the PI etching (Figure 3.7(e)) was almost complete. Calculating the etch rate till that point, the remaining PI was removed and a small over etch was also performed to make sure the surface of the opening was clean.

Using Dektak 150 the PI thickness on top of the contact pad and the device openings were measured. Since the electron beam evaporation is directional deposition, there is minimum side wall deposition. So to make continuous metal line from contact pad to the top of the wall, a thick metal comparable to the PI thickness needed to be deposited. If the PI sidewall was about 1um thick, at least 700nm metal thickness would be needed to get a high yield from the process in terms of connected devices. Sputtering would be an option to have good side wall coverage, but then lift off becomes practically impossible especially for small patterns. In electron beam evaporation, if 700nm was deposited in the same run, the chamber would become too hot and damage the resist. In that case, the lift off would be very difficult and sometimes no lift off was possible at all. For these reason the tedious path of patterning and depositing and lifting off multiple times was chosen. At each run, no more that 300nm was deposited. A thick PR NR9-3000 was spun at a slow speed to get an easy lift off. 100nmTi and 200nm Au was evaporated using Temescal BJD 1800 electron beam evaporator (Figure 3.7(f)). The deposition rate was very slow about 1A°/s to allow a very directional deposition by minimizing sputtering which is very important for a good lift off. Also the base pressure before starting the deposition was chosen to be low (<2e-7 Torr) to minimize scattering
for minimum sidewall deposition. The sample was rested in acetone for 5 mins, rinsed for 5 mins and then sprayed hard till most of the metal came off. If the lift off seemed hard, the sample was soaked in acetone overnight and lifted off next day. A very gentle rub of a cleanroom swab was used to remove pieces of metal films that were attaching by the end and won’t go by other means. To this point the devices (Figure 3.8) were taken ready to measure the dc current-voltage characteristics between the top 2 contacts. The I-Vs of several devices were measured to see if the second metal layer was connected to the devices, or still are open. If most of them seemed disconnected, another metal layer would be added by repeating the same procedure used above, and the thickness was determined by seeing the metal in the Scanning Electron Microscope. If there were some devices that showed normal I-V, we knew, only a thin metal layer would be sufficient. Attention was given while redoing the metal layer to maintain minimum misalignment.

Figure 3.8.: Device at 100X magnification after second metal layer deposition[43].
Once the yield was satisfactory, the next step was to define the apertures. The apertures were defined, to make the light hit the single photon sensitive area. Apertures were defined in different locations for different devices with the hope that one of the locations would be the most sensitive region. For this an insulator layer (Figure 3.7(g)) and a subsequent metal layer with aperture/opening (Figure 3.7(h)) was designed.

For the insulator layer, a photo definable polyimide HD4014 was used. This material was chosen over the other PI-2610 for simplicity because the pattern size is large, and HD-4104 is good enough for patterning large structures with 5micron resolution. The solution was spun at 1000rpm for 10s and 4500rpm for 30s and prebaked at 100°C for 3 min. Then it was exposed with the mask aligner for 22s, and after a rest period of 5 min, the patterns were developed in PA-401D for 60s by rinsing the sample in up and down direction. The sample was then rinsed in IPA for 40s and checked under the optical microscope to see if the pattern came out good. It is important not the rinse sample in water immediately after developing to avoid any residual layer forming on top of the sample. After this, the samples were cured in the oven for 30min at 200°C, and 60min at 375°C. The ramp up time was 18 min to 200°C, 18min from 200°C to 375°C and finally ramp down to room temperature in 36 min. The thickness of the film was then measured using Dektak 150 and was found to be about 5microns, which was thick enough to put a metal layer on top while not leaking to the second metal layer. After this step the device was ready to be metalized with the aperture on it.

NR9-3000 was spun at 2500 rpm for 40s for a very thick photoresist layer. It was then baked at 150°C for 2 mins, exposed for 9s, and post baked at 100°C for 90s. Afterwards,
it was developed for 22s to define the apertures. After a $O_2$ plasma descum, the device was put inside the electron beam evaporator for a 120nm Ti layer deposition. The deposition rate was about 1A/s. After lifting off the resist in solvents, a cleanroom swab was used gently to open the apertures, where the PR was harder to remove.

To make a backside contact Indium solder was pasted in the backside of the substrate (Figure 3.7(i)). A backside metallization could also be done as well for this purpose.

![Image](image.png)

Figure 3.9: Packaged device ready for testing.

Devices were tested initially through electrical probe measurements and Semiconductor Parameter Analyzer Agilent 4155B for I-V characteristics. Suitable devices with low dark current and good photo response were then chosen for wire bonding. The chip was cleaved into smaller pieces to fit into ceramic dual-in-line packages. The die backside was attached to the package (Figure 3.9) using conductive silver paste (Pelco conductive silver 187) and baked for 30mins at about 120°C to drive
out any of the solvents. Then, wire bonding was done with a West Bond ball bonder to attach gold wires from the on-die contact pads to the package. Then SMA connectors were attached to the pins by soldering so that the package can fit into a cryo-chamber for low temperature measurements.

3.3 Device Characterization

3.3.1 Experimental Setup

The setup including optical and electrical components is depicted in Figure 3.10.

![Figure 3.10: Experimental setup for sensitivity measurement [50].](image)

The packaged device is placed in a commercial cryochamber Janis VPF-100 for cooling control from as low as liquid nitrogen temperature cooling to as high as 400K temperature. A temperature controller Lakeshore 325, which has a PID controller for controlling a thermal diode was used. One thermal diode was placed near the heater
which will be controlled by the heater and the other was located closer to the actual sample to have a more accurate temperature reading.

To focus a small beam spot on the small aperture openings of the device, an imaging system was built which was placed on top of a fine control micrometer stage with 3-D linear control and rotational control. The challenge was to see the device and the beam spot on the same screen at the same time, to be able to align the focused beam spot with the device aperture. A 50X Mitutoyo M Plan Apo NIR Infinity-Corrected Objective (NA=0.42, working distance=17.0 mm) was used to focus the collimated light source to the device residing in the cryochamber. A pulsed laser Id Quantique id300 Fabry-Perot was used as the source of photons, which has a wavelength of 1550nm and pulse duration of 300ps, and a peak power of 1mW. The output was fiber coupled and connected to fiber mounted graded-index lens, the output of which was a collimated light beam.

A single plano-convex lens was placed 11 cm away from the objective. A NIR camera (Spiricon phosphor coated CCD focal plane array) was placed 13 cm away from the objective, laser source, and imaging lens. To illuminate the sample, a 10/90 beamsplitter was placed between a fiber bundled halogen light source and the optical path (Figure 3.10). A variable optical attenuator was placed in front of the laser source to adjust the number of photons per pulse and an optical power meter (Newport 1830-C) was used to calibrate that photon number.

Two dc source meters (Keithley 2400) were used to bias the center electrode and bottom electrode. The output from both are send to a digital oscilloscope Tektronix (tek 3000 series) with 500MHz bandwidth. Between the scope and the device output the
capacitive port of a bias-Tee (ps-5530A) was placed in series to block the dc component of the response to the oscilloscope. The inductive port of the bias tee was connected in series between the source meter and the device output. Agilent 81110A pattern generator was used to trigger the light source and the oscilloscope. So any response that was correlated with the trigger signal would be the response from the detector (Figure 3.11).

![Diagram](image)

Figure 3.11: Correlation between trigger signal, laser signal and output response.

### 3.3.2 Sensitivity Measurement

The number of photons per pulse is calculated using the equation. \( N_{ph} = \frac{I_{primary}}{q \times f} \)

where \( I_{primary} \) is the primary absorbed photo current, \( q \) is a unit electron charge, and \( f \) is the repetition rate of the laser. For a separate absorption–multiplication region
structure, there is a punch through voltage, at which the absorbed photo generated holes will reach the multiplication region and experience impact ionization. Below the punch through voltage, there is almost no electric field in the absorption region (Figure 3.12), and all the photo generated electron hole pairs get recombined before the holes reach the pin region and experience gain. As the voltage is increased, the electric field of the n region starts to penetrate into the absorption region, and now the photo generated holes will be pushed to the multiplication region. For this reason there is a sudden increase in the photocurrent at that point (green line) as can be seen in Figure 3.12.

Figure 3.12: Current-Voltage characteristics of an SAM structure depicting the punch through condition.
Total current measured from an APD is $I_t = I_d + M \times I_{\text{primary}}$

Here, $I_d$ is total dark current, $M$ is multiplication gain and $I_{\text{primary}}$ is primary absorbed photocurrent which depends on the source photocurrent and the total quantum efficiency. Right after the punch through voltage, where the photo generated holes reach the multiplication region, the gain $M$ is 1. So at punch through

$$I_{t,PT} = I_{d,PT} + I_{\text{primary}}$$

$$I_{\text{primary}} = I_{t,PT} - I_{d,PT} \tag{3.1}$$

Eqn. (3.1) implies that, the measured photocurrent, which is the difference of the total current and dark current at punch through voltage, is a measure of the absorbed primary photo current. So, from the measured I-V and a clear punch through point, one can get the primary absorbed photocurrent and calculate the absorbed number of photons. From this procedure we know, how many photons are hitting the device. For sensitivity measurement, the next task is to find out what is the minimum photon number that gives a detectable output pulse. For that, we used the setup explained in the previous sub section and recorded the output pulse for each photon number. The lowest photon number that gave us a detectable pulse was our ultimate sensitivity for each iteration.

### 3.3.3 Result Analyses for iterations

This section will present and discuss the experimental results obtained by the methods explained in the previous sections. For the first iteration, only dc measurements were done. The leakage current level was very high with that process iteration and that led us to modify the process for the iteration 2. Iteration 2 gave us a better understanding
since we did complete sensitivity measurement with those devices and that knowledge helped us design another epilayer (Design 2) for improved sensitivity. The process flow was same as the iteration 2 for this case.

3.3.3.1 Iteration 1

The dc current-voltage characteristics were measured for the devices fabricated from iteration 1 using the commercial parameter analyzer Agilent 4155B. The DUT (Device Under Test) was cooled down due to the high dark current and analyzed for different photon number inputs and also dark condition. Figure 3.13 shows the I-V measured from the bottom electrode, in other words the electrode that is used to control the avalanche gain process. The punch through voltage is evident near 27V which is due to the electric field penetration into the absorption region.

Figure 3.13: Bottom Electrode I-V. Measurement temperature, $T = 180K$, $V_{\text{center}} = 0.9V$. 
Figure 3.14 plots the center electrode I-V at the same conditions as above, and the current level found at only a small bias of 0.9V is very high. The leakage current is in the order of 1.34mA in this case, which is indicative of a turned on channel rather than an npn structure. The background current being so high aids the flow of accumulated electrons and therefore the photocurrent is also of the same order. However, due to large shot noise from high current the sensitivity of these devices are very far from the targeted sensitivity.

Figure 3.14: Center Electrode I-V. Measurement temperature, $T = 180K$, $V_{\text{center}} = 0.9V$.

Figure 3.15 and Figure 3.16 plots the dc gain derived from the dc photocurrent and the primary photocurrent. Comparing the two gain plots, we see that the total gain has a large factor over the coupled gain, and the excess gain is believed to come from the FET gain mechanism. As explained above the already present electron current aids
the flow of new coming electrons and results in a high transistor gain. However, this gain is not useful because it does not give us the sensitivity. To get the true sensitivity from the integrated gain mechanism, the leakage current needs to be cut down by several orders of magnitude.

Figure 3.15: Coupled Avalanche-Bipolar gain. Measurement temperature, $T = 180K$, $V_{center} = 0.9V$. 
Figure 3.16: Total gain for different photon number inputs. Measurement temperature, $T = 180K$, $V_{center} = 0.9V$.

Figure 3.17: Simulated structure for examining the effect of interface traps in the channel current.
The structure in Figure 3.17 was simulated using commercial software Silvaco Atlas. The channel current was observed with and without interface traps for the npn structure.

![Conduction band diagram along the channel with and without interface states at a low gate or bottom electrode bias. Acceptor trap density, \( N_a = 1e14 \text{cm}^{-2} \).](image)

From Figure 3.18 it can be seen that for an ideal npn structure there is a potential barrier between the electron source and the channel material, whereas for the cases with interface traps the barrier is absent and that means electrons from the source can easily transport to the channel material and cause a significant leakage current at low gate voltage. This is believed to be the source of the high leakage current from iteration 1 devices. In our real structure there is a PECVD oxide layer deposited on the channel. This step being a high temperature process can create trapped charges in the insulator layer [5], and these can potentially be the source of the non-ideal surface condition. To examine the case, the oxide from the surface was etched away using BOE solution and a dramatic decrease in the channel current was observed. However, since etching the oxide removed some oxide from below the contact pads and some oxide which was used
for isolation, the device was not stable any longer and therefore a new iteration with modified process flow was required.

3.3.3.2 Iteration 2

Iteration 2 was designed with a slightly different process flow with the same epilayer. Instead of protecting the channel surface with oxide which has trapped charges from the plasma enhanced deposition process, this time the mask was designed in a way that polyimide will be deposited on the channel during the passivation step, and this layer will act to isolate the center and outer electrodes. The process flow is shown in section 3.2.2.

To examine the leakage current three npn test structures were designed (Figure 3.19) which were generated during the actual device process steps. One structure had oxide layer in the interface which was deposited during the oxide contact pad deposition step. Another test structure did not have any insulator layer and the third structure had Polyimide layer which was deposited during the mesa sidewall passivation step. Drain voltage of 0 to 3 Volts was applied across the structures, and the current was measured using Agilent 4155B at room temperature.
Figure 3.19: npn test structures for testing leakage current by applying a voltage between the two electrodes.

![Figure 3.19: npn test structures for testing leakage current by applying a voltage between the two electrodes.](image1)

Figure 3.20: Leakage current from n-p-n test structures with different channel interfaces.

![Figure 3.20: Leakage current from n-p-n test structures with different channel interfaces.](image2)

From Figure 3.20, it can be seen that the structure with oxide has at least 3 orders of magnitude larger leakage current than the other two structures. With no insulator and Polyimide insulator the leakage current is of the same order, which indicated that
Polyimide is a good material for the electrode isolator layer and also the passivation layer or protection layer for the channel.

![Graph](image)

Figure 3.21: Coupled Avalanche-Bipolar gain for 6250 photons from iteration 2.

Figure 3.21 and Figure 3.22 plots the coupled avalanche-bipolar gain and the total gain for iteration 1, which demonstrates the successful integration of the three gain mechanisms. The gain values in these figures were dc gain values obtained from the dc photocurrent from each electrode. The FET gain is derived from the ratio of the total gain and the coupled gain and is plotted in Figure 3.23.
The sensitivity of this iteration devices are about 5000 photons per pulse at a low temperature of 155K. Although a high total gain was demonstrated, the distribution of the different gain components were not favorable for a high sensitivity. Figure 3.21 clearly shows a low coupled gain, meaning an even lower pure avalanche gain for these
devices. Such a small avalanche gain is not sufficient to produce a reasonable trigger for the other gain mechanisms to pick up. Even though the FET gain is significant but that alone cannot produce a high sensitivity.

Figure 3.24: Analog response for 10,000 photons per pulse at $= 155K$.

The reason behind the low avalanche gain is the too strong negative feedback, which kept the multiplication region electric field below the required level at a given applied voltage. For large number of photons (>5000), the total hole accumulation was high and they pulled electrons from the emitter and that way a high overall gain was developed. However, for low photon numbers (<5000), the accumulated holes were lower and not sufficient to trigger bipolar gain and FET mechanism and therefore remained undetected. Since the avalanche gain is bias dependent, one can argue that increasing the bottom electrode bias can improve the avalanche gain and therefore the sensitivity. This method works up to a certain bias. When the bias becomes too high,
the device enters an unstable mode and starts oscillation due to the coupled bipolar-avalanche mechanism. This is because the bipolar gain or electron emission is too strong for this design. Figure 2.4 which was based on the analytical model, showed a stability chart for iteration 3 devices, where one can see that for a higher bipolar gain, the devices reach the unstable mode at a lower bias. Once the unstable mode is triggered, the output response would start oscillating as shown in Figure 3.25. In this mode, the device loses the power to respond to any dark or photon signals.

Figure 3.25: Oscillatory behavior at high applied bias for iteration 2 devices.

3.3.3.3 Iteration 3

To further improve the sensitivity, a new epitaxial layer was designed (Design-2), where the valence band offset between the p-InP and the TCB layer was decreased to reduce the strength of the negative feedback mechanism. This would result in a larger avalanche gain but a lower FET gain. The FET gain will lowered due to the faster
hole escape time which means smaller electron lifetime in the channel. Also, to reduce
the electron emission strength, another electron barrier was added by inserting a thin
InGaAs layer (Table 3.2) in the emitter region. In the revised design, the avalanche gain
should increase and the bipolar and FET gain should decrease and this should increase
the sensitivity of the devices. The dc I-Vs in Figure 3.26 and 3.27 indicate a successful
device fabrication with a clear punch through voltage.

Figure 3.26: DC current-voltage characteristics of bottom electrode.
Figure 3.27: DC current-voltage characteristics of center electrode.

Figure 3.28: Peak Current distribution for different number of photons per pulse [43].
To obtain the sensitivity of these devices, the setup explained in section 3.3.1 was used. A minimum of 10 photons per pulse was detectable from this device. Any photon signal with the dark signal. Figure 3.28 shows the distribution of output signals under different average photon numbers in the input pulse. Each curve can be fitted to a Gaussian function and the variation of the Gaussian distribution represents the overall noise. Figure 3.29 shows the bit-error-ratio (BER) calculated from the Gaussian fit in Figure 3.30. The results indicate that at 200K for reduced dark current, the device can achieve an error ratio of $10^{-2}$ with 10 photons per signal. Under very low photon numbers, the $BER$ from an “ideal noise free photoreceiver” can be given by the simple expression, $BER = 0.5 \exp(-N)$, where $N$ is the average photon number in each “1” bit[14]. The number $N$ is 4 photons for a BER of $10^{-2}$. Therefore, our device is only 6 photons away from the true noise-free photo receiver limited only by photon statistic fluctuations.

![Figure 3.29: Bit Error Rate for different photon numbers [43].](image)
Figure 3.30 and 3.31 plot the measured gain values for different number of photons per pulse for the device operating at 200K. The Avalanche-Bipolar gain was obtained from the bottom electrode photocurrent entering the bottom electrode of the APD. The net photocurrent, including the contribution of the FET amplifier, was obtained from the current leaving the center electrode and from that the total gain was derived. Figure 3.30 and 3.31 experimentally manifest the presence of FET gain coupled with the avalanche-bipolar gain.

![Figure 3.30: Coupled Avalanche Bipolar gain from iteration 3 device [43].](image)

When photon number within the signal increases, the avalanche-bipolar gain decreases due to negative feedback effect from the earlier arriving carriers. However, the output signal intensity increases monotonically and consistently with the intensity of input signal in a power law similar to the response of CMOS image sensors, thus producing a decent dynamic range.
This chapter presented the detail of the device implementation including the epitaxial design and fabrication steps for the different iterations. The experimental setup was described with the detail electrical and optical arrangements and the procedure of measurements. Finally the results from each iteration was presented and the reasoning for the transition to the next iterations were clarified.

This chapter, in part, is a reprint of material as it appears in the following publications:

Samia Nawar Rahman, David Hall, Zhe Mei and Yu-Hwa Lo, “Integrated 1550nm photoreceiver with built-in amplification and feedback mechanisms”, Optics Letters 38, 4166 (2013). The dissertation author was the primary investigator and author of this material.
Samia Nawar Rahman, David Hall, Zhe Mei and Yu-Hwa Lo, “Negative feedback and multiple gain mechanisms for sensitivity improvement in 1550nm optical detection”, IEEE Photonics Conference (IPC), 364 (2013). The dissertation author was the primary investigator and author of this material.
Chapter 4

Unified Device Modeling

In this chapter, a Monte Carlo simulation is presented which is based on the physical model of the invented device. The objective of the simulation is to analyze the device’s fundamental characteristics and also compare the results with conventional SPADs with different quenching mechanisms (passive quenching and self-quenching). Our model takes into account all three stochastic processes that control the fluctuations of photocurrent of the proposed device, and the model for different types of SPADs may be considered as a special case of the comprehensive model. By comparing the proposed device with passively quenched and self-quenched SPADs, the invented device shows a significant increase in detection efficiency, orders of magnitude lower error rates for single photon detection, and much greater dynamic range, attributed to the integration of avalanche multiplication and bipolar amplification as two coupled gain mechanisms in conjunction with the self-quenching feedback to regulate the output signal. The model justifies that device concept opens up a new avenue for single photon detection in Non-Geiger mode operation.
4.1 Motivation

The previous model discussed in Chapter 2 demonstrated the integration of bipolar gain mechanism with an APD. By using deterministic rate equations, the current response to a single carrier was shown. The model included experimental input parameters such as applied bias and operating temperature, device design parameters such as barrier height and epi-layer thicknesses, along with material parameters such as impact ionization rates and effective masses. However, this model does not take into account the stochastic nature of the avalanche multiplication process, hole escape mechanism and electron emission process. To learn the true sensitivity, the nature of noise or current fluctuations is an important piece of information. The deterministic model shows that the gain of the device can be high below Geiger mode. However, that is only useful if the excess noise or the gain fluctuation is under control. Therefore, this Chapter takes the effort to do the Monte Carlo simulations with all the necessary stochastic mechanisms included which is important for exploring the true sensitivity.

4.2 Model Algorithm

To simulate the avalanche buildup, a 1-D random path length model for carriers is used [54]. The ionization rates for the carriers are based on the local field approximation, with an effective dead space effect based on a hard ionization threshold [55]. The schematic structure used for the Monte Carlo model is shown in Figure 4.1. In the one dimensional structure, the coordinate extends from $x = 0$ at the absorption
region-pin region interface towards the top. Figure 4.1 is the structure for the complete modeling of a MAGIC detector. However, if we take off the emitter region, the structure becomes a self-quenched APD structure which is one of the cases we will study.

Figure 4.1: MAGIC detector schematic structure with InGaAs/InP material. $\Delta E_c$ and $\Delta E_v$ are the conduction band and valence band offsets at the material interfaces. $V_b$ is the applied voltage bias. The coordinates in the Figure correspond to the coordinates in the device model.

In the model the multiplication region extends from $x = 0$ to $x = W$. Holes drift in the $x$ direction and are collected at the hole energy barrier at $x = W$, while electrons drift in the $-x$ direction and are removed from the simulation at $x = 0$. Holes are initially injected from the absorption region into the multiplication region at $x = 0$. At every time step $dt$, each carrier in the multiplication region moves by $v_{sat} dt$, where $v_{sat}$ is the saturation velocity of the carrier. Based on the random path length model
[54], if the total distance the carrier travels causes its cumulative ionization probability, \( F_{c(e,h)} \) (eqn. 4.1) to be larger than its “luck,” it will create another electron hole pair at that location, new lucky numbers (uniformly distributed between 0 and 1) will be rolled for the carriers and their distance traveled will be reset.

\[
F_{c(e,h)} = 1 - \exp(-\alpha_{e,h}(x - x_0 - d_{e,h}))
\]  

(4.1)

Here, \( \alpha_{e,h} \) is the ionization coefficient of electron or hole, \( x - x_0 \) is the distance traveled by the carrier and \( d_{e,h} \) is the dead space of the carrier given by the threshold ionization energy divided by the local electric field [55], [56].

For self-quenching analysis, we need to calculate the time-dependent field distribution in the multiplication region. After the carrier locations are updated according to the random path model, any holes that travel past the barrier interface at \( x = W \) will be added to the sheet charge concentration \( \sigma \). To relate \( \sigma \) to the multiplication voltage \( V_m \) when the device is under a constant bias \( V_b \), we need to calculate the voltage drop \( V_{TCB} \) across the InAlGaAs/InAlAs transient carrier barrier (TCB) layer and then update \( V_m \) each time when \( dt \) passes:

\[
V_{TCB} = E_\sigma L = \frac{q\sigma}{\varepsilon L}
\]  

(4.2)

\[
V_m(t + dt) = V_m(0) - V_{TCB}
\]  

(4.3)

Where \( V_m(0) \) is the voltage drop in the multiplication region at \( \sigma = 0 \), \( E_\sigma \) is the field drop due to the accumulated charges, and \( L \) is the thickness of the TCB layer. The initial \( V_m \) is given by:

\[
V_m(0) = V_b
\]  

(4.4)

After each time interval \( dt \), new holes will be arriving at the barrier and
depending on the hole escape statistics, some accumulated holes will escape the barrier, which will give us an updated $\sigma$. This will result in a change of the electric field distribution according to eq. (2) and eq. (3). The hole escape, which is a stochastic mechanism, has been treated in a similar fashion as the avalanche multiplication. Once each hole arrives the barrier, it is assigned a random lucky number (uniformly distributed between 0 and 1). At each $dt$, if the lucky number is smaller than the cumulative escape function $F_{he}$ (eqn. 4.5), the hole escapes and otherwise remains in the barrier. The cumulative hole escape function is given by,

$$F_{he} = 1 - \exp\left(-\frac{(t-t_0)}{\tau}\right)$$  \hspace{1cm} (4.5)

Where, $t - t_0$ is the duration of the carrier in the barrier and $\tau$ is the average hole escape time determined by the barrier height $\Delta E_v$.

For bipolar gain mechanism, we have taken into account the emission statistics of the electrons reaching the multiplication region. For a BJT the mean emitter current [57] is a function of the voltage drop across the TCB layer, the conduction band offset, $q\Delta E_c$, carrier velocity $v$, emitter area, $S$, and emitter doping $N_n$.

$$I_{mean} = q \, v \, N_n \, S \, \exp\left(-\frac{q\Delta E_c}{kT}\right) \times \exp\left(\frac{qV_{TCB}}{kT}\right)$$  \hspace{1cm} (4.6)

The increase in the accumulated hole concentration will result in an increase in $V_{TCB}$ (eqn. 4.2), which will induce more electrons from the emitter according to eq. (6). A stronger $\Delta E_c$ means electrons face a stronger barrier towards the multiplication region and a reduced emitter current. Both the TCB voltage drop, $V_{TCB}$ and the band offset $\Delta E_c$ are the controlling parameters for the emitted electron current.

After the hole accumulation is updated from the hole escape statistics, the
updated $V_{TCB}$ is used to calculate the mean emitter current, $I_{mean}$ for that instant. The mean number of electrons emitted from the emitter over a time interval $dt$ can be calculated as

$$n_e = \frac{I_{mean} \cdot dt}{q} \quad (4.7)$$

A Poisson distribution is generated with a mean of $n_e$. A dice is rolled to sample a number from that distribution at that instant and the sampled electrons $n'_e$ will be added to the interface between the multiplication region and the TCB region (i.e. at $x = W$). This becomes the new source for further impact ionization process.

![Diagram](image)

Figure 4.2: Passive quenched SPAD [34]. A series resistor $R$, is connected with the APD to quench the single carrier response from avalanche multiplication.

For a passive quenched device (Figure 4.2), every time step, $dt$ the voltage drop across the resistor, $R$ will be deducted from the total applied voltage to obtain the multiplication region voltage, which will control the impact ionization rate and
quenching in the next time step.

\[ V_m(t + dt) = V_m(0) - V_R, \quad V_m(0) = V_b \]  
(4.8)

Figure 4.3: Monte Carlo simulation flow chart for passively quenched APDs, self-quenched APDs, and the MAGIC detector.
With the updated $V_m$, and current calculation the simulations will be repeated by a time step of $dt=0.1\text{ps}$ till no carriers are left in the multiplication region. The complete simulation produces a current response to each photon and this is repeated numerous times to obtain the distribution. The process flow for different mechanisms is depicted in Figure 4.3.

4.3 Simulation Results

The matlab model file for a single instance is attached in Appendix A. The first two pages of the .m file lists all the experimental, device and material parameters and simulation conditions needed to run a certain condition and is explained in a comment next to each input. Note that, the simulation in the sample file is for InP material, but it can be easily changed to other material systems by changing the material property related inputs.

In the simulations, the operating temperature $T$, multiplication width $W$, and TCB length $L$ are kept constant. The temperature is kept at 200 K for all three types of devices. A single photon is assumed to be absorbed and produce an electron-hole pair at $x = 0$ at $t = 0$. For self-quenched SPADs and the proposed device, the multiplication width and the transient carrier buffer (TCB) layer are assumed to be 0.8 $\mu$m and 0.4 $\mu$m, respectively. Typical ionization parameters for InP from [55] are used. For the proposed device, all three stochastic processes: avalanche multiplication, carrier escape, and electron emission, are employed. The applied voltage is set to be at 5% below the breakdown voltage, $V_{br}$, whereas the bias voltage for passively quenched and self-
quenched SPADs is 4% above breakdown. The breakdown voltage $V_{br}$ is obtained by solving the denominator of multiplication gain, $M$ in eq. (4.9) [6] to be zero so that the DC gain becomes infinite. Here $k$ is the ratio of hole and electron ionization coefficients.

$$M = \frac{1-k}{\exp(-\delta(1-k))-k}$$  \hspace{1cm} (4.9)

$$k = \frac{\alpha_h}{\alpha_e}, \quad \delta = \int_0^W \alpha_e dx$$  \hspace{1cm} (4.10)

To simulate self-quenched SPADs, the bipolar injection is turned off ('BJT=0'), but the avalanche statistics and the hole escape statistics are kept on. The devices are operated at 4% above $V_{br}$ since below breakdown, the gain is not sufficient to produce sufficient current output to detect single photons. The simulation of passively quenched SPADs includes only avalanche statistics and the devices are biased at 4% above breakdown as well for the same reason.

Figure 4.4 shows the simulated distributions of photon current response to single photons for three different device designs.
By repeating the simulation for single photons 500 times, the single-photon current distribution for each device structure was obtained. From Figure 4.4, the passively quenched device has the widest distribution, resulted from the increasing excess noise. The increasing excess noise is due to the absence of an effective feedback mechanism that carries information of cumulative carrier history (i.e. lacking an integration over the time lapse since the avalanche is triggered). With the self-quenching mechanism, the current fluctuation is much better controlled compared to the passively quenched SPADs, so the distribution of output signals is narrowed significantly. The better noise performance is primarily due to the internal negative feedback mechanism that regulates the gain, and such effect has been demonstrated in [20]. With the MAGIC detector we clearly observe substantial improvements in the overall gain manifested by
the mean value of the output current, and the greatly suppressed noise represented by
the much tighter distribution of the output current. The underpinning physics for such
improvements is (a) the split and coupling of two amplification mechanisms (i.e.
average multiplication and bipolar gain) so that neither mechanism has to be pushed
to the excessive value and suffer from the associated excess noise, and (b) the presence
of the internal feedback mechanism to regulate the overall gain. Note that although the
coupling between the avalanche gain and bipolar gain can be viewed as a positive
feedback since increasing one will cause the increase of the other, thus allowing for a
highly efficient buildup of a strong response to a single photon, the internal overall gain
control mechanism, most effectively implemented by the band discontinuities of
heterojunctions, regulates the device response.

Table 4.1: SPDE and BER of different technologies.

<table>
<thead>
<tr>
<th>Device type</th>
<th>SPDE(%)</th>
<th>BER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passive quenched APD</td>
<td>48.6</td>
<td>0.3137</td>
</tr>
<tr>
<td>Self-quenched APD</td>
<td>66</td>
<td>0.0247</td>
</tr>
<tr>
<td>MAGIC detector</td>
<td>98</td>
<td>1e-7</td>
</tr>
</tbody>
</table>

Table 4.1 summarizes the single photon detection efficiency (SPDE) and the bit
error ratio (BER) of single-photon signals for passively and self-quenched SPADs and
the MAGIC detector. The threshold current was set at 250nA. To calculate the SPDE
and BER, we have optimized the bias conditions of each device individually, having the
two SPADs biased above the breakdown voltage and the MAGIC detector biased below
breakdown.

The BER [14] was calculated based on the data in Figure 4.4 but fitted to
appropriate distribution functions. The error probability can be calculated as
\[ BER = \frac{1}{2} \text{erfc} \left( \frac{\mu_1 - \mu_0}{\sigma_1 + \sigma_0} \right) \] (4.11)

Here \( \mu_1, \mu_0, \sigma_1 \) and \( \sigma_0 \) are the signal amplitudes and spreads in light and dark conditions. The results indicate that the MAGIC detector outperforms SPADs significantly regarding both the single-photon detection efficiency and the bit error ratio for single photon signals. The results confirm the physical arguments underlining the MAGIC detector discussed earlier.

To examine the dynamic range and linearity of MAGIC detector, Monte Carlo simulations have been performed with different number of photons per pulse. Figure 4.5 shows the mean current obtained for different photon input numbers. Neither passively quenched nor self-quenched SPAD shows any appreciable dependence of its mean output current on the input photon number. In sharp contrast, the output signal of the MAGIC is approximately linearly proportional to the input photon number, demonstrating a substantial level of dynamic range that is lacking in all known Geiger mode devices.
4.4 Conclusion

This chapter developed and analyzed a Monte Carlo model to study the statistical noise of the different device technologies and it supported the concept and arguments made in Chapter 2 and 3. The model is easily adjustable to the analysis of devices with all or any of the mechanisms presented with any material system.

This chapter, in part, is a reprint of material as it is submitted in the following Journal:

Samia Nawar Rahman, David Hall and Yu-Hwa Lo, “Non-Geiger mode single photon detector with multiple amplification and gain control mechanisms", Journal of Applied Physics. The dissertation author was the primary investigator and author of this material.
Chapter 5

Conclusions

This chapter will provide brief summary and outlook on the material presented in the dissertation.

5.1 Thesis Summary

The thesis presented here demonstrates a highly sensitive photo detector technology namely “MAGIC Detector” which is advantageous over the conventional APD and SPAD devices in many ways. Design, fabrication and measurements of the device prototype have several key contributions which are summarized below.

1. Design & fabrication of MAGIC detector devices through addition of a TCB barrier and an electron emitter internally in the epitaxial layer of an APD.
2. Physical understanding of the coupled gain mechanisms and the negative feedback mechanism and their contributions to the sensitivity and dynamic range of the detector. These understandings were demonstrated by the implementation of new iterations and a deterministic analytical model.
3. Sensitivity measurement technique design for analysis of the device behavior.
4. A Unified Device Model development to understand the current and gain fluctuation of the MAGIC detector devices.
The near infrared detector presented in this dissertation has been able to provide a sensitivity which is 50-100 times better than a conventional linear mode APD. The reason is that the total gain of the device was high even though the operation voltage was below breakdown. The integration of two more gain mechanisms namely, bipolar gain and FET gain allowed the use of a smaller avalanche gain compared to the SPAD devices. A negative feedback mechanism was implemented to take care of the current and gain fluctuation.

An analytical model was developed which demonstrated the coupled avalanche-bipolar gain mechanism.

Three iterations of fabrication runs with two epitaxial designs were executed to learn the device physics and design rules. A weak hole barrier resulted in higher avalanche gain, lower FET gain and weaker negative feedback, whereas a strong emitter resulted in a higher bipolar gain and a very strong emitter could push the device to unstable mode. By optimizing the TCB layer design and the electron emitter design, the balance between the three gain mechanisms can be obtained.

The Unified Device Model showed a clear improvement in excess noise by implementing the negative feedback mechanism in the self-quenched SPAD devices. The MAGIC detector which also has that mechanism showed a tight distribution and a greater signal strength because of the higher total gain, and above all, a significantly improved single photon detection efficiency approaching 100%. 
5.2 Outlook

Experimentally, ten photons were detected from the 3rd device iteration implemented in InP based design. By analyzing the devices, we learnt the qualitative design trends of the detectors for further sensitivity improvements. From design 1 to design 2, the avalanche gain was made stronger and the electron emission was made weaker, and that helped the sensitivity to improve by orders of magnitudes. The next required improvement from ten photons to one photon for near infrared light detection needs smaller change in the same direction.

One major advantage of the negative feedback or self-quenching mechanism is that implementing a large array of single photon detectors is practically possible. Large array of SPADs for visible light detection is not much commercially available yet. Implementing the concept of MAGIC detector for Si based devices [58] for visible light detection can be a promising solution for CMOS compatible 3D image sensors [59], [60]. By band gap engineering of Si-SiGe heterostructure layers for integrated gain and feedback mechanisms, single photon detection from a silicon integrated circuit should be feasible.
References


Appendix A: Matlab Code for Unified Device Model

% Monte Carlo simulation of APD
% v1 2013-09 Samia
% Using matlab 7.8.0 (R2012)

function [Vm, Nfinished, spde, jitter, gain, peak, width] = MCSPAD()

% experimental parameters %
BJT=0;                % BJT='1'--> Bipolar gain on, '0'--> Bipolar gain off
Voverp =5;            % [%], applied overbias, as percentage
Nn=1e20;             % [cm^-3] Emitter doping
Tp=100e-9;            % [s] Hole escape time
Ni=10;                % number of photons per pulse [1, 2, 3,……]
Temperature = 240;    % [K], lattice temperature
T=Temperature;
input_th = 1e-8;        % [A] Current threshold to determine if input is detected
Nsweeps = 1000;        % number of trials to run

% device and physical parameters %
Ec=0.4;                % [eV], conduction band offset
Wm =1e-4;             % [cm], width of multiplication region
Wtcb = 0.45e-4;       % [cm], width of multiplication region
S = 1e-7;             % [cm^2], device effective area (capacitance)
Edelta = 4.30e5;       % [V/cm], Efield change due to n-charge layer ~ q*nDoping*Wc/eps

dEc = Ec;            % [eV], conduction band offset
vs = 1e7;            % [cm/s], saturation velocity

Eie = 2.05;            % [eV], threshold energy for impact ionization of electrons % Saleh 2001
Eih = 2.20;            % [eV], threshold energy for impact ionization of holes
% *** ionization coefficients [cm^-1] [cm/V] []
a0 = 3.01e6; cn = 2.45e6; mn = 1.08; b0 = 4.29e6; cp = 2.08e6; mp = 1.12;

q = 1.6e-19;
eps0 = 8.854e-14; % [F/cm]
eps = 12.4*eps0;
me = 0.08*(9.109e-31); % [kg], electron effective mass
h = 4.136e-15; % [eV*s], Planck's constant
hbar = h/2/pi; % [eV*s], = 6.582e-16

k = 8.61734e-5; % [eV/K], Boltzmann constant
kT = k*Temperature; % [eV]
vnb = 1e7;
Nc = 2*(2*pi*me*kT/h^2/q/1e4)^1.5; % [/cm^3], = 5.737e17 @ 300K, me=0.08

% numerical parameters %
dt = 0.1e-12; % [s], calculation step size (0.1ps)
et = 1000e-12; % [s], simulation end time (max); may stop earlier if no more carriers
Z = 6.0e7;%1.5e4 % working array size, e.g. no. of carriers in m-region must be less than this
Glim = 1e9; % gain limit: max number of carriers to calculate per avalanche pulse
mat = 1e-12; %[s], moving average time, to smooth out current response
mas = round((mat/dt-1)/2)*2+1; % moving average span, based off above. must be odd.

DEBUG = 0;
DEBUG_GFX = 0;
K=1.38e-23;
tic;

% %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% ***** Equations %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function y = M_denom(V) % denominator term of avalanche gain eqn.
Fm = V/Wm;
alpha = a0*exp(-(cn/Fm)^mn);
beta = b0*exp(-(cp/Fm)^mp);
k = beta/alpha;
y = exp(Wm*(beta-alpha)) - k;
end

VmBr = fzero(@M_denom, 35); % find breakdown voltage across multiplication region
VBr = VmBr/Wm*(Wm+Wtcb)-Edelta*Wtcb; % total voltage on device at breakdown
V = VBr*(1+Voverp/100); % set overbias voltage

Fm0 = (V+Edelta*Wtcb)/(Wm+Wtcb); % e-field in multiplication when there is no charge trapped
% {rearrange V=Em*Wm+(Em-Ed)*Wtcb}
Vm0 = Fm0*Wm; % initial voltage in multiplication region
Ftcb0 = Fm0-Edelta;
if Ftcb0<0, display('Error: Edelta too large');return;
end

function y=Etcb(Em) % Gives e-field in TCB given e-field in multiplication
y = (V-Em*Wm)/Wtcb;
end

% %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% ***** Setup %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

finished = false(1,Nsweeps); % did avalanche simulation finish?
pde = false(1,Nsweeps); % detection efficiency:: (I>input_threshold)
jitter = zeros(1,Nsweeps); % time to reach input_threshold
gain = zeros(1,Nsweeps); % avalanche gain
peak = zeros(1,Nsweeps); % avalanche pulse peak
width = zeros(1,Nsweeps); % avalanche pulse width
nWs = zeros(1,Nsweeps); % nW at end of each sweep
inic; % simulation runtime timer

for sweep = 1:Nsweeps

% ******************************************
% ***** Initialize Device ****************************************
% ******************************************

Vm = Vm0; % reset Vm
Ae = false(1,Z); % carrier type: 0:nothing, 1:electron
Aex = zeros(1,Z); % pos: W:tcb/M interface, W: hole injection
Aex0 = zeros(1,Z); % initial injection location

Ah = false(1,Z); % carrier type: 0:nothing, 1:hole
Ahx = zeros(1,Z); % pos: W:tcb/M interface, W: hole injection
Ahx0 = zeros(1,Z); % initial injection location

Ah(1:Ni) = 1; % hole injection @ x=0
Ahx(1:Ni) = 0;
Ahx0(1:Ni) = Ahx(1);

Re = rand(1,Z); % random number: ionization probability
Rh = rand(1,Z);
Be = false(1,Z); % used to store temp boolean comparisons
Bh = false(1,Z);
Ne=false(1,Z);

tx = zeros(1,Z);
tx0 = zeros(1,Z);

G = 1; % total number of generated carriers, eg. gain
fe = 1;  % index of first electron still in Mregion
ge = 1;  % index of last electron still in Mregion
fh = 1;  % index of first hole still in Mregion
gh = Ni; % index of last hole still in Mregion
f=1;
g=1;
eti = floor(et/dt); % simulation end time, as index
t = 1; % time index
Mt=zeros(1,eti);  % m(t)
Gt=zeros(1,eti);  % G(t)
Vmt=zeros(1,eti); % Vm(t)
J=zeros(1,eti);  % hole current leaving x=0
nW = 0; % number of electrons at TCB interface (x=W)
nWt=zeros(1,eti); % nW(t)
Jtt=zeros(1,eti); % Jt(t)
Oh=false(1,Z);
Th=false(1,Z);
Rb=rand(1,Z);
Ch=zeros(1,Z);
Dis=zeros(1,eti);
Di=zeros(1,eti);
Ie_mean=zeros(1,eti);
ght=zeros(1,eti);
get=zeros(1,eti);
Oht=zeros(1,eti);
Bht=zeros(1,eti);
Bet=zeros(1,eti);

success = true;  % whether avalanche is completely simulated successfully

% ********************
% ***************************************
% ***** start of time calculation ****************************
% *****************************
% ****************************

for t = 1:eti

% -update Vm
Vm = Vm0 - nW*q/(eps/Wtcb*S);
if Vm < 10, display('crit Vm'), success=false; break; end
Vmt(t)=nW*q/(eps/Wtcb*S);

% -update electric field
Fm = Vm/Wm ;  % multiplication electric field
alpha = a0*exp(-(cn/Fm)^mn);
beta  = b0*exp(-(cp/Fm)^mp);
de = Eie/Fm;  % [cm] ionization dead space for electrons
dh = Eih/Fm;  % [cm] ionization dead space for holes

tx=tx+dt;
Aex = Aex - vs * dt;  % update position of e/h
Ahx = Ahx + vs * dt;

% -check impact ionization: correct carrier exists in Mregion, & < CDF

Be(fe:ge) = (Aex(fe:ge)>0) & ((1-exp(-alpha*(Aex0(fe:ge)-Aex(fe:ge)-de))) > Re(fe:ge));
nne = \text{sum}(Be) ; \quad \% \text{total Number of New ionization events generated by Electrons for this dt}
nnh = \text{sum}(Bh);

\text{if } ((gh+nne+nnh) > Z) \text{ || } ((ge+nne+nnh) > Z), \text{display('out of bounds error'), success=false; break; end}

\%if ((G+nne+nnh) > Glim), display('gain limit reached'), success=false; break; end

\% -create child carriers if any ionization events occurred
\text{if } (nne > 0)

\begin{align*}
Ae(ge+1:ge+nne) &= \text{ones}(1,nne) ; \quad \% \text{as electrons} \\
Aex(ge+1:ge+nne) &= \text{Aex}(Be) ; \quad \% \text{set current position of new carriers to parent (electron) loc} \\
Ah(gh+1:gh+nne) &= \text{ones}(1,nne) ; \quad \% \text{as holes} \\
Ahx(gh+1:gh+nne) &= \text{Ahx}(Bh) ; \\
\text{Re(Be)} &= \text{rand}(1,nne) ; \quad \% \text{reset parent injection loc} \\
\text{get(t)} &= ge; \\
ge &= ge + nne; \\
gh &= gh + nne; \\
G &= G + nne; \\
nne;
\end{align*}

\text{end}

\text{if } (nnh > 0)

\begin{align*}
Aex(ge+1:ge+nnh) &= \text{Ahx}(Bh) ; \quad \% \text{set current position of new carriers to parent (hole) loc} \\
Ae(ge+1:ge+nnh) &= \text{ones}(1,nnh) ; \quad \% \text{as electrons} \\
Ahx(gh+1:gh+nnh) &= \text{Ahx}(Bh) ; \\
Ah(gh+1:gh+nnh) &= \text{ones}(1,nnh) ; \quad \% \text{as holes} \\
Aex(ge+1:ge+nnh) &= \text{Aex}(ge+1:ge+nnh) ; \quad \% \text{copy current loc to injection loc} \\
Ahx(gh+1:gh+nnh) &= \text{Ahx}(gh+1:gh+nnh) ; \\
Ahx(0,Bh) &= \text{Ahx}(Bh) ; \quad \% \text{reset parent injection loc}
\end{align*}
Rh(Bh)=rand(1,nnh); % reset parent luck
ge = ge + nnh;
gh = gh + nnh;
ght(t)=gh;
G = G + nnh;
nnh;
end

% -collect holes at interface
Bh(fh:gh) = Ah(fh:gh) & (Ahx(fh:gh)>Wm);
Bht(t)=sum(Bh);

tx0(Bh)=tx(Bh);
Ah(Bh)=0; % remove holes from m-region
    % add to sum at interface

Rb(Th)=rand(1,sum(Th));
Oh(Th)=exp(-(1/Tp)*((tx(Th)-tx0(Th))>Rb(Th));
Th=Bh|Oh;
Oht(t)=sum(Oh);
nW=sum(Th);
nWt(t)=nW;
%
save('Oht.mat','Oht');
save('nWt.mat','nWt');
save('Bht.mat','Bht');
save ('Vmt.mat','Vmt');
save ('Mt.mat','Mt');
save('Bet.mat','Bet');
save('ght.mat','ght');
save('get.mat','get');
\% -count Jp(x=0,t)
Be(fe:ge) = Ae(fe:ge) \& (Aex(fe:ge)<0);
Ae(Be)=0; \% remove electron from m-region
Bet(t)=sum(Be);

J(t) = -sum(Be(fe:ge));
Jp(t)=sum(Th) ;

\% bookkeeping
Gt(t)=G;

m = sum(Ae(fe:ge)) + sum(Ah(fh:gh)); \% number of carriers left in multiplication region
Mt(t)=m;
if (m < 1), display('no carriers left'), break;
end; \% no carriers left in Mregion, but !set success=false
Vm = Vm0 -nW*q/(eps/Wtcb*S);

if (BJT && mod(t,1)==0)
\quad Ie\_mean=Nn*vnb*q*S*exp(-Ec/(K*T/q))\*exp((Vm0-Vm)/(K*T/q));
\quad Ne\_mean=Ie\_mean\*1*dt/q;
\quad Dis=poissrnd(Ne\_mean);
\quad Di=round(Dis);
\quad Dit(t)=Di;
\quad Ae(ge+1:ge+Di)=ones(1,Di);
\quad Aex(ge+1:ge+Di)=Wm;
\quad Aex0(ge+1:ge+Di)=Wm;
\quad Re(ge+1:ge+Di)=rand(1,Di);
\quad ge=ge+Di;
end

\% Nd=sum(Dit);
% numerical speedup: compact f
if (mod(t,100)==0) % 'find' takes a while, so don't run every time
    fe = find(Ae,1,'first');
    if isempty(fe)
        fe=1;
    end
% At some point we need to shift A, in order to keep its size small.
% We can't do this every time, to balance the cost of shifting vs keeping size(A) small
if (fe > 2e3)
    % shift by f-1
    Ae(1:ge-fe+1)=Ae(fe:ge);
    Aex(1:ge-fe+1)=Aex(fe:ge);
    Aex0(1:ge-fe+1)=Aex0(fe:ge);
    Re(1:ge-fe+1)=Re(fe:ge);
    % reinitialize tail
    Ae(ge-fe+2:ge)=zeros(1,fe-1);
    Aex(ge-fe+2:ge)=zeros(1,fe-1);
    Aex0(ge-fe+2:ge)=zeros(1,fe-1);
    Re(ge-fe+2:ge)=rand(1,fe-1);
    % reset markers
    ge = ge-fe+1;
    fe = 1;
end
fh = find(Ah,1,'first'); % repeat for hole dat
if (fh > 2e3)
    Ah(1:gh-fh+1)=Ah(fh:gh);
    Ahx(1:gh-fh+1)=Ahx(fh:gh);
    Ahx0(1:gh-fh+1)=Ahx0(fh:gh);
    Rh(1:gh-fh+1)=Rh(fh:gh);
    Ah(gh-fh+2:gh)=zeros(1,fh-1);
    Ahx(gh-fh+2:gh)=zeros(1,fh-1);
    Ahx0(gh-fh+2:gh)=zeros(1,fh-1);
Rh(gh-fh+2:gh)=rand(1,fh-1);
gh = gh-fh+1;
fh = 1;
end
Be = false(1,Z); % zero out B when f changes
Bh = false(1,Z);
end % compacting
if DEBUG && (mod(t,1000)==0)
display(t);
toc
end
end % time iteration
toc

% %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% ***** Analysis %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% **%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

dVmdt = diff(Vmt(1:t-1))/dt;
dVmdt = [0 dVmdt];
Ip = abs(smooth(J(1:t-1)*q/dt,mas)); % hole current [A]
I = abs(smooth(J(1:t-1)*q/dt + dVmdt*(eps/Wm*S),mas)); % total current [A]

t = t - floor(mas/2); % shorten t interval since smoothing seems to add kinks at the end
I = I(1:t-1);

Ip = I(1:t-1);
tt = [1:t-1]*dt;
finished(sweep) = success;

if success

    peak(sweep) = max(I);
    % set(gca, 'Xscale', 'log');
    gt = find(I > input_th);  % find the indices where the current is above the threshold

    if sum(gt)>0
        pde(sweep) = true;           % current pulse detected
        jitter(sweep) = gt(1);
        halfmax = find(I > peak(sweep)/2);  % find the half max points
        width(sweep) = (halfmax(end)-halfmax(1))*dt;
        gain(sweep) = ((G-1)/2);
    end  % gt true

    nWs(sweep) = nW;
end % success
end % Nsweep

% keep data only for pulses detected
B = false(1,Nsweeps);
B = finished & pde;
jitter = jitter(B)*dt;
peak = peak(B);
width = width(B);
gain = gain(B)/Ni;
nWs = nWs(B);
spde = sum(pde)/Nsweeps;
Nfinished = sum(finished)/Nsweeps;
% Distribution generation for different figures of merits
% hist(peak,100);
% ylabel('Count','fontsize',14);
% xlabel('Peak current','fontsize',14);
% figure;
%
% hist(gain,100);
% ylabel('Count','fontsize',14);
% xlabel('Gain','fontsize',14);
% figure;
%
% hist(jitter,100);
% ylabel('Count','fontsize',14);
% xlabel('Jitter','fontsize',14);
% figure;

plot(I,'linestyle',':'); % plot the transient current

% ylabel('Current','fontsize',14);
% xlabel('time in 0.1ps','fontsize',14);

% real jitter number: std dev of time of detection
% jitterMax = (max(jitter)-min(jitter))*dt;
% jitterSigma = std(jitter);
% jittermean=mean(jitter);
% currentsigma=std(peak);
% currentmean=mean(peak);
% gainsigma=std(gain);
% gainmean=mean(gain);
end