EPFL

Thèse n° 10 381

Advanced Silicon and SWIR Single-Photon Avalanche Diodes: Design, Simulation, and Characterization

Présentée le 19 avril 2024

Faculté des sciences et techniques de l'ingénieur
Laboratoire d'architecture quantique
Programme doctoral en photonique

pour l'obtention du grade de Docteur ès Sciences

par

Ekin KIZILKAN

Acceptée sur proposition du jury

Prof. O. Martin, président du jury
Prof. E. Charbon, directeur de thèse
Prof. A. Tosi, rapporteur
Dr F. Bussières, rapporteur
Prof. A. Fontcuberta i Morral, rapporteuse

École
polytechnique
fédérale
de Lausanne

2024

One should own nothing.
Then, one owns the world.
— Panait Istrati

To my parents…

Acknowledgements

This thesis would not have been possible without the guidance and support of many people who have contributed their valuable time and insights to my academic and research endeavors. I am deeply grateful to each one of them.

First and foremost, I extend my sincerest gratitude to Prof. Edoardo Charbon, who welcomed me into the AQUA laboratory as a part-time working master's student in 2017. This opportunity marked a turning point in my research career. Prof. Charbon, as my supervisor, consistently challenged me to push my boundaries and motivated me to delve deeper into my research field. His guidance has been instrumental in shaping my academic journey.

I am thankful to Dr. Claudio Bruschini for his unwavering positivity, critical feedback, and valuable discussions during our meetings. His insights have greatly enriched my understanding and approach to my research.

My gratitude extends to Dr. Myung-Jae Lee, who guided me through the intricacies of my master's thesis and taught me the fundamentals of single-photon-avalanche diodes. His mentorship has been crucial in my academic development.

I owe a debt of gratitude to Prof. Serdar Kocaman for introducing me to the fascinating world of semiconductors. Working in a cleanroom and gaining hands-on experience under his guidance has been an invaluable part of my education.

A special thanks to Utku Karaca, whose support has been a cornerstone of my PhD journey. Since our paths crossed in 2015 at METU, Utku has been a pillar of reliability and support. His invaluable suggestions on the InGaAs/InP SPAD work have been indispensable. I am truly fortunate to have such a valuable friend and colleague by my side throughout this process.

I am grateful to Yunqiao Xu for her unwavering support during the thesis writing period. Your encouragement and companionship have been a great comfort to me.

I would also like to express my appreciation to my fellow AQUA lab members for their camaraderie and collaboration. To Francesco Gramuglia and Pouyan Kesheverzian, with whom I shared an incredible collaboration in 55 nm SPAD design, your friendship and teamwork have been a source of joy and inspiration. To Vladimir Pesic, whose brilliance in research became apparent soon after he joined as my master's student, your contributions to the InGaAs/InP SPAD simulation work have been invaluable. To Ming-lo Wu, my office mate, thank you for the enriching discussions on research and the many moments of laughter we shared. To other members, Ivan Michel Antolovic, Samuel Burri, Feng Liu, Michael Alan Wayne, Augusto Carimatto, Arın Can Ülkü, Preethi Padmanabhan, Scott Lindner, Ermanno Bernasconi, Andrada Muntean, Andrei Ardelean, Jad Benserhir, Barış Can Efe, Simone Frasca, Tommaso Milanese, Andrea Ruffino, Halil Kerim Yıldırım, Ghalia Dhaoui, Bedirhan Ilık, Jiuxuan Zhao, Paul Mos, Chang Liu, Chufan Zhou, Yasemin Uzun, Batuhan Keskin, Won Yong Ha, Kodai Kaneyasu, Samuele Bisi, and Yating Zou. To our lovely administrative assistants for solving every problem quickly Brigitte Khan and Begonia Tora.

To my friends, Aslı Şahin, Mustafa Kangül, Selen Keleş, Aras Mumcuyan, Şilan Şahan, Ulaş Dinç, Ashley Reynolds, Deniz Yalçın, Ece Özelçi, Thorsten Rutzen, Floreana Miesen, Furkan Ayhan, Tugay Karagüzel, Ayhan Şavklıyıldız, Nilay Şavklıyıldız, Ecem Akkaya, Buğra Arıkan, Andaç Yiğit, Orçun Gümüş, Caner Boz, Merve Boz, İsmail Uzunlar, Begüm Uzunlar, Adilcan Uğurlu, Bahar Taşkesen, Yiğit Özer, Kübra Çırçır, Muammer Kozan, and Halit Dolaş, your continuous support and friendship have been my stronghold.

Finally, to my parents, your unconditional love and unwavering support have not only shaped the person I am today but have also been the foundation upon which I've built my academic pursuit.

Lausanne, 07 March 2024 E. K.

Abstract

Low-level light detection with high spatial and timing accuracy is a growing area of interest by virtue of applications such as light detection and ranging (LiDAR), biomedical imaging, time-resolved Raman spectroscopy, and quantum applications. Single-photon avalanche diodes (SPAD) offering the capability of detecting picosecond transients at the single-photon level are becoming a new trend for these needs. Silicon-based SPADs have demonstrated superior performance in the visible wavelength window benefiting from the mature complementary-oxide-metal semiconductor technology. Nevertheless, silicon-based SPADs exhibit a significant efficiency drop in the near-infrared (NIR) and short-wave infrared (SWIR) regions due to their bandgap (1.1 eV) and since silicon is an indirect bandgap material. Extending SPAD's efficiency towards the NIR/SWIR region is a key requirement for eye-safe LiDAR and fiber optic-based telecommunication applications.

This thesis explores high-performance single-photon detectors using Silicon and InGaAs/InP-based SPAD technologies for QKD and LiDAR applications. It aims to develop compact detectors operable at near room temperature. The research extensively characterizes SAG-based and double-diffusion InGaAs/InP SPADs, alongside Silicon SPADs in 55 nm BCD technology, assessing DCR, PDP, timing jitter, and uniformity. The SAG-based design offers a novel InGaAs/InP SPAD structure, reducing the electric field at the edges and improving DCR and uniformity. Moreover, this method does not utilize the standard shallow diffusion and improves the fill factor of the SPAD. The thesis also investigates smaller diameter devices for further DCR reduction.

In addition, the thesis introduces a unique simulation environment for PDP, DCR, IV, and breakdown voltage of SPADs using TCAD tools and focusing on 2D simulations. The simulation successfully predicts these performance metrics, allowing an easier and more robust design environment. It will be useful to optimize the device performance for specific target metrics and reduce fabrication iterations to achieve the best performance. Apart from SAG-based InGaAs/InP SPADs, double-diffusion-based InGaAs/InP SPADs are also examined, using both shallow and deep diffusion processes for device fabrication. This technique helps in managing the electric field within the SPAD, thereby reducing-edge breakdown. The research explores the impact of diffusion depths on device performance, especially DCR. Adjusting these depths affects the electric field and TAT generation, influencing the SPAD's efficiency. The work also looks forward to enhancing SPAD performance and developing SPAD arrays integrated with ROIC.

Finally, the thesis explores the implementation and advantages of Silicon SPADs. Leveraging mature silicon fabrication technologies, Silicon SPADs offer ease of implementation, advanced back-end processing, and low defect ratios. Despite their limited efficiency in the NIR range, the benefits of monolithic integration, cost-effectiveness, and widespread availability make them an attractive alternative. Four different SPAD designs were fabricated using 55 nm BCD technology, addressing the challenges and solutions in designing with standard doping layers. A particular focus is given to deep SPAD designs with and without a PW layer, revealing that the inclusion of the PW layer significantly improves efficiency.

Key words: 3D integration, afterpulsing, BCD, CMOS, BSI, FSI, DCR, PDP, PET, SPAD, time-of-flight, timing jitter, TSCPC, III/V, InGaAs, InP, SAG.

Résumé

La détection de la lumière à faible niveau avec une grande précision spatiale et temporelle est un domaine d'intérêt croissant en raison d'applications telles que la détection et la télémétrie par la lumière (LiDAR), l'imagerie biomédicale, la spectroscopie Raman résolue dans le temps et les applications quantiques. Les diodes à avalanche à photon unique (SPAD) offrant la capacité de détecter des transitoires de l'ordre de la picoseconde au niveau du photon unique deviennent une nouvelle tendance pour répondre à ces besoins. Les SPAD à base de silicium ont démontré des performances supérieures dans la fenêtre de longueur d'onde visible en bénéficiant de la technologie mature des semi-conducteurs à base d'oxyde complémentaire et de métal. Néanmoins, les SPAD à base de silicium présentent une baisse d'efficacité significative dans les régions du proche infrarouge (NIR) et de l'infrarouge à ondes courtes (SWIR) en raison de leur bande interdite (1,1 eV) et du fait que le silicium est un matériau à bande interdite indirecte. L'extension de l'efficacité du SPAD vers la région NIR/SWIR est une exigence clé pour les applications LiDAR sans danger pour les yeux et les applications de télécommunication basées sur la fibre optique.

Cette thèse explore les détecteurs à photon unique de haute performance utilisant les technologies SPAD basées sur le silicium et l'InGaAs/InP pour les applications QKD et LiDAR. Elle vise à développer des détecteurs compacts fonctionnant à une température proche de la température ambiante. La recherche caractérise de manière approfondie les SPAD à base de SAG et à double diffusion InGaAs/InP, ainsi que les SPAD au silicium en technologie BCD 55 nm, en évaluant le DCR, le PDP, la gigue temporelle et l'uniformité. La conception basée sur le SAG offre une nouvelle structure de SPAD InGaAs/InP, réduisant le champ électrique sur les bords et améliorant le DCR et l'uniformité. De plus, cette méthode n'utilise pas la diffusion standard peu profonde et améliore le facteur de remplissage du SPAD. La thèse étudie également des dispositifs de plus petit diamètre pour une réduction supplémentaire du DCR.

En outre, la thèse introduit un environnement de simulation unique pour le PDP, le DCR, l'IV et la tension de claquage des SPAD en utilisant les outils TCAD et en se concentrant sur les simulations 2D. La simulation prédit avec succès ces mesures de performance, permettant un environnement de conception plus facile et plus robuste. Il sera utile d'optimiser les performances du dispositif pour des mesures cibles spécifiques et de réduire les itérations de fabrication afin d'obtenir les meilleures performances. Outre les SPAD InGaAs/InP à base de SAG, les SPAD InGaAs/InP à double diffusion sont également examinés, en utilisant des processus de diffusion à la fois superficiels et profonds pour la fabrication des dispositifs. Cette technique permet de gérer le champ électrique à l'intérieur du SPAD, réduisant ainsi le claquage des bords. La recherche explore l'impact des profondeurs de diffusion sur les performances des dispositifs, en particulier le DCR. L'ajustement de ces profondeurs affecte le champ électrique et la génération de TAT, influençant ainsi l'efficacité du SPAD. Les travaux portent également sur l'amélioration des performances des SPAD et sur le développement de réseaux de SPAD intégrés au ROIC.

Enfin, la thèse explore la mise en œuvre et les avantages des SPAD en silicium. S'appuyant sur des technologies matures de fabrication du silicium, les SPAD en silicium offrent une facilité de mise en œuvre, un traitement final avancé et de faibles taux de défauts. Malgré leur efficacité limitée dans le domaine du proche infrarouge, les avantages de l'intégration monolithique, de la rentabilité et de la disponibilité généralisée en font une alternative attrayante. Quatre conceptions différentes de SPAD ont été fabriquées à l'aide de la technologie BCD 55 nm, abordant les défis et les solutions dans la conception avec des couches de dopage standard. Une attention particulière est accordée aux conceptions SPAD profondes avec et sans couche PW, révélant que l'inclusion de la couche PW améliore de manière significative l'efficacité.

Mots clefs : Intégration 3D, afterpulsing, BCD, CMOS, BSI, FSI, DCR, PDP, PET, SPAD, temps de vol, gigue temporelle, TSCPC, III/V, InGaAs, InP, SAG.

Contents

List of Figures

List of Tables

Acronyms

APDs avalanche photodiodes

AR anti-reflection

Au Gold

BCD Bipolar-CMOS-DMOS

BNW buried n-well

CIS CMOS image sensor

CMOS complementary metal–oxide–semiconductor

cps counts per second

DCR Dark count rate

dToF direct time of flight

DUT device under test

FEM finite element method

FWHM Full width at half maximum

Ge Germanium

GRs guard rings

H_2O_2 hydrogen peroxide

H_3PO_4 phosphoric acid

HCl hydrochloric acid

InGaAs indium gallium arsenide

InGaAsP indium gallium arsenide phosphide

InP indium phosphide

iToF indirect time of flight

K Kelvin

LiDAR Light detection and ranging

MOCVD metal organic chemical vapor deposition

NbN niobium nitride

NbTiN niobium titanium nitride

NIR near infrared

PCB printed circuit board

PDP Photon detection probability

PET Positron emission tomography

PMTs Photomultiplier tubes

PQAR passive quench-active recharge

PW p-well

QKD Quantum key distribution

Radar radiation detection and ranging

ROIC readout integrated circuit

SAG Selective area growth

SAGCM separate absorption-grade-charge-multiplication

SAM separate absorption and multiplication

SEM scanning electron microscope

SIMS secondary ion mass spectrometry

SiN silicon nitride

SMU source measurement unit

SNSPDs Superconducting nanowire single-photon detectors

SPW shallow p-well

SRH Shockley-Read-Hall

SWIR short-wave infrared

TAT trap-assisted tunneling

TCAD Technology computer-aided design

TCSPC time-correlated single-photon counting

TDC time-to-digital converter

VP punch-through voltage

WSi tungsten silicide

Zn Zinc

1 Introduction

Max Planck's studies in the early 20th century on black body radiation have revolutionized the understanding of physics. His formula, referred to as Planck's law, explained the experimentally observed spectral energy density of a blackbody [1]. In his equations, he used the concept of energy quanta proposing energy changes could occur only in fixed amounts. Building upon Max Planck's work, Albert Einstein proposed that light also comprises energy pockets (photons) or light quanta [2]. These studies paved the way for the quantum theory of physics predicting various phenomena such as quantum entanglement and superposition. Consequently, the fields of quantum computing and cryptography have been developed in pursuit of leveraging the understanding of quantum theory into technological advancements. Today, there are commercially available quantum cryptography systems by ID Quantique that ensure secure communication based on the concept of quantum entanglement. Emerging quantum information applications escalated the interest in robust, high efficiency, low noise, and low timing jitter single-photon detectors [3–5]. Single-photon avalanche diodes (SPADs) could meet these expectations. As a result, they have been widely used in quantum information applications [6–12]. The application space of SPADs is broadened thanks to their remarkable capabilities for photon counting and accurate marking of photon arrival time. Several works have demonstrated that SPADs could replace p-i-n or avalanche photodiodes (APDs) in characterizing optical fiber links [13–15], while SPADs are widely used in Light detection and ranging (LiDAR) [16–21], and Fluorescence lifetime imaging [22–24].

According to the Rp-Photonics encyclopedia the near infrared (NIR) wavelength ranges from 700 nm to 1400 nm and the short-wave infrared (SWIR) wavelength ranges from 1400 nm to 3000 nm [27]. Operating in NIR/SWIR bands has various advantages for SPAD-based applications. We can divide the light-travelling media into two being; free space and fiber optics. Free-space applications using active illumination sources such as free-space LiDAR are subject to optical power attenuation. Factors like water vapor, ozone, carbon dioxide, and other gases in the atmosphere can absorb specific wavelengths, influencing the transmission. As seen in Figure 1.1, the transmittance of the atmosphere slightly increases towards NIR/SWIR infrared wavelengths offering an infrared transmission window between 1.5 μm and 1.75 μm

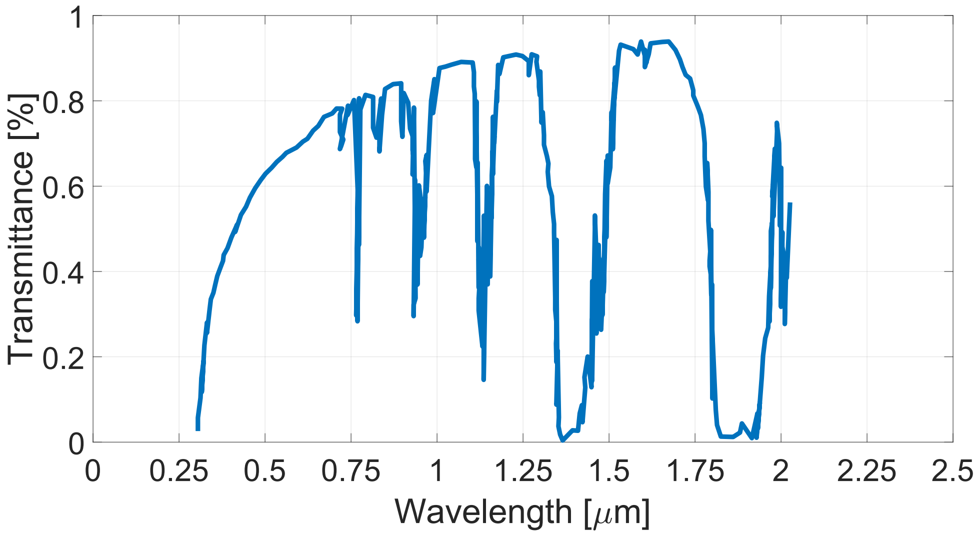


Figure 1.1: Atmospheric transmission spectra for a vertical length from ground to space. Replotted after [25].

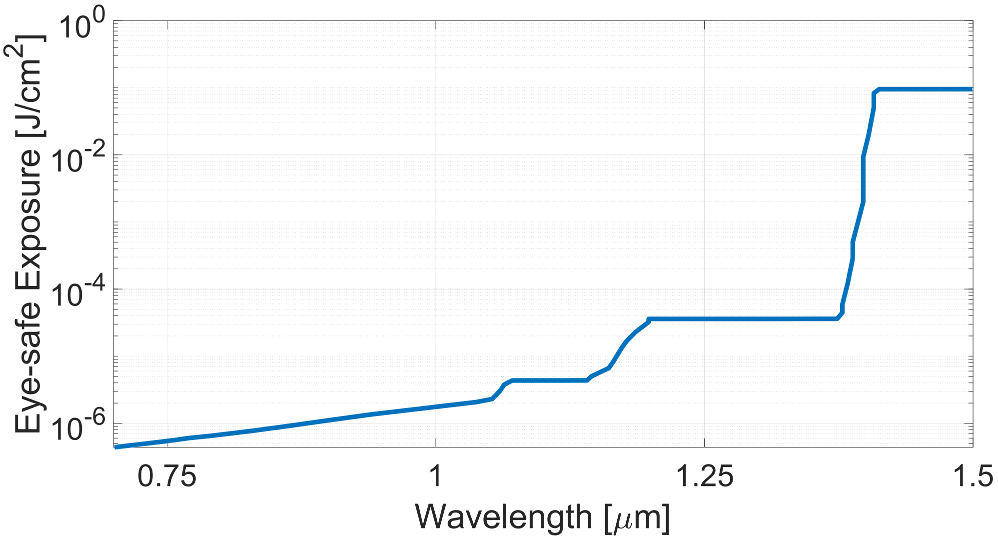


Figure 1.2: Eye-safe laser limit for a 1 ns pulse duration [26].

wavelengths. A higher transmission rate allows photons to travel longer before being absorbed, thus allowing an increased operation distance. Serving a similar purpose, the eye-safe laser power limit increases in the SWIR range. As seen in the Figure 1.2, the maximum permissible exposure limit is less than 10^{-6} J/cm^2 around 750 nm wavelengths and increases to 10^{-1} J/cm^2 around 1500 nm wavelength. This five orders of magnitude difference allows higher input laser power levels. For a given transmission rate, photons will travel longer starting from a high number of photons, eventually making distant objects visible.

Another benefit of operating in the NIR/SWIR wavelength range using active illumination emerges from reduced background light. Regarding free-space LiDAR or free-space quantum communication applications, ideally, the receiver should detect the photons emitted from the input laser of the system. Detection of photons caused by dark carriers or light in the environment results in false counts. They will disturb the system's performance by introducing noise. The Sun is the main source of background light and its emission spectrum could be approximated to a blackbody of 6000 Kelvin (K). Figure 1.3 shows solar spectral irradiance at the top of the atmosphere and at sea level. Surprisingly, the Sun's emission peaks around visible wavelengths and decreases towards infrared wavelengths. The SWIR detectors placed at sea level will benefit from ~ five times fewer background photons while it will be ~ 10 times if they are placed outside the atmosphere. However, this advantage would depend on the dark noise of the device which varies with the operation temperature, frequency, and device type. The devices that exhibit lower dark noise compared to ambient light would benefit from reduced Sun emission.

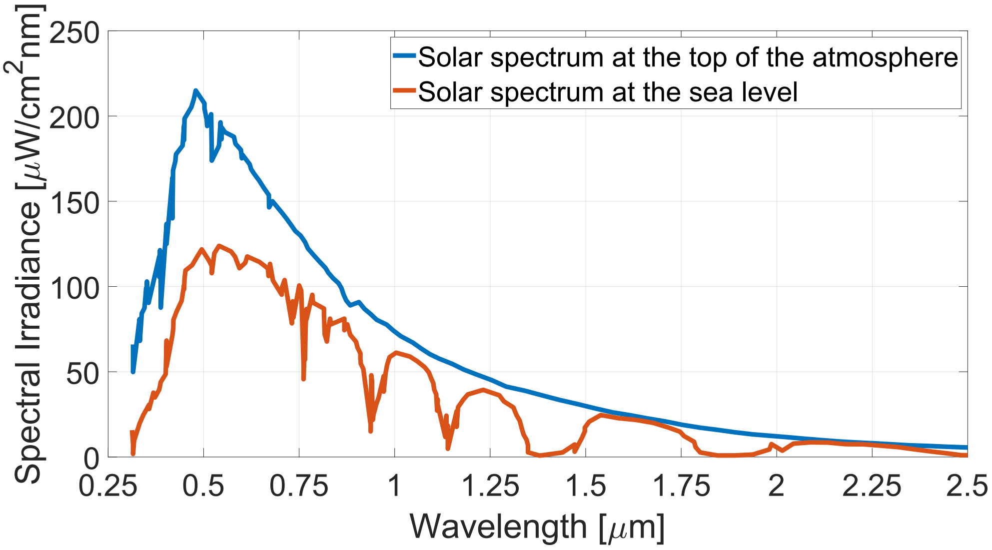


Figure 1.3: Solar irradiance at the top of the atmosphere and sea level. Re-plotted after [28].

Finally, for applications that use fiber optic cables for transmission media, the NIR/SWIR range covers the low-loss windows. Low-loss windows in optical fiber cables refer to specific wavelength ranges where the fiber has minimal signal loss, allowing light to travel longer distances with less attenuation. These windows are at wavelengths where the material of the fiber — usually silica — exhibits minimal absorption and scattering. The most common low-loss windows are centered around 850 nm, 1300 nm, and 1550 nm. The 1550 nm window is particularly favored for long-distance communication because it offers the lowest loss, thus greatly enhancing the efficiency of the optical fiber as depicted in Figure 1.4. Moreover, Erbium-doped fibers are used in fiber optics to enhance signal strength over long distances. Erbium is a rare earth element that, when added to the fiber, can amplify light at wavelengths around 1550 nm without the need to convert it back to an electrical signal for amplification, thus preserving signal quality and reducing noise.

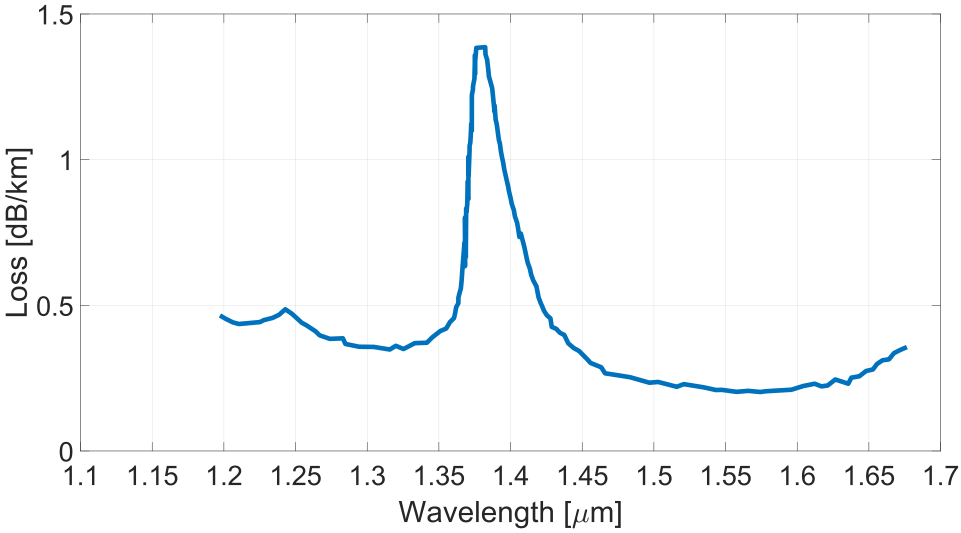


Figure 1.4: Loss spectrum of silicon dioxide glass fibers. Peaks around 1250 nm and 1370 nm wavelengths correspond to hydroxide absorption. Re-plotted after [29].

1.1 Single-photon Detectors

1.1.1 Photomultiplier Tubes

Photomultiplier tubes (PMTs) are first-generation, high-gain light detectors that can detect single photons. Figure 1.5 depicts the subparts of a PMT. As photons enter through the faceplate, they are absorbed by the photocathode, a specialized material chosen for its ability to emit electrons upon exposure to light due to the photoelectric effect. These photoelectrons are then directed toward the electron multiplier structure by the focusing electrode. The dynodes, arranged in a sequential chain within the PMT, serve as the heart of the electron multiplication process. Each dynode is set at a successively higher positive voltage, so when the electrons from the photocathode strike the first dynode, several secondary electrons are released. These secondary electrons are then directed to the next dynode, continuing the multiplication process. This cascade of electron multiplication occurs across each dynode stage, resulting in an exponential increase in the number of electrons similar to the avalanche multiplication effect in SPADs. After the final amplification stage at the last dynode, the swarm of electrons reaches the anode. Here, they are collected and the total charge is converted into a current pulse. This pulse is proportional to the initial number of photons that struck the photocathode, providing a powerful signal indicative of the light's intensity. The signal is then transmitted through the stem pin, which is the electrical output of the PMT.

The photocathode sets the limit of the quantum efficiency of a PMT. Photocathodes are usually made of Alkali metals. They can achieve 40% efficiency in the visible range and 45 ps timing jitter [30]. These photocathodes exhibit a steep decrease in quantum efficiency towards NIR/SWIR wavelengths. In order to gain efficiency in these wavelengths, III/V pn junction-based photocathodes are introduced [31, 32]. However, their efficiency is limited to 2% at 1550 nm. The major disadvantages of PMTs are their bulky structure, high voltage requirement (100 V - 3000 V), and low detection efficiency. Nevertheless, they are widely used in the detection of gamma-rays in Positron emission tomography (PET) scanners and neutrino detection for fundamental physics experiments [33–35].

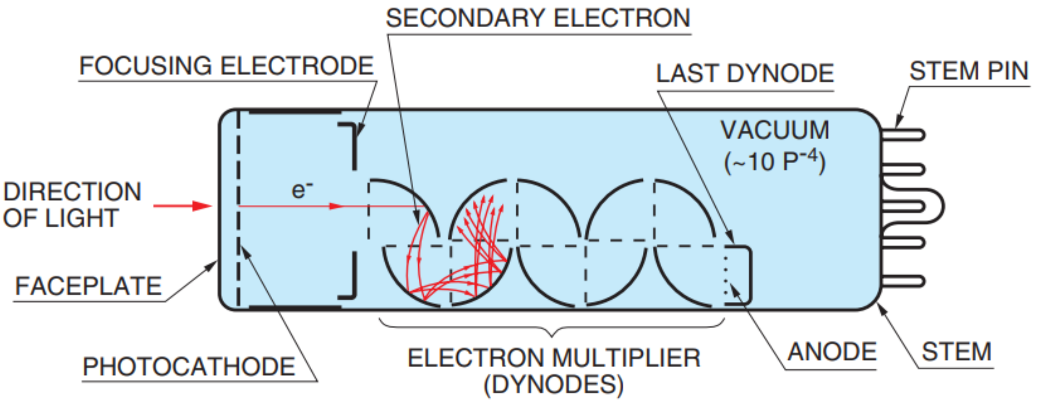


Figure 1.5: Structure of a Photomultiplier tube [30].

1.1.2 Superconducting Nanowire Single-Photon Detectors

Superconducting nanowire single-photon detectors (SNSPDs) operate based on the principle of superconductivity. When a superconducting material, typically a thin film of niobium nitride (NbN), niobium titanium nitride (NbTiN), or tungsten silicide (WSi), is cooled to a sufficiently low temperature, it enters a superconducting state where it can conduct electricity with zero resistance. To form a detector, this superconducting film is first shaped into a nanowire. Then biased with an electrical current close to the maximum current it can carry while remaining superconductive (critical current).

When a photon enters the detector and strikes the nanowire, absorption of the photon causes a temporary non-superconducting region or hotspot. Because the current is nearly at the critical value, this disturbance in superconductivity cannot be sustained. Then the hotspot spreads to its surroundings eventually forming a resistive slot. This process results in a measurable change in the resistance of the nanowire, and because the wire is part of an electrical circuit, this change in resistance causes a voltage pulse. This pulse is then amplified and read out by the detector's electronics. After the detection event, the hotspot rapidly cools back down to below the critical temperature, and superconductivity is restored, readying the SNSPD for the next photon detection event.

The entire process from photon absorption to hotspot formation and restoration of superconductivity occurs in picoseconds to nanoseconds, allowing for rapid and precise measurements. SNSPDs are highly sensitive and capable of detecting individual photons with high efficiency. The very first demonstration of NbN SNSPDs was achieved in 2001 and already exhibited 20% quantum efficiency at 810 nm wavelength with negligible dark noise operating at 4.2K [36]. In less than 20 years, their efficiency reached the orders of 90% peaking at 99.5% at 1500 nm wavelength with 34 ps timing jitter [37–42] operating below 3 K. In order to boost their absorption, SNSPDs are fabricated with Bragg reflectors and cavities. Moreover, a record level of 2.6 ps timing jitter in visible wavelengths and 4.3 ps at 1550 nm wavelength has been demonstrated [43, 44]. A key obstacle in the advancement of SNSPDs is constructing extensive two-dimensional arrays to enhance spatial resolution. As the number of pixels in these arrays grows, the heat load on the transmission cables approaches the maximum cooling capacity of standard cryostats [45]. Despite this challenge, arrays with dimensions as large as 32x32 have been successfully developed [46–48]. Moreover, recently an array of 800x500 SNSPDs with a pixel size of 5 μm by 5 μm has been demonstrated [49]. The primary drawback of SNSPDs, however, lies in their need for operation at cryogenic temperatures (~ few Kelvin). The cryostats required for this purpose are typically large, complicated, and costly.

1.1.3 Single-photon avalanche diodes

Single-photon avalanche diodes (SPADs) are compact, p-n junction-based diodes operating above their breakdown voltage. They exhibit very high sensitivity and timing accuracy. Details of SPAD operation, research trends, and figures of merits are discussed in the next Chapter.

1.2 Applications for NIR/SWIR Single-Photon Detectors

This section briefly introduces two main applications benefiting from single-photon sensitivity in the NIR/SWIR range. Quantum key distribution (QKD) is a key research area in the pursuit of developing quantum internet. It aims to provide the most secure form of communication. Using already established fiber infrastructure which operates with the lowest loss in SWIR wavelengths drives these systems to utilize NIR/SWIR single-photon detectors. LiDAR is the other application that is analyzed here. Over the past decade, the possibility of employing them in augmented reality/virtual reality and autonomous vehicles has collected considerable interest in LiDAR. The Near-NIR/SWIR bands allow the use of higher laser power, enabling LiDAR systems to function over greater distances with an improved signal-to-noise ratio.

1.2.1 Quantum Key Distribution

Quantum key distribution is an advanced cryptographic protocol that leverages the principles of quantum mechanics to enable two parties to generate a shared, secret key, which is secure against any computational or mathematical attack. The essence of QKD lies in quantum entanglement and the fact that observing a quantum system inevitably alters its state. In QKD, information about the key is encoded in the quantum states of particles, such as the polarization states of photons. A common protocol used in QKD is BB84, where a sender (often called Alice) transmits a sequence of photons, each randomly polarized in one of four orientations [50]. Each orientation corresponds to a bit value (0 or 1), but crucially, these bits are encoded in two incompatible bases. Bases can be thought of as different measuring standards such as measuring the polarization of the photon through two different polarization angles. The receiver (Bob) measures the incoming photons using bases chosen at random, some of which will align with Alice's bases. After the transmission, Alice and Bob publicly compare their bases and keep only the bits where their bases match. In an ideal case of no error and no one eavesdropping (Eve), half of their measured bits have to match. If that is the case, Alice can rely on the security of the communication channel and encode her message using a subset of matching bits.

The security of QKD arises from the fact that any Eve trying to intercept the key cannot measure the quantum states without disturbing them. Since Eve doesn't know the bases used for each bit, her measurements will introduce errors in the key bits, detectable by Alice and Bob. They can estimate the level of eavesdropping by comparing a subset of their key bits. If the error rate exceeds a certain threshold, they know the key is compromised and discard it.

One of the main challenges of QKD is the need to maintain the integrity of the quantum states over long distances, as interaction with the environment can cause loss or decoherence. Current systems use optical fibers or free-space channels for transmission, with the distance limited to a sub-thousand kilometer due to optical losses [51–55]. A recent study has demonstrated that combining a 2000 km fiber link with 2600 km of free space link (with a secret key rate of 47.8 kilobits per second), a 4600 km total channel distance could be achieved [56]. Another push forward in QKD development is to improve the secret key rate. Secret key rates of 64 and 115.8 Mbps have been demonstrated using SNSPDs [57, 58]. The low noise of SNSPDs allows them to detect photons under higher loss. This translates to a longer detection distance for a given fiber loss rate. However, their sub 4 K operation requirement increases the system complexity and the cost. An operation temperature above 225 K could be achieved by thermo-electronic cooling taking away the need of using cryostats. InGaAs/InP SPADs could operate with a reasonable noise at these temperatures and various studies have demonstrated QKD application with this method [12, 53, 59–61]. There is still room to improve the performance of InGaAs/InP SPADs pushing them to reach a similar performance with SNSPDs at higher temperatures.

1.2.2 Light detection and Ranging

Light detection and ranging has a similar principle as radiation detection and ranging (Radar) in the sense that a signal is sent to the scene and its reflection is collected by the receiver to acquire distance information. Radar operates in longer wavelengths and as a result, has lower spatial resolution. It is effective for large-scale environmental scanning, like tracking weather patterns or large objects. LiDAR usually operates in the visible or NIR/SWIR bands with high spatial resolution allowing to construction of 3D models of the scene. Methods used in Lidar to obtain 3D information could be divided into two main families called; direct time of flight (dToF) and indirect time of flight (iToF).

In terms of implementation, dToF and iToF systems present distinct challenges and considerations. dToF systems, which rely on measuring the time taken for a laser pulse to return, require precise timing mechanisms and high-powered lasers to emit short, intense light pulses [62]. This requires sophisticated electronics for accurate time measurement and robust laser sources capable of producing the necessary pulse intensity. On the other hand, iToF systems operate by emitting a continuous light source modulated at specific frequencies and measuring the phase shift of the returned light. The continuous wave nature of the light source in iToF systems means that they do not require the high-power laser pulses of dToF systems, allowing for more compact and energy-efficient designs. However, in iToF systems, there exists a compromise between operational range and resolution, leading to a lower resolution when extending the operational distance [63].

Using SPADs in dToF systems brings several significant advantages, such as long-range operation, high-depth resolution, and compactness. SPADs are known for their exceptional sensitivity and timing accuracy which is crucial in dToF systems for accurately measuring the time of flight of photons. Operation range depends on the minimum light level the receiver could detect and the single-photon sensitivity of the SPADs could successfully address this issue. Furthermore, the depth resolution in dToF systems is intrinsically linked to the timing precision of the receiver. To illustrate, a timing jitter of 100 ps equates to a depth resolution of about 3 centimeters, considering the speed of light. Recently, advancements in SPAD technology have showcased an impressive timing jitter of just 12.1 ps, opening the door to achieving sub-centimeter resolution and long distance when employed by a LiDAR system [64]. The operational range of SPAD-based LiDAR systems faces a significant challenge from background sunlight. To mitigate this, coincidence detection is used, allowing these systems to distinguish true signals from background noise effectively. This technique has enabled the development of Silicon SPAD-based LiDAR systems that can reach ranges up to 100 meters [65, 66]. Thanks to reduced background light and higher eye-safety limit in the NIR/SWIR range, distances of tens of kilometers could be reached by utilizing indium gallium arsenide (InGaAs)/indium phosphide (InP) SPAD-based LiDAR systems [67, 68].

1.3 Thesis Contributions

This thesis aims to develop high-performance NIR/SWIR SPAD pixels to be incorporated in high-resolution, low-pixel pitch LiDAR cameras and multichannel QKD applications. To this end, the contributions of the thesis could be listed as follows:

- A novel implementation of SWIR InGaAs/InP SPAD was demonstrated, providing both low noise and high efficiency. Furthermore, this method eliminated the need for guard rings to occupy extra area in the SPAD pixel, significantly improving the fill factor and overall sensitivity of the detector.
- A simulation environment was developed to estimate the device performance of InGaAs/InP SPADs, analyze limiting factors of their performance, and reduce fabrication iterations.
- Double diffusion-based InGaAs/InP SPADs were designed and characterized to reduce pixel pitch and optimize device performance.
- NIR efficient Silicon SPADs were developed in 55 nm Bipolar-CMOS-DMOS (BCD) technology by optimizing the doping profile, thereby significantly improving the collection efficiency of the device. The underlying physical phenomenon of this improvement was analyzed with simulation results.

1.4 Organization of the Thesis

Chapter 2 explains the principle of operation in SPADs and explains the trends in Silicon and SWIR SPAD development, together with SPAD's figures of merits. Chapter 3 presents the novel Selective area growth (SAG)-based SPAD depicting its structure, characterization results, and considerations for pixel size shrinking. Chapter 4 describes the simulation environment to estimate the noise and efficiency of an InGaAs/InP SPAD. Simulation results are compared with the measurement results to test the accuracy of the environment. Design, fabrication, and characterization details of a standard double-diffusion-based InGaAs/InP SPAD are reported in Chapter 5. In Chapter 6, four different SPAD implementations in 55 nm BCD technology to achieve high NIR efficiency are discussed by demonstrating their design considerations and performance results.

2 Background

Signal amplification at the receiver side of an optical detection system becomes necessary where the unamplified signal level stays below the noise level of the receiver electronics [69]. In photodiodes, this amplification could be achieved by the impact ionization phenomenon. This effect occurs when a carrier attains enough energy under a high electric field and collides with a lattice atom to release other free electron-hole pairs. Studies have been performed as early as the 1950s to understand this effect [70–72]. It is followed by theories explaining fluctuations in the multiplication gain and conditions of avalanche breakdown establishing the fundamentals of avalanche photodiodes [73, 74]. Finally, McIntyre and Webb suggested single-photons could be detected based on impact ionization [75]. Single-photon avalanche diodes were made commercially available as early as 1989 [76].

2.1 SPAD operation

2.1.1 Geiger mode operation

PN junction photodiodes have two distinct modes of operation in reverse bias. Linear mode is where the device operates below the breakdown voltage. In this mode, the photocurrent generated by the photodiode is linearly proportional to the number of incident photons. The spectral responsivity of the device in linear mode is defined as: $R = \frac{\lambda\eta}{hc}qg$ where λ is the wavelength, η is the quantum efficiency, h is Planck's constant, c is the speed of light, q is the electron charge, and g is the current gain [77]. Responsitivity has a unit of Ampere per Watt and it is the metric indicating the efficiency of the photodiode in linear mode. If the device operates at a low reverse bias (few Volts), the gain is unity and the maximum responsivity of the device is defined as $R_{max} = \frac{\lambda}{1.24}$ where λ is in micrometers. Photodiode operation under varying reverse bias is illustrated in Figure 2.1. As reverse bias increases, the electric field in the depletion region becomes higher allowing carriers to acquire enough energy to break bonds when they collide with a bonded electron. It gives rise to the gain from unity, and the device is called an avalanche photodiode (APD). In the APD mode, there is a finite amount of secondary electrons that could be created resulting in a finite gain. There is a voltage

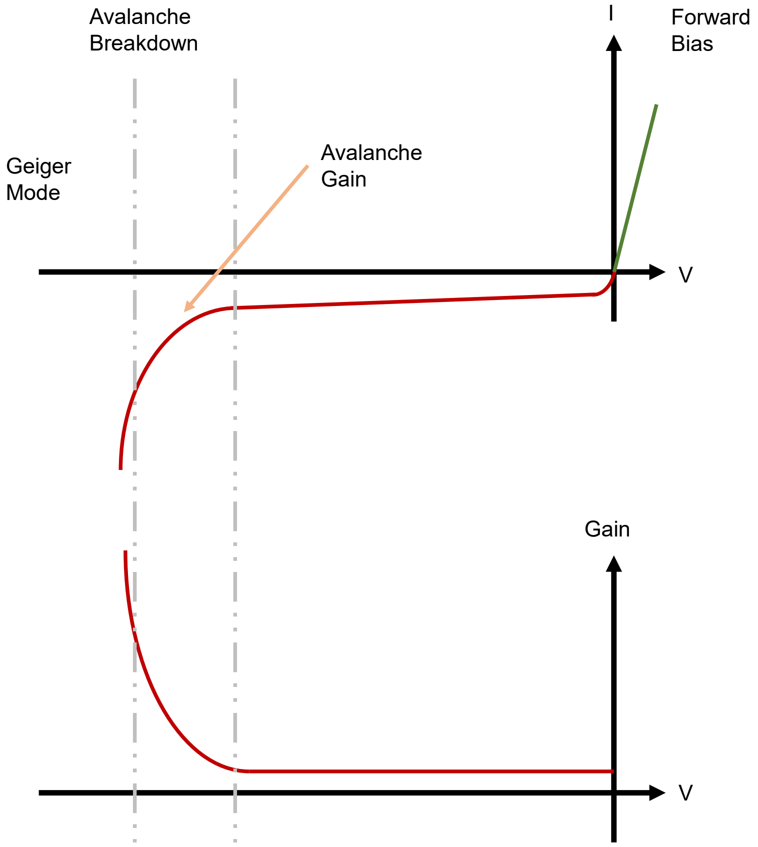


Figure 2.1: Various modes of operation of a pn junction diode. The diode is on during forward bias. At reverse bias, gain stays unity until avalanche multiplication takes place. It rises to infinity above the breakdown voltage.

point above which a single carrier causes a so-called self-sustaining avalanche process. Gain becomes infinite at this point. Without special care, the current tends to diverge to infinity causing the device to heat up leading to its eventual destruction. Assuming only electric field-dependent ionization rates for electrons and holes, McIntyre developed an equation that describes the gain (Equation 2.1).

$$M(x) = \frac{\exp\left(-\int_x^w (\alpha - \beta)\,dx'\right)}{1 - \int_0^w \alpha \exp\left(-\int_{x'}^w (\alpha - \beta)\,dx''\right) dx'}, \tag{2.1}$$

where x stands for electron-hole generation position within the depletion region, w is the depletion width, α and β are impact ionization coefficients for electrons and holes respectively. This equation also establishes the breakdown voltage point that makes the denominator of the Equation 2.1 zero.

If the avalanche current is not quenched, the device could get damaged due to the heat generated. A series-connected quenching resistor could implement a very simple quenching mechanism. The high current develops a voltage on the resistor. As a consequence, the diode voltage decreases to near breakdown voltage. Unlike linear mode photodiodes, that operate below breakdown, SPADs operate in a dynamic regime divided into three distinct phases: avalanche, quench, and recharge. The SPAD is initially in an idle state. In this state, there are no carriers initializing the avalanche multiplication process. When a photon or a dark carrier triggers an avalanche, a macroscopic current starts to build up quickly. The same current flows through the resistor as well generating a voltage close to excess bias voltage on the resistor. Once the voltage on the SPAD reaches a voltage below breakdown, the avalanche is said to be quenched. During the recharge phase, in a passively quenched SPAD, further trigger events can occur. If there are no other incidents in the recharge phase, the voltage on the quenching resistor decays with a time constant depending on the load capacitance [78, 79]. Details of quenching and recharge are analyzed in the following section.

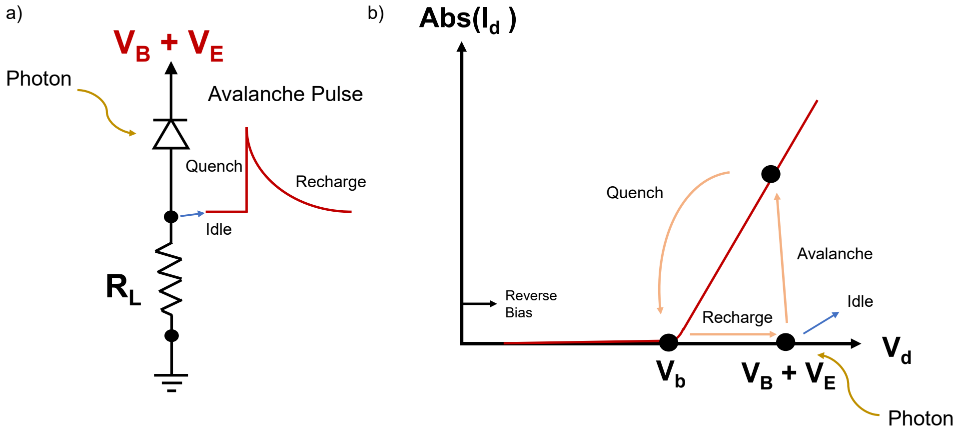


Figure 2.2: a) A SPAD with passive quenching. When the SPAD is triggered a voltage pulse is generated at the resistor. b) Schematic shows dynamic operation modes of a SPAD. Prior to trigger, SPAD stays in an idle condition. Avalanche current flows and the voltage developed on the resistor put SPAD around breakdown voltage. In the absence of further triggers, SPAD recharges to idle point.

2.1.2 Quenching & Recharge

Quenching and recharge are two key processes in SPAD operation. Quenching ensures that the SPAD is not damaged due to the heat generated by the avalanche current. The recharge process resets the SPAD operation and sets it ready for detecting the following photon. Passive quenching is the most straightforward way to achieve quenching. This is illustrated in Figure 2.3, together with an equivalent circuit of a SPAD [80]. R_D and C_D (around a few pF) stand for SPAD resistance and junction capacitance. McIntyre lists the contributors of the R_D as spreading resistance, space-charge resistance, and thermal resistance [81]. R_D can have a value of a few hundred ohms up to a few kilo-ohms for a SPAD diameter range of 10 μm to 50

μm [80]. C_L represents the load capacitance associated with the counting circuit used to read out the SPAD pulse. Using discrete components could add up to the parasitic capacitance of this node while a monolithic integration would be the ideal approach to reduce C_D.

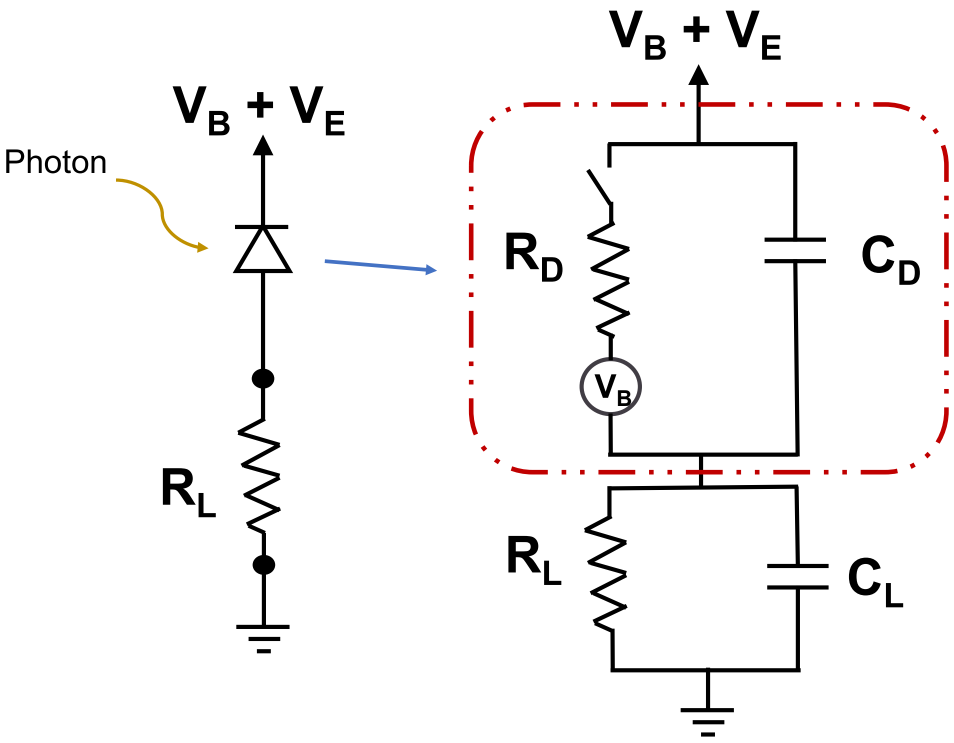


Figure 2.3: Passive quenching (left). SPAD model (right).

Figure 2.4 shows the change of the SPAD model during quench and recharge phases together with the evolution of avalanche pulse in time. In the quenching phase, the switch in the SPAD model closes. A large current $I_d, max = \frac{V_E}{R_D}$ starts to flow through the diode which has an asymptotic value of:

$$I_d(t \rightarrow \infty) = I_f = \frac{V_E}{R_L}. \tag{2.2}$$

Similarly, the voltage at the load node tends to increase to:

$$V_L(t \rightarrow \infty) = V_f = R_L I_f = R_L \frac{V_E}{R_L} = V_E. \tag{2.3}$$

However, there is a current value below which avalanche is not self-sustaining anymore. This current level is not very well defined and can cause variations at the peak value of V_L [80]. If the value of R_L is not sufficiently high value of I_f can surpass I_q. This can cause the avalanche to not quench properly, resulting in varying pulse times or V_L staying at a constant voltage level. A rule of thumb of 50 kohm per excess bias voltage is proposed in [80]. By considering the equivalent circuit seen by the node V_L, the rise time of the quenching phase could be calculated as:

$$\tau_q = R_D \parallel R_L(C_D + C_L) = R_D(C_D + C_L), \tag{2.4}$$

since $R_L \gg R_D$. As a conclusion if a large enough quenching resistor is used, the quenching time of the avalanche pulse should not depend on R_L.

Initial conditions for the recharge phase are illustrated in Figure 2.4 (c). In this phase switch of the SPAD model opens up and capacitor C_L discharges. The time constant associated with the recharge phase could be written as follows:

$$\tau_r = R_L(C_D + C_L). \tag{2.5}$$

One should note that, with passive quenching, the SPAD will be above breakdown voltage during recharge. The effective excess bias will be lower while C_L discharges but the SPAD could be triggered. Therefore, over-illuminating the SPAD in passive quenching could result in an almost constant DC voltage output at the load node.

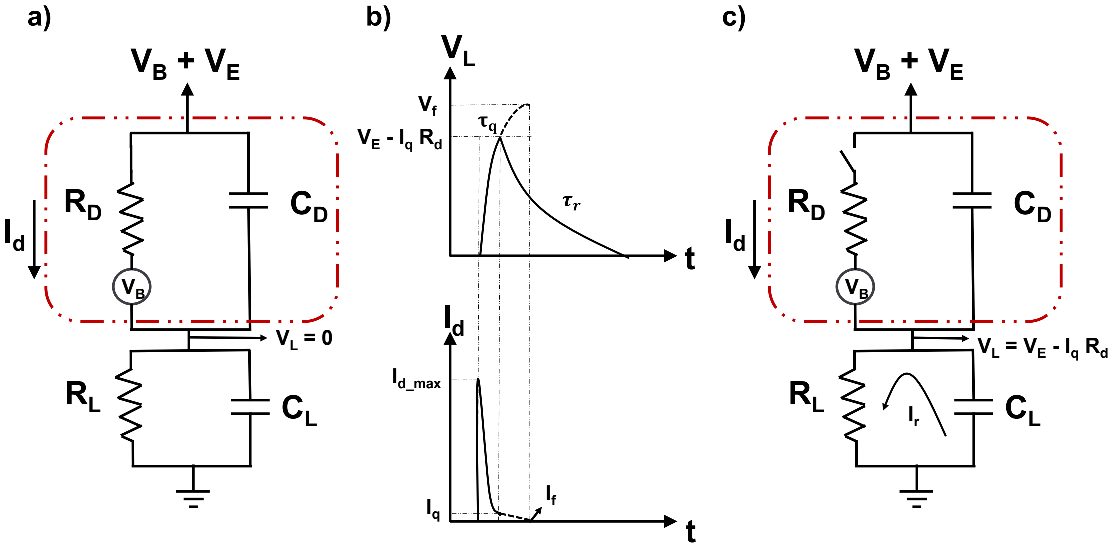


Figure 2.4: a) SPAD in quench phase. b) Load voltage and diode current evolution of avalanche pulse. c) SPAD in recharge phase.

2.2 Silicon SPAD Technology

Silicon SPADs are highly viable for applications in visible and NIR ranges thanks to their bandgap and well-established fabrication techniques. Their implementation methods could be divided into two being custom and standard technologies [82]. Custom silicon technologies refer to specialized semiconductor fabrication processes tailored to specific applications or performance requirements. These processes often involve unique structures and manufacturing techniques to achieve desired electrical, optical, or mechanical properties. In particular, for Silicon SPADs, custom technologies allow for implementations of very wide depletion regions. Reach through type Silicon SPAD is one of the early structures incorporating a very wide depletion. They usually utilize n+, p, p-, p+ type junction with having the n+ contact on the surface and p+ on the rear side of the substrate. The high field required for avalanche multiplication is formed between the n+ and p layers while the p- region is kept fully depleted to collect carriers from a wide area. This architecture significantly boosts NIR efficiency (40 % at 900 nm wavelength [82]) at the cost of a quite high breakdown voltage [83] and power dissipation. Planar-type implementations provided better control over the depletion region, thus reducing breakdown voltage to 45 - 55 V with 20% efficiency at 900 nm [84].

Standard technologies allow for monolithic systems, such that the SPAD and the read-out electronics, including quenching and recharge, could co-exist on the same chip. This enables designing small pixels, and large SPAD arrays up to several megapixels, and minimizes the parasitic capacitance between SPADs and read-out circuits [64, 85, 86]. The main limitation of working with standard technologies is the limited access to process modifications. The designer has to work with available doping layers and epi-silicon thickness. This imposes a fundamental limit on available SPAD structures. Design methods using Standard technologies could be further divided into two as: isolated and non-isolated. The former utilizes a deep or buried n-well to form a reverse p-n junction between the SPAD and p-type epi silicon, thus preventing the carriers from epi from penetrating the avalanche region. This approach limits the photo collection area and reduces NIR efficiency. However, since the photocarriers generated in the substrate cannot enter the avalanche region through random walk, the diffusion tail is significantly improved. Moreover, only the dark carriers generated inside the area defined by the n-well can trigger an avalanche, as a result, SPAD noise decreases [87]. In non-isolated designs, the n+ contact is placed at the center of the device and the avalanche region is formed at a junction between an n-well and the p-type epi silicon. With this method, peak PDPs up to 72% have been demonstrated thanks to larger photo collection areas [88, 89].

2.3 SWIR SPAD Technology

The first parameter to consider when designing a device for a certain wavelength range is the absorption coefficient of the semiconductor material. Even the high-quality Silicon fabrication technologies have been highly matured over many years, Silicons' efficiency is quite low in NIR and almost negligible in SWIR.

Materials

Unstrained Silicon has a 1.12 eV bandgap and can absorb photons up to 1100 nm wavelength (Figure 2.5). This pushed researchers to investigate alternative semiconductor materials with a lower bandgap. Germanium (Ge) and InGaAs are the main materials that could address this requirement. Germanium being an indirect bandgap semiconductor material like Silicon, exhibits a lower absorption coefficient in SWIR compared to InGaAs. On the other hand, Germanium could be monolithically integrated on a Silicon substrate together with read-out electronics, potentially promising a lower fabrication cost. However, the lattice mismatch between Silicon and Germanium poses challenges to their integrated fabrication. This delayed the emergence of low-noise Germanium SPADs, even today they need to operate at a lower temperature to reach the same noise level as their InGaAs counterparts [90].

InGaAs is a ternary semiconductor and has been widely used for fiber-based applications for detecting photons at 1550 nm wavelength. It is a direct bandgap material and demonstrates a high absorption coefficient from 1700 nm to visible wavelengths. It is grown at a compound ratio of $\mathrm{In_{0.53}Ga_{0.47}As}$, so that it is lattice-matched to InP substrate. The bandgap of $\mathrm{In_{1-x}Ga_xAs}$ could be calculated by the following equation [91]:

$$E_g(x) = 0.36 + 1.064x. \tag{2.6}$$

One could also adjust the bandgap of InGaAs by adding phosphorus to the compound. This allows InGaAsP to have a bandgap between 1.32 eV (bandgap of InP) and 0.74 eV (bandgap of InGaAs) while still being lattice-matched to InP.

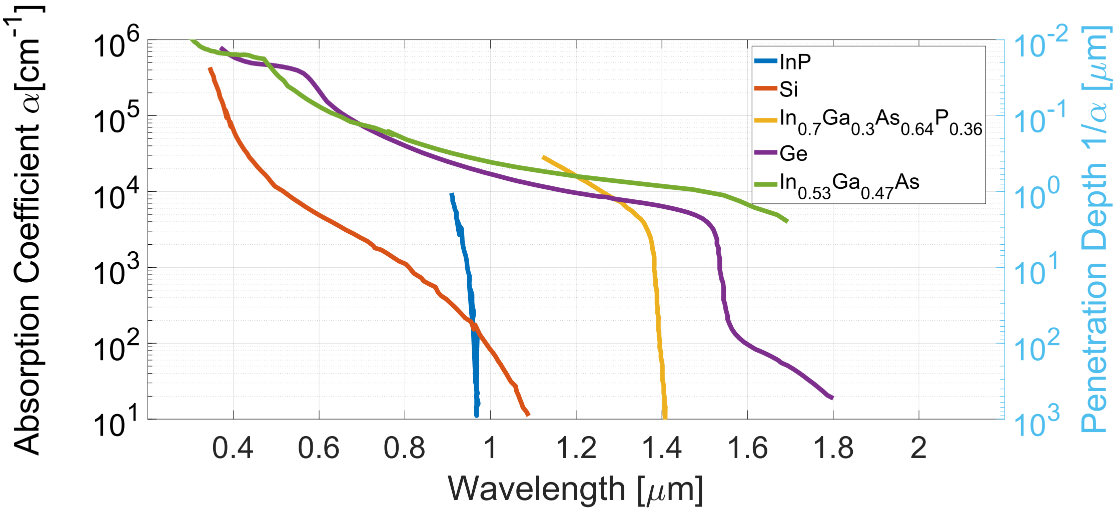


Figure 2.5: Absorption coefficients of various semiconductors [92, 93].

Device Structures

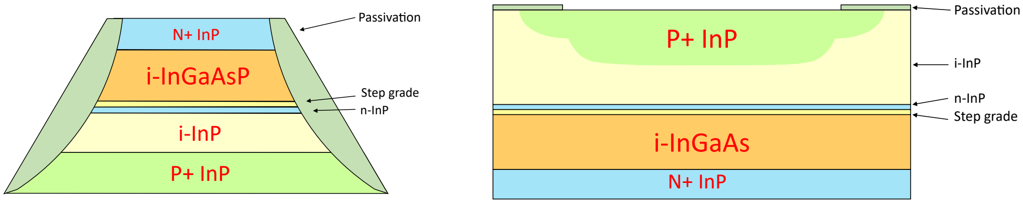
InGaAs p-i-n photodiodes have historically exhibited a non-abrupt breakdown characteristics, yielding limited amplification capabilities within the range of 10-100 when operated in linear APD mode [94, 95]. Subsequent research attributed this phenomenon to tunneling-facilitated carrier generation within the InGaAs substrate under the influence of elevated electric fields. Recognizing this as a consequence of inherent material properties necessitated a reconsideration of diode design. Nishida *et al.* addressed this limitation by introducing an innovative layered architecture, successfully increasing the device's gain to an impressive 3000 [96]. The proposed structural revision involved the spatial separation of the absorption and multiplication zones by repositioning the p-n junction inside InP. Within this configuration, the requisite high electric field for triggering impact ionization is localized at the junction within the InP, while carrier photogeneration is confined to the InGaAs region. These carriers are subsequently swept into the multiplication area, where the avalanche is initiated. This architectural concept has been called a separate absorption and multiplication (SAM) type diode.

Further enhancements to the SAM design include the integration of grading and charge control layers. A disparity in the valence band energies between InP and InGaAs induced a charge accumulation effect within the SAM diode, which was mitigated by the incorporation of a gradient layer. This layer involved a compositional transition within the ternary and quaternary materials, optimizing the lattice constant to align with InP. The preferred material for this gradient layer is the indium gallium arsenide phosphide (InGaAsP) quaternary compound, which notably allows for bandgap tuning from 0.74 eV to 1.34 eV while maintaining a consistent lattice match with InP. By precisely modulating the compound ratios, a seamless valence band transition is facilitated, eliminating the initial charge buildup issue.

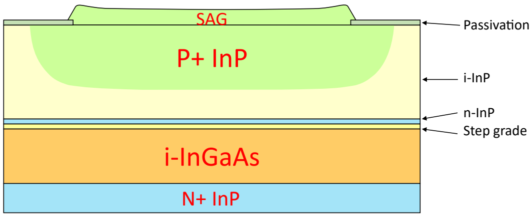
The final refinement introduced to the SAM structure is the addition of a charge layer. As previously noted, a substantial electric field within the InGaAs layer promotes pronounced tunneling-assisted carrier generation. The interposition of a charge layer between the InGaAs absorption region and the InP multiplication zone effectively modulates the electric field, depending upon the layer's thickness and doping concentration. This configuration enables control over the electric field within the InGaAs layer, thereby optimizing device performance. The final device structure is often referred to as the separate absorption-grade-charge-multiplication (SAGCM) type.

In order to define an isolated active area two main strategies could be followed: mesa and planar. Figure 2.6 shows these techniques. In mesa-type structures, etched trenches are implemented to electrically isolate the device from the surrounding devices. This approach is relatively easier to implement compared to the planar approach which requires optimization for the doping process. However, sidewalls formed due to trench etchings could behave as noise sources due to the dangling bonds at the sidewalls of the mesa. Although passivating the surface by depositing a dielectric material could reduce noise generation at the surface, it is a highly challenging process. Another fabrication issue that needs consideration is the shape or texture of the sidewalls. While it may be less critical for a linear photodiode, SPADs could suffer from premature edge breakdown. Therefore, the sidewalls of the mesa have to avoid forming a jagged surface profile to eliminate local high electric field regions.

Planar-type structures are implemented by p doping in a certain area of the intrinsic top layer. Since the intrinsic layer exhibits high resistance, neighboring pixels are electrically isolated.



(a) Mesa (b) Planar Double Diffusion



(c) Planar Single Diffusion

Figure 2.6: InGaAs/InP SPAD types in mesa and planar.

In general, the planar approach utilizes the double-diffusion method to prevent premature edge breakdown while isolating the devices. The shallow diffusion region reduces the radius of curvature of the deep diffusion region in the pursuit of eliminating a high electric field at the device periphery. Nonetheless, the diffusion process in InP is a challenging step and requires careful calibration. For the double diffusion approach, the second diffusion process also causes dopants to move due to its high-temperature nature which significantly enhances the complexity of the overall process. The single diffusion method replaces one of the diffusion steps with a masked epi-layer growth which is simpler to control. Selective-area growth is performed prior to the P-diffusion step to reduce the radius of curvature of the doping profile. Details of the SAG approach are discussed in the Chapter 3.

2.4 SPAD Figures of Merits

2.4.1 Noise

Noise in optoelectronics devices establishes the minimum optical signal level that could be resolved. In SPADs, noise manifests itself as false counts, i.e. counts not photogenerated. Shot noise of these false counts could overshadow photocounts, consequently putting a minimum detectable photon level. Under the assumption that background light is not the dominant false count generation mechanism, the noise of a SPAD device is established by dark count generation in the semiconductor. Dark counts are divided into two main categories; primary and secondary dark counts. Primary dark counts result from the generation mechanism of thermally generated carriers, while afterpulsing is responsible for secondary dark counts.

Dark count rate (DCR)

There are various primary dark carrier generation mechanisms in semiconductors, such as Auger, band-to-band tunneling, trap-assisted tunneling (TAT), and trap-assisted thermal generation also called Shockley-Read-Hall (SRH) generation rate. The former two are rarely reported in SPAD devices as significant contributors and are excluded in the context of this work. The main mechanisms contributing to the DCR are illustrated in Figure 2.7.

Alamo *et al.* argues that thermal excitation directly from the valence band to the conduction band is very unlikely because of the high energy barrier [97]. An energy level introduced at the forbidden bandgap region due to a defect or impurity could significantly enhance thermal generation probability. These trap levels could capture carriers which could later be released and contribute to conduction band. Although this mechanism is similar to the energy levels introduced by dopant atoms, associated energy levels are different. Energy levels introduced by dopants are very close to either conduction and valence band while, trap-assisted thermal generation is mostly affected by the impurities that form energy levels at the midgap. The model to estimate trap-assisted thermal generation rate is given by [98]:

$$G_{srh} = \frac{n \cdot p - n_i^2}{\tau_p \cdot \left(n + n_i \cdot e^{\frac{E_{trap}}{kT}}\right) + \tau_n \cdot \left(p + n_i \cdot e^{\frac{-E_{trap}}{kT}}\right)}, \tag{2.7}$$

where n, p are the carrier concentrations, n_i is the intrinsic carrier concentration, τ_n, τ_p are the carrier lifetimes, k is the Boltzmann constant, T is the temperature, E_{trap} is the energy difference between trap and Fermi level, and G_{srh} is the dark carrier generation rate. Carrier lifetimes correspond to the average time a carrier can travel before recombination. SRH generation is more significant in the depletion region since the energy difference between a trap and the Fermi level is higher outside. In contrast to tunneling generation, SRH generation depends strongly on temperature. Therefore, for a device with SRH dominant noise at room temperature, cooling could improve the device's performance.

Under high reverse bias, conduction and valence energy bands bend significantly due to the high electric field. This bending reduced the energy barrier for covalently bonded electrons to escape. Band-to-band tunneling becomes important at field values of $7\text{x}10^5\,\frac{\text{V}}{\text{cm}}$ [99]. This is usually not the case for SPADs with a breakdown voltage larger than 15V. Trap-assisted tunneling on the other hand takes place when a carrier tunnels with the assistance of trap energy level. Therefore, it is more probable to happen and could contribute to the DCR of the device. It is a quite complicated process to model and different models are proposed in the literature. Hurkx and Schenk proposed their own version to modify the lifetime in the SRH generation equation such that lifetimes depend on the local electric field [99, 100].

$$\tau_{n,p} = \frac{\tau_{dop}}{1 + g_c(F)}. \tag{2.8}$$

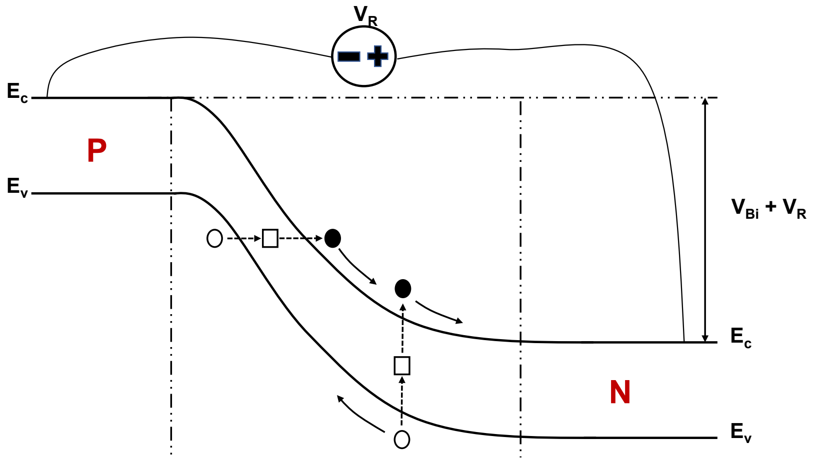


Figure 2.7: Dominant primary dark carrier generation mechanisms in SPADs. Trap-assisted tunneling (left) and Trap-assisted thermal generation (right) are shown.

Increasing the electric field lowers the lifetime. Thus, enhancing the generation rate. These models are relatively easy to implement in commercial simulation tools. Donelly *et al.* proposed another method to calculate the tunneling current density for InP as [101]:

$$J_{\text{tun-Av\_def}}(x) = \frac{AF(x)^2 N_T \exp\left(-\frac{B_1 E_{B1}^{3/2} + B_2 E_{B2}^{3/2}}{F(x)}\right)}{N_v \exp\left(-\frac{B_1 E_{B1}^{3/2}}{F(x)}\right) + N_c \exp\left(-\frac{B_2 E_{B2}^{3/2}}{F(x)}\right)}, \tag{2.9}$$

where $F(x)$ stands for electric field at position x, E_g is the bandgap, $A = \frac{q^3(2m_r/(qE_g))^{1/2}}{(4\pi^3\hbar^2)}$, $B_1 = \frac{\pi(m_{\text{th}}/2)^{1/2}}{(2q\hbar)}$, $B_2 = \frac{\pi(m_{\text{c}}/2)^{1/2}}{(2q\hbar)}$, $E_{\text{B1}} = aE_{\text{g}}$, and $E_{\text{B2}} = (1-a)E_{\text{g}}$ represents tunneling barrier heights from valence band to trap and trap to the conduction band,N_{v} and N_{c} are light hole valence and conduction band density of states, N_{T} is trap density, m_{r}, m_{th}, and m_{c} are reduced effective mass, light hole valence effective mass and conduction band effective mass respectively, h is Planck's constant. Tunneling generation has a weak dependence on the temperature. Devices with a dominant DCR contribution from SRH at room temperature become tunneling-limited at lower temperatures. Once this condition is met, further cooling the device does not reduce DCR significantly.

Afterpulsing

Different from primary dark counts, afterpulsing is a secondary, correlated noise source. When the DCR of a device is measured, one cannot distinguish the contributions from primary dark counts and afterpulsing. However, since afterpulsing is a correlated source, it is possible to quantify the two sources by looking at the inter-avalanche timing histogram of dark pulses. The correlated nature of afterpulses will result in a deviation from Poissonian distribution [102].

Once a dark carrier or photon triggers an avalanche and impact ionization processes take place, a very large number of carriers flows through the avalanche region. Some of these carriers are captured by the deep energy level introduced by traps. Finally, these carriers are released after a certain lifetime that depends on trap energy levels, temperature, and the local electric field [103–105]. The majority of the published work assumes discrete energy levels and associated decay lifetimes to model afterpulsing probability over time [103–109]. This model assumes afterpulsing probability over time respects the model from Equation 2.10.

$$P_{ap}(T_{h0}) = A_0 + A_1 e^{-\frac{T_{h0}}{t_1}} + A_2 e^{-\frac{T_{h0}}{t_2}} + A_3 e^{-\frac{T_{h0}}{t_3}} + \ldots, \tag{2.10}$$

where A_0 is the primary dark count rate, A_i is the coefficient of exponential, and t_i is the lifetime for ith trap. However, Itzler *et al.* claims results obtained by this approach highly depend on the measured time range and time constants have no physical relevance [110]. The aforementioned work suggests using broadband trap energy distribution instead of discrete trap levels.

2.4.2 Single-photon Sensitivity

Photon detection probability (PDP) is the parameter that measures the single-photon sensitivity in a SPAD. PDP reflects the number of pulses a SPAD will generate, on average, when subjected to a certain amount of photons. It is a function of wavelength since the SPAD sensitivity strongly depends on the wavelength of light. In order to analyze the factors impacting PDP a simplified one-dimensional Silicon device is considered. A cross-section of this simplified device is shown in Figure 2.8. Under this assumption, PDP could be modeled by the equation [82]:

$$PDP = \int_0^{z_\infty} (1-R)\eta_{abs}(z) P_{coll}(z) P_b(z)\,dz, \tag{2.11}$$

where R stands for reflection, $\eta_{\text{abs}}(z)$ is the absorption density, $P_{\text{coll}}(z)$ is the collection probability, and $P_{\text{b}}(z)$ is the breakdown probability of a carrier generated at depth z.

Reflection

When photons arrive at the device surface, some portion of the photons immediately reflect from the surface without contributing to the absorption process in Silicon. If the dielectric stack only consists of a silicon dioxide layer, Fresnel equations could be used to calculate the reflection at the Silicon/Silicon dioxide interface [102]. If the thickness of the Silicon dioxide is close to the wavelength of light, there will be an interference pattern on the reflection spectrum. This pattern will also manifest itself on the PDP curve. In modern CMOS process, many dielectric layers are used to isolate different metal levels. If they are not designed carefully, they could enhance the reflection and reduce PDP. In CMOS image sensor (CIS) processes, the dielectric layers are usually optimized to achieve maximum transmission by utilizing, among others, an anti-reflection coating as a final dielectric layer.

Absorption Density

Absorption density is a measure of the optical generation rate within Silicon. It is usually modeled by Beer-Lambert law:

$$\eta_{\text{abs}}(z) = \alpha e^{-\alpha z}, \tag{2.12}$$

where z is the depth position with respect to the surface, and α is the absorption depth. As shown in Figure 2.5, the absorption depth depends on the wavelength and it has a varying behavior for different semiconductors. In the case of Silicon, the absorption coefficient has a subtle increase trend towards low wavelengths due to its indirect bandgap. Penetration depth is the inverse of absorption coefficient and it helps to put it into a clearer perspective. Penetration depth indicates the depth that around 64% of the incident photons (after reflection) are absorbed. It is around 100 nm at 400 nm wavelength for Silicon. Thus, most of the 400 nm photons will be already absorbed at very close to the Silicon surface. To boost the efficiency around this wavelength, the device design should be made in a way that carriers near the surface could drift or diffuse to the avalanche region. As we move forward towards longer wavelengths, penetration depth gets even smaller. Eventually, the carriers will recombine at the dielectric/Silicon interface establishing a lower limit for PDP at smaller wavelengths. At the other end of the spectrum, if we look at the 800 nm wavelength with a 10 μm penetration depth, its absorption takes place at a quite long distance. Those carriers could be collected by a wide depletion, drift region, or diffusion region.

Collection Probability

Collection probability is a valuable concept to analyze if a carrier could reach the avalanche region. Figure 2.8 highlights four distinct regions. Each region has a different mechanism affecting the ability of a carrier to move towards the high-electric field region. Both electrons and holes generated in the avalanche region have a unity collection probability since they are already in the high-field region. The electric field in the drift region pushes carriers towards the avalanche region, giving them again a unity collection probability. Holes in the drift region are pushed away from the avalanche region resulting in a sharp collection probability decrease for holes at the drift/avalanche interface. In the diffusion region, the electric field is nearly zero. The carriers in this region perform a random walk until they recombine or enter the avalanche region. The distance carriers could travel before the recombination is governed by the recombination lifetime. It translates to diffusion length by the Equation 2.13 [91].

$$L_{p,n} = \sqrt{D_{p,n}\tau_{p,n}}, \tag{2.13}$$

where $D_{p,n}$ is the diffusion coefficient, and $\tau_{p,n}$ is the recombination lifetime for holes or electrons. Lifetimes depend on the doping of the region. In Silicon, for a doping larger than $3x10^{19}$ cm^{-3}, hole lifetime can be smaller than 1 ns [111]. If the diffusion length is much larger than the distance between the surface and the junction, this mechanism could be ignored [102].

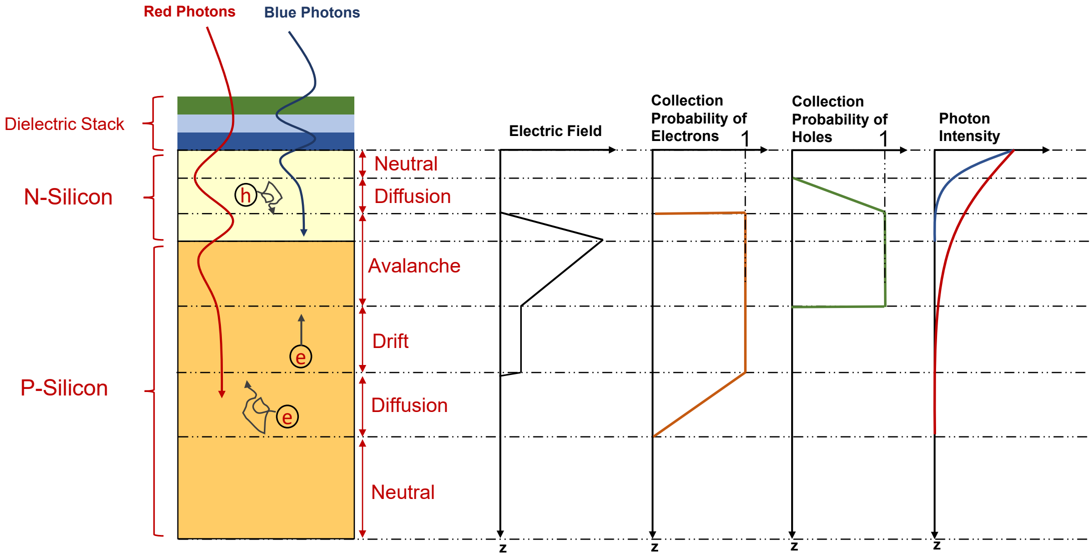


Figure 2.8: A non-isolated 1D SPAD with different regions contributing to the PDP are highlighted (left). Representative sketches of the electric field, collection probability, and photon intensity are shown (right).

Breakdown Probability

The probability that a carrier generated at a specific location in the device will trigger an avalanche is modeled by Equation 2.14 [82, 112]. This equation assumes that the triggering probabilities of holes and electrons are independent. Further details of breakdown probability calculation are discussed in Chapter 4.

$$P_b(z) = P_e(z) + P_h(z) - P_e(z)P_h(z). \tag{2.14}$$

2.4.3 Timing Jitter

Timing jitter indicates how precise the arrival time of the photon can be measured by a given SPAD. It is usually measured by a pulsed picosecond (or femtosecond) laser directed to the

SPAD after proper attenuation to suppress pile-up and a fast discriminator to compute the time difference between the avalanche pulse and the laser clock. The measurement is repeated a large number of times to construct a histogram of the time response. The Full width at half maximum (FWHM) or sigma of this histogram is a common metric to compare the timing performance of different SPADs. Even in the ideal case that laser clock jitter and optical pulse width are too small, the avalanche pulse exhibits fluctuations in time giving rise to timing jitter. Various sources of timing fluctuation contributors to overall detection time could be listed as follows [113]:

$$\sigma^2(T_{\text{det}}) = \sigma^2(T_{\text{drift}}) + \sigma^2(T_{\text{diffusion}}) + \sigma^2(T_{\text{build}}) + \sigma^2(T_{\text{spread}}), \tag{2.15}$$

where T_{diff} and T_{drift} are arising from the time needed for a carrier to reach the avalanche region while T_{build} and T_{spread} are related with the time required for the avalanche current to reach macroscopic levels once it is initiated. Note that all jitter contributions are summed in quadrature, assuming that they are statistically independent random variables.

As discussed in the previous section, the parts of a hypothetical SPAD could be divided into four distinct regions. Avalanche and drift regions are depleted and a high electric field exists in these regions. Photocarriers generated in these regions get quickly accelerated. Depending on the position of the carrier, the time it takes to initiate an avalanche event varies resulting in T_{drift}. This process contributes to the Gaussian shape in the timing jitter curves. A smaller junction width is expected to reduce the contribution from T_{drift} to overall SPAD jitter. However, it also decreases the PDP. $T_{\text{diffusion}}$ stands for the timing variations of the carriers generated in the diffusion region triggering an avalanche. Since the electric field to accelerate the carriers is absent in this region, it is a slower process compared to T_{drift}. This process forms an exponential tail in the timing jitter measurements and is usually referred to as diffusion tail. Once the avalanche is triggered, there are two main processes to consider; build-up and spread. The build-up is the phase in which carriers perform impact ionization parallel to the electric field. The number of carriers exponentially grows in a small area where the avalanche is initiated. Nevertheless, the current does not reach a level to be picked up by the read-out electronics [78]. The statistical nature of the impact ionization process causes variations in T_{build}. The final major process is the spread of the avalanche. Hot carriers generated during build-up result in a high concentration of carriers locally. These carriers start to diffuse laterally initiating the following build-up processes. Furthermore, hot electrons generated by the impact ionization could lose energy by photon emission. These photons are reabsorbed in the surrounding area helping the spread process. Authors have shown that avalanche pulse grows faster if it is triggered at the center by using a focused laser beam at the center and the periphery of the SPAD [114]. In the case of a flood illumination of the total SPAD area, photons initiate avalanches at laterally random locations. As a result, there will be varying delays in the avalanche pulse manifesting itself as T_{spread}. For InGaAs/InP SPADs a 13 ps better SPAD jitter is reported at 5V excess bias when the laser is focused to a 6 μm spot compared to flood illumination [115].

3 Selective-area-growth based InGaAs/InP SPAD

Efforts to enhance InGaAs/InP SPADs have focused on refining Zinc (Zn) diffusion profiles, layer dopings, and thicknesses [115–118]. Traditional designs utilize a standard double diffusion guard ring, limiting the active pixel area, and tend to have an increased electric field at the edges at lower excess biases. Alternatives include one-step diffusion InGaAs/InP SPADs through recess-etching [119, 120], though their performance falls short of double-diffusion variants. This chapter introduces a novel InGaAs/InP SPAD structure using Selective-Area Growth (SAG) [121], which prevents edge breakdown with a single diffusion process, enabled by the SAG epitaxy's surface topography. This design ensures a uniform electric field and the potential for a reduced pixel pitch without sacrificing the active area, promising better spatial resolution for larger pixel arrays. The SAG technique is reproducible, simpler in terms of diffusion depth calibration, and shows potential for next-generation LiDAR applications. The chapter initially demonstrates the obtained SPAD figures of merits from a 70 µm active diameter device, then discusses the performance change when the device diameter is reduced.

This work has been published in [122].

The author contributed to this work by performing device simulations, characterization, and data analysis.

3.1 Device Structure

The device's configuration incorporates a lattice-matched InGaAs semiconductor material for photon absorption and InP for avalanche multiplication based on SAGCM configuration, as depicted in Figure 3.1. The layer stack is composed of an approximately 1.8 µm thick i-InGaAs absorber and roughly 3 µm thick i-InP cap layers, where a 2 µm Zn diffusion process was completed. The 1.8 µm thick i-InGaAs layer enables the absorption of nearly all incident light passing through the absorber twice, facilitated by the back-side metallization. To maintain a low DCR without compromising device jitter, a 1 µm thick multiplication width was selected [123]. The i-InP cap layer was thickened adequately for Zn diffusion, thereby preventing sharp junction curvature and achieving a low field enhancement effect at the device edge. The approximately 0.1 µm thick n-InP charge layer was doped within the 2-2.5×10^{17} cm^{-3} range to maintain a 25 V ± 5 V difference between the punch-through and breakdown point, ensuring complete depletion of the absorber even at low temperatures. Following the epitaxial growth and Zn diffusion process, the device's backside was fully covered with Gold (Au) metal to form the cathode contact. Additionally, a ring-shaped, ohmic metallization of 5 µm width and 56 µm inner diameter was deposited around the device's periphery for the anode metallization. However, this configuration partially shades the device's active area with the ohmic metal, a design choice made for ease of testing with front-side illumination. Nonetheless, for eventual smaller diameter devices or pixelated arrays, a backside-illuminated geometry would be more advantageous to maximize the anode electrical contact area without causing optical shading.

In the case of standard InGaAs/InP SPADs, a technique involving two diffusion steps is applied to prevent early breakdown at the device edge. This entails a deep Zn diffusion to define the active area and a shallow diffusion around the periphery, which prevents abrupt changes in the depletion front. Numerous groups have found this approach effective in preventing edge breakdown. However, the expanded area covered by the shallow diffusion increases the device's size without enhancing its efficiency, leading to a less-than-optimal fill factor. In the proposed design, the prevention of edge breakdown is achieved through a single diffusion step using SAG. Before the Zn diffusion, an approximately 300 nm thick undoped InP layer is grown at the device's center through SAG. The thickness of this layer gradually increases from the center to the edge due to an accelerated growth rate [124], reaching approximately 700 nm at the device's edge. Zn diffuses through this tapered InP surface, ensuring a gradual transition of the doping profile inside InP towards the periphery. This ensures that the electric field around the device's edge remains lower than that at the center. The diameter of the diffusion window for this specific device is 70 µm. Additional information about the fabrication process

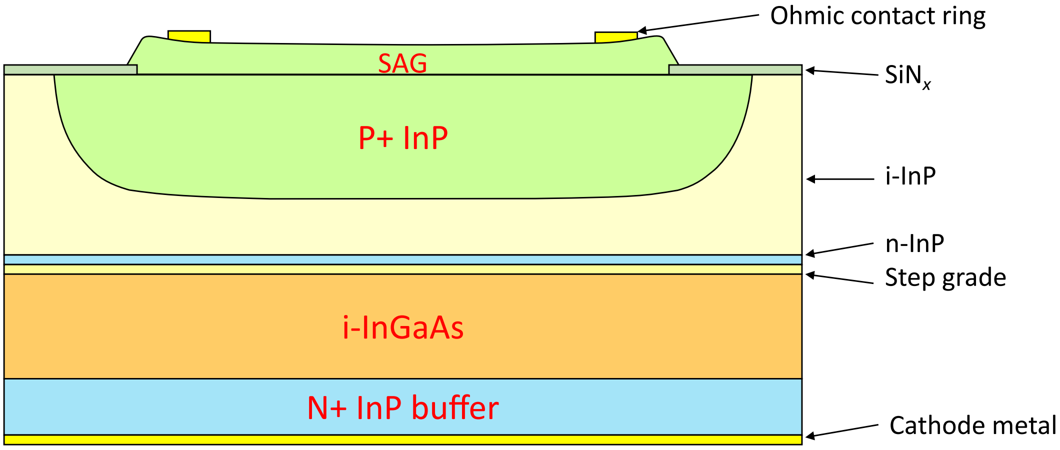


Figure 3.1: Schematic of SAG-based InGaAs/InP SPAD. Green parts highlight the InP doped by Zn. The image is not to scale.

for the device is provided in [121]. To test the operability of the same process conditions in Geiger mode and conduct a thorough analysis of the active area, this study is performed for large-diameter devices. The significant changes in the SAG's topography and resulting Zn diffusion depth profile occur over a distance of approximately 5 µm or less.

3.2 Device Simulations

The device's electrical simulations were conducted using the commercial software Synopsys Technology computer-aided design (TCAD). For Zn diffusion, charge, and grading regions, the doping values from the secondary ion mass spectrometry (SIMS) measurement in [125] were implemented. In order to accurately represent the Zn diffusion profile, the curvature of the Zn diffusion edge in the simulated device was adjusted to approximately match the curvature observed in the scanning electron microscope (SEM) images of the device from [121] by using error functions in TCAD. In terms of avalanche generation, the quasi-physical model from [101] was first used to calculate the impact ionization coefficients in InP. Subsequently, the Okuto-Crowell model [126] was employed in the TCAD environment for the avalanche process, with fitting parameters fine-tuned to align the ionization coefficients between [101] and TCAD under various electric fields. The simulation framework is explained more in detail in Chapter 4. In Figure 3.3, the observed 0.5 V difference between the measured and simulated breakdown voltage likely stems from variations in the doping of the charge region. Even minor fluctuations in this doping level, only a few percent, can noticeably impact the breakdown voltage. Furthermore, a slight variance between the simulated and actual multiplication width might have also contributed to this divergence. To obtain the I-V curve below the breakdown voltage, standard lifetime values for InGaAs [127] and InP [128] materials were utilized from the literature, typically within the order of microseconds. Additionally, a surface current was defined between the cap i-InP and SiN layers by introducing traps at the interface. These traps were assumed to be positioned at the mid-gap, with the trap concentration and cross-section used as fitting parameters to achieve a good match with experimental I-V results.

Figure 3.2 illustrates the device's electric field distribution at an excess bias of 5 V and a temperature of 300K. The diagram shows that the high electric field remains effectively contained beneath the diffused region. Furthermore, the deliberately designed diffusion boundary at the device's edge has successfully eliminated premature edge breakdown. Within the absorber region, the electric field demonstrates a linear gradient, owing to the n-type background doping. The electric field in the absorber is maintained below 150 kV/cm, which effectively prevents tunneling noise. At the interface of the InP buffer, the electric field reaches 15 kV/cm, ensuring complete depletion of the absorber [129].

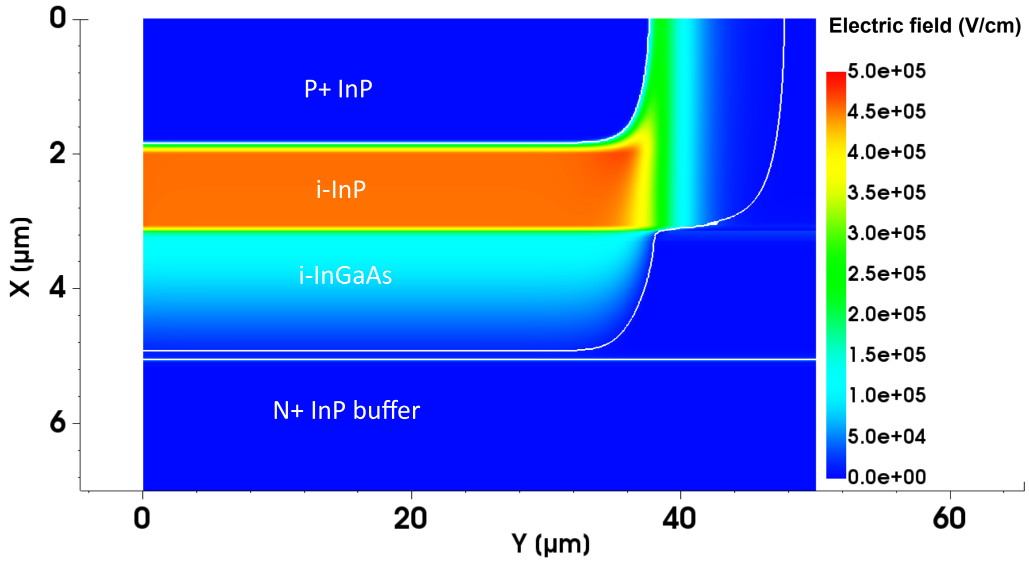


Figure 3.2: 2D electric-field simulation of half device at 300K and at 5 V excess bias.

3.3 I-V Characteristics

Figure 3.3 shows the I-V curves obtained using the Agilent source measurement unit (SMU) B2902A at room temperature, both under light and dark conditions. The device was illuminated with a 1550 nm laser to determine the punch-through voltage from the front side. As this

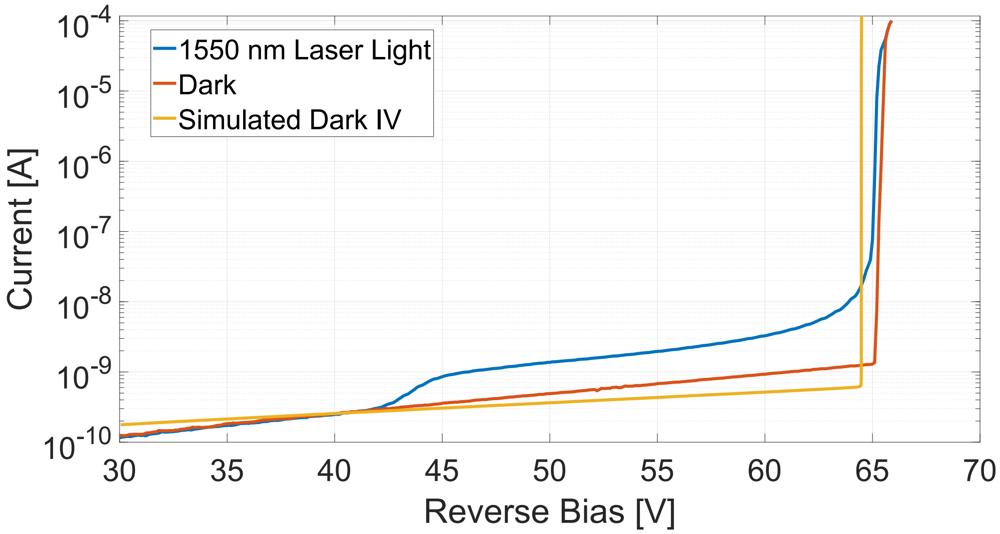


Figure 3.3: I-V measurements of SAG-based InGaAs/InP SPAD in dark and under 1550 nm wavelength light illumination together with simulated I-V curve at 300K.

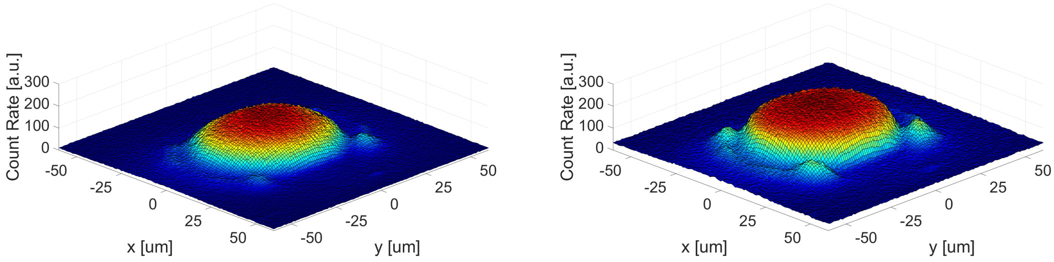
wavelength is not absorbed by the InP layer, the voltage at which photocurrent emerges shows the depletion of the InGaAs absorber region. The device demonstrates a distinct avalanche breakdown at 65 V and a punch-through voltage of approximately 43 V, ensuring the depletion of the absorber region across a wide temperature range.

3.4 Active Area Scan

In order to explore the uniformity of the electric field across the device, a 1550 nm laser beam was directed onto the device via a confocal microscope. The Mitutoyo apochromatic objective from Thorlabs is used as the focusing element. This objective has a resolution of 0.6 μm at 550 nm wavelength as provided by Thorlabs in the objective specifications. If we assume a linear relationship between resolution and wavelength, it translates to a beam size of around 1.7 μm at 1550 nm. The focused light beam traveled in a 120 μm × 120 μm window, with steps of 1 μm, and the SPAD's count rate was recorded at each beam position. Photomaps serve as an assessment of the electric field within the device, as the areas exhibiting a photoresponse indicate where avalanche multiplication occurs.

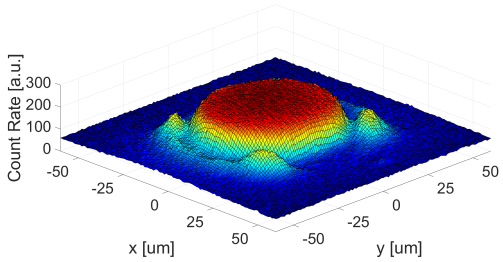
In addition to the device illustrated in Figure 3.1, another device with two floating guard rings (GRs) was also tested to evaluate GR impact on device performance. The cross-section of this device is displayed in Figure 3.5. Figure 3.4 showcases the measured 2D photomap of the device lacking GRs at excess biases of 3, 5, and 7 V, while Figure 3.5 presents the photomap for the device with two floating GRs at excess biases of 3 and 5 V. A notably flat response is observed at 3 V excess bias. However, at 5 and 7 V excess biases, irregular side lobes emerge for the device without a GR. The maximum sidelobe-to-center response ratios at 3 and 5 V excess biases are 55% and 70%, respectively. Nevertheless, in the device with floating GRs, the side lobes dominate the photoresponse at 3 V excess bias. At 5 V excess bias, the field at the sidelobe saturates, causing the peak responses at the center and the sidelobes to have a similar height. Consistent with the photocurrent maps referenced in [121], the sidelobes appear along the [100] and [010] crystal directions. The uniformity of the SAG surface surrounding the mask edge was disrupted along these directions, leading to deeper Zn diffusion and a higher electric field. The SAG process mitigates edge effects by leveraging growth enhancement at the device's edge, facilitated by the near-surface diffusion of growth precursors over the masked areas of the wafer [130].

In the floating GR devices, this growth enhancement is most prominent in the outermost open areas, namely the floating GRs, with less growth enhancement at the edge of the active area compared to devices lacking floating GRs. This may account for the comparatively less effective suppression of edge enhancement in the floating GR devices, leading to the observed irregular lobe features along specific directions at the device's edge. Finally, the effective active area is determined for the scan results in Figure 3.4 by integrating the regions with a response exceeding half the maximum. The calculated diameters based on the effective active area are 53 μm at 3 V, 59 μm at 5 V, and 63 μm at 7 V excess biases, respectively. The presence of partial



(a) 3 V excess bias

(b) 5 V excess bias



(c) 7 V excess bias

Figure 3.4: Active area scans of the device without GR under focused 1550 nm laser light at 300K.

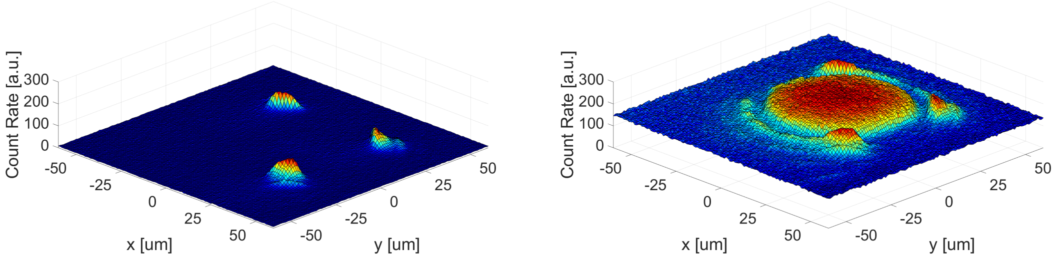
shading induced by the anode metallization is visible as a blue region in Figure 3.5b, situated between the central active region and the sidelobes.

3.5 Noise Performance

The noise floor of a SPAD is crucial in determining the lowest discernible photon rate that the device can detect.

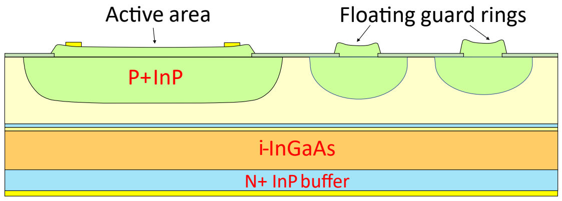
3.5.1 DCR

The primary DCR measurements were conducted at a gate-off time of 250 µs to remove afterpulsing contribution. Gating is performed by generating an excess bias voltage pulse from a waveform generator. This voltage is then added to the DC voltage (set 1 V below the breakdown voltage) with the help of a bias tee. The voltage across the quenching resistor is measured with an oscilloscope that also counts the number of pulses. Results for various excess biases within the temperature range of 225K to 300K are presented in Figure 3.6 and Figure 3.7. For the device lacking a GR, the DCR at 225K and 5 V excess bias is 20 kcps. As the temperature increases to room temperature, the DCR rises to 1.6 Mcps due to thermal generation in the absorber region. Conversely, the device equipped with a GR shows DCR values approximately one order of magnitude higher than those of the guard ringless device



(a) 3 V excess bias

(b) 5 V excess bias



(c) Schematic of the device with floating GR

Figure 3.5: Active area scans of the device with two floating GRs under focused 1550 nm laser light at 300K. The scan could not be performed at 7 V excess bias due to high DCR that is probably arising from the high electric field on the sidelobes. While there are also floating guard rings on the left side of the device, they are not shown in the schematic.

across all temperatures.

For a more detailed comparison of the DCR between the two devices, refer to Figure 3.8. Notably, at 3 V excess bias, while only small sidelobes are active in the device with floating GRs, its DCR matches that of the device without a GR. In essence, the DCR contribution from the small sidelobes is comparable to the DCR arising from the entire active area of the device without a GR. At 5 V excess bias, when the central regions of both devices are active, the DCR ratio between the two devices becomes approximately 4, indicating that a significant portion of the DCR in the device with a GR results from the enhanced electric field at the sidelobes. This aspect is further investigated through the analysis of activation energy curves.

If SRH generation dominates DCR, the following relation holds between DCR and temperature:

$$DCR \quad \propto \quad T^2 \cdot e^{\frac{-E_a}{kT}}, \tag{3.1}$$

where E_a is the activation energy and k is Boltzmann's constant [131]. Activation energies at 5 and 7 V excess biases are obtained by plotting $\ln(DCR/T^2)$ $vs.$ $1/kT$ and fitting a linear curve as in Figure 3.9.

At 5 V excess bias, the device without GR exhibits an activation energy of 0.37 eV, corresponding to half the bandgap of InGaAs. This measurement suggests that the primary mechanism contributing to DCR is the SRH generation within the absorber region, and this mechanism remains dominant down to 230K. In contrast, the device with a GR demonstrates an activation energy of 0.34 eV (less than 0.37 eV) at 5 V excess bias. This lower activation energy indicates that the DCR is not exclusively controlled by the SRH mechanism but is rather a combination of SRH and TAT. At 7 V excess bias, both devices experience a decrease in activation energies due to the intensified electric field. It is worth noting that the sidelobes, at 7 V excess bias,

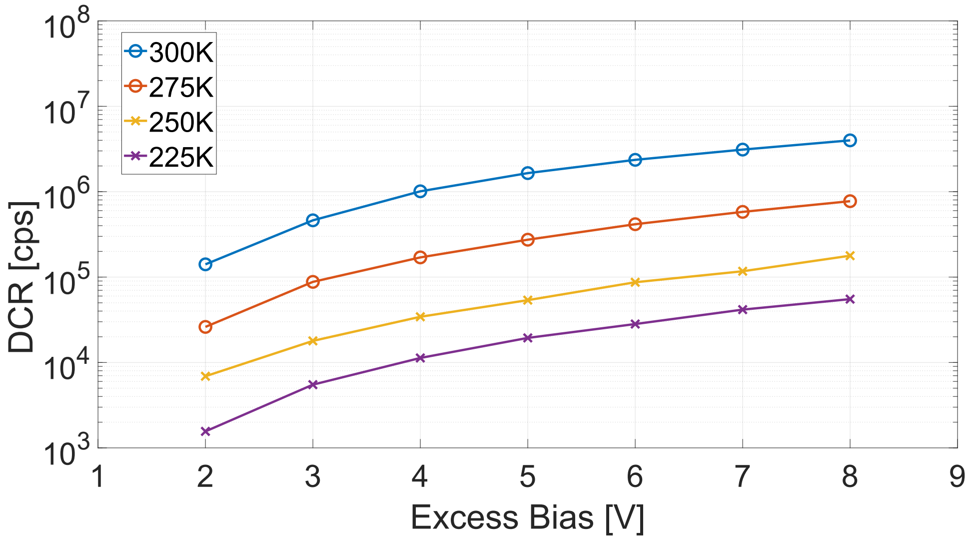


Figure 3.6: DCR measurement of the device without GR from 225K-300K and 2 V-8 V.

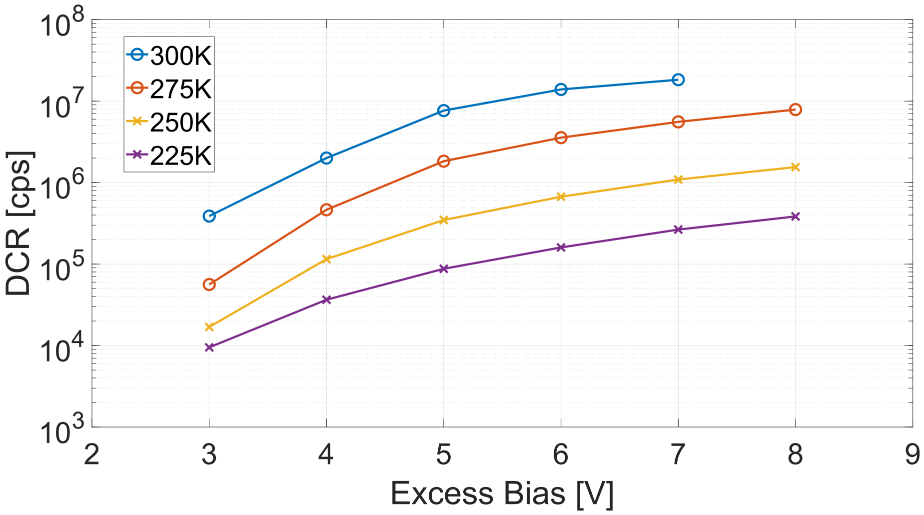


Figure 3.7: DCR measurement of the device with two floating GRs from 225K-300K and 3 V-8 V. Due to count rate saturation at 8 V excess bias at 300K, the result is not plotted at this particular point.

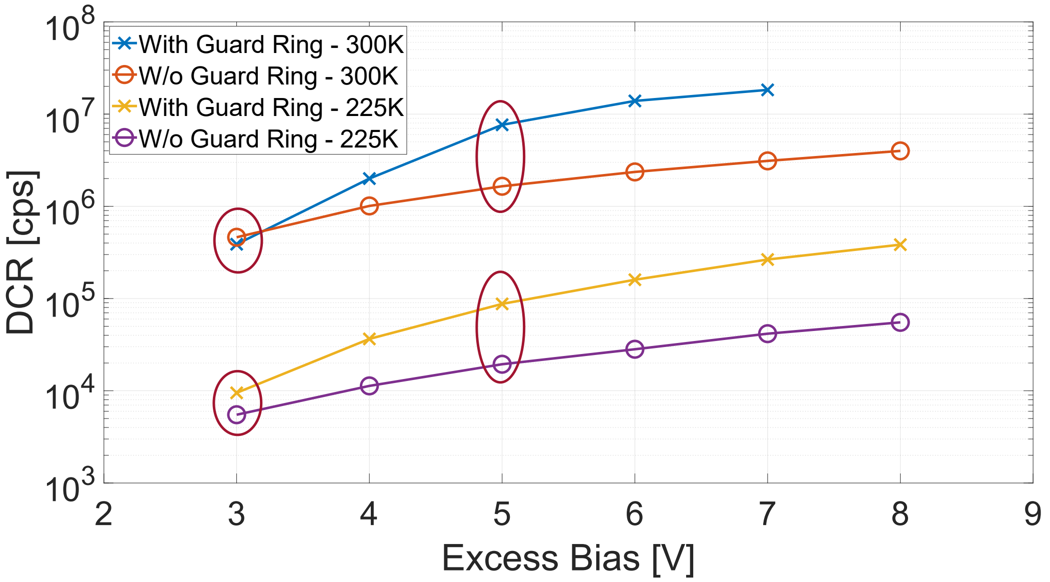


Figure 3.8: DCR results of both devices at 225K and 300K. Values at 3 V and 5 V excess bias are highlighted.

show a response level comparable to the center of the device without a GR. The DCR generated by these sidelobes negatively impacts the device's performance. Hence, there is potential for enhancing the DCR of the device without GR by improving the SAG process to eliminate sidelobes.

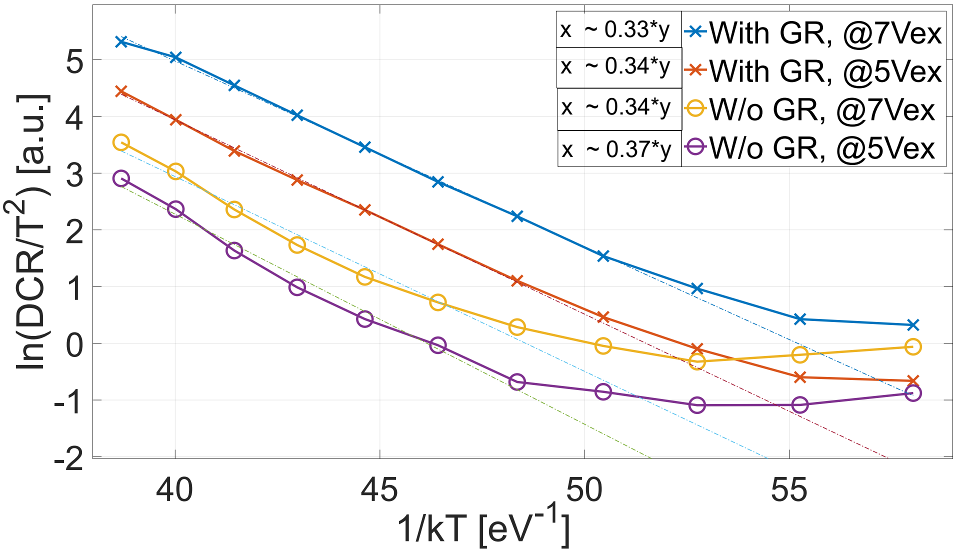


Figure 3.9: Activation energy curves of both devices at 5 V and 7 V excess biases.

3.5.2 Afterpulsing

Eliminating the contribution of afterpulsing on primary DCR and PDP is crucial for accurately characterizing the device's performance. Afterpulsing has the potential to lead to an overestimation of DCR and PDP by causing elevated count rates. To ensure the complete elimination of afterpulsing, it is essential to employ a sufficiently long gate-off time that allows for the release of all trapped carriers. Figure 3.10 displays the relationship between DCR and gate-off time at temperatures ranging from 225K to 300K. Across all temperatures, there is a noticeable rise in DCR for gate-off times shorter than 20 µs. This increase is more pronounced at lower temperatures due to longer trap lifetimes. The specific point at which DCR starts to increase significantly depends on the number of trapped carriers, which is directly proportional to the charge flowing during an avalanche. Additionally, the quantity of charge is influenced by the duration of the pulses. A shorter gate-on time would lead to a shift in the curve toward a shorter gate-off duration. However, at longer gate-off times, the DCR remains constant, indicating that afterpulsing does not impact the count rate. Therefore, a gate-off time of 250 µs at 50 ns is adopted for PDP measurement, similar to the approach used for primary DCR measurements. This ensures that afterpulsing does not influence the obtained results.

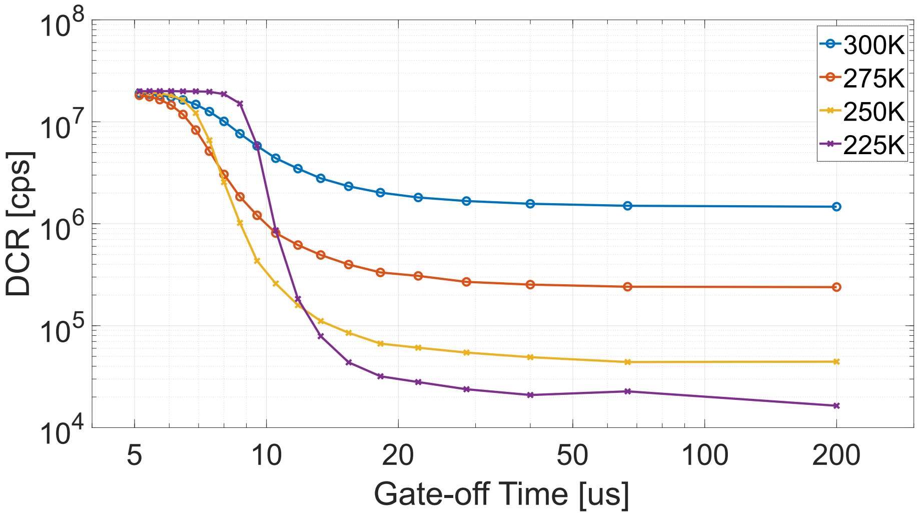


Figure 3.10: DCR *vs*. Gate-off time for the device without GR at 50 ns gate-on time and at 5 V excess bias.

3.6 Photon Detection Probability

PDP measurements were conducted using the setup depicted in Figure 3.11. A laser pulse with a duration of approximately 40 picoseconds at a wavelength of 1550 nanometers was employed. Synchronization between the laser pulse and the electrical gate was achieved via the waveform generator, ensuring that photons reached the SPAD while its voltage was above the breakdown. The incident light power was divided using a beam splitter, with the high output end connected to a power meter (Thorlabs PM20A) to determine the number of photons reaching the device. The power at the low output end was further attenuated and then focused onto the SPAD using a confocal microscope.

Using a focused beam in the PDP measurement provides a significant advantage as it helps to avoid uncertainties associated with the active area. When the entire chip area is illuminated for PDP measurement, inaccuracies in estimating the active device area can impact the results [117]. Additionally, charge persistence could introduce further uncertainties in the determination of the active area, especially at lower temperatures [116, 132, 133].

PDP calculation is performed by

$$PDP = \frac{1}{\mu} \cdot ln\left[\frac{1 - C_d/f}{1 - C_t/f}\right], \tag{3.2}$$

where C_d stands for the number of dark counts per second (cps), C_t for the total number of counts per second under illumination, f is the gate frequency, and μ is the average number of photons per optical pulse [131]. The measured PDP reached 32% at an excess bias of 5 V and increased to 43% at 7 V (Figure 3.12). These measured PDP values exhibited no significant variation between temperatures of 225K and 300K, which is consistent with findings in the existing literature [131]. The increase in PDP with rising excess bias can be attributed to the higher electric field in the avalanche region and the subsequent increase in the probability of breakdown. However, as the impact ionization coefficients saturate at higher excess biases, the PDP reaches a saturation point [134].

3.7 Timing Jitter

Timing jitter is a crucial parameter that significantly impacts the precision of depth measurements in LiDAR applications. To evaluate timing jitter, an experimental setup similar to the one depicted in Figure 3.11 was employed, with some minor adjustments. A reference signal was connected to the oscilloscope to mark the arrival of photons, allowing for the creation of a histogram that illustrates the time differences between the avalanche pulse and the reference signal. These measurements were conducted at room temperature, utilizing a 1550 nm wavelength, and were performed under single-photon conditions.

The results of the jitter measurements revealed FWHM values of 154 ps and 116 ps for 5 V and 7 V excess bias, as shown in Figure 3.13. These FWHM values include the effects of the laser pulse width. After deconvolution with a 40 ps optical pulse, the actual timing jitter of the device was determined to be 149 ps and 109 ps. Figure 3.13 also presents the jitter result at 5 V excess bias, along with a Gaussian fit that exhibits a standard deviation of 65 ps. Additionally, the measurements indicated diffusion tails of 110 ps and 92 ps at 5 V and 7 V excess bias, respectively. In comparison to the results reported in [115], the diffusion tail observed in this study was longer.

It's important to note that conducting jitter measurements using a focused beam, rather than wide-area illumination, can lead to lower jitter estimates by eliminating lateral diffusion, as

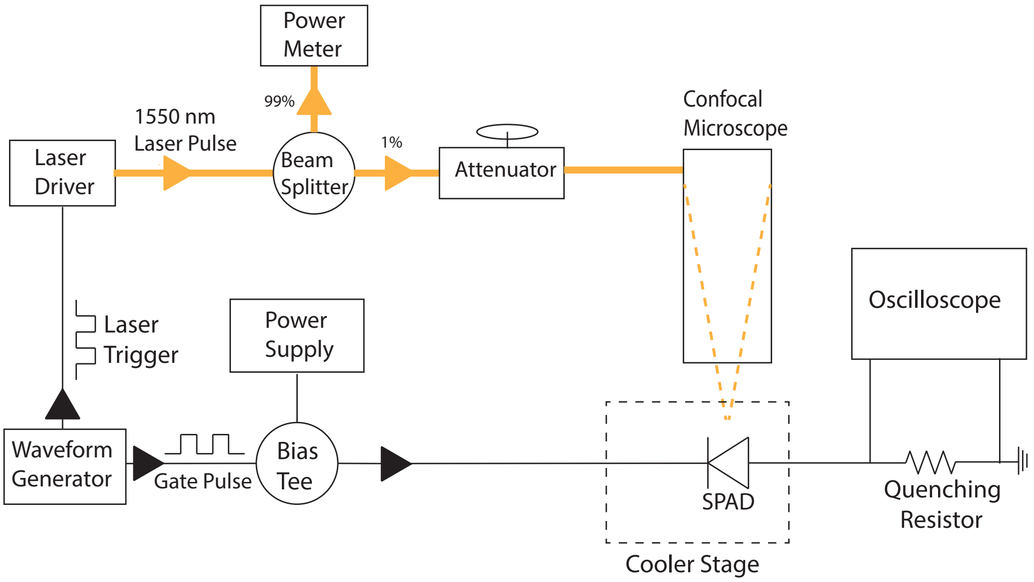


Figure 3.11: Schematic of PDP measurement setup. Black lines show electrical connections while yellow ones correspond to the optical path.

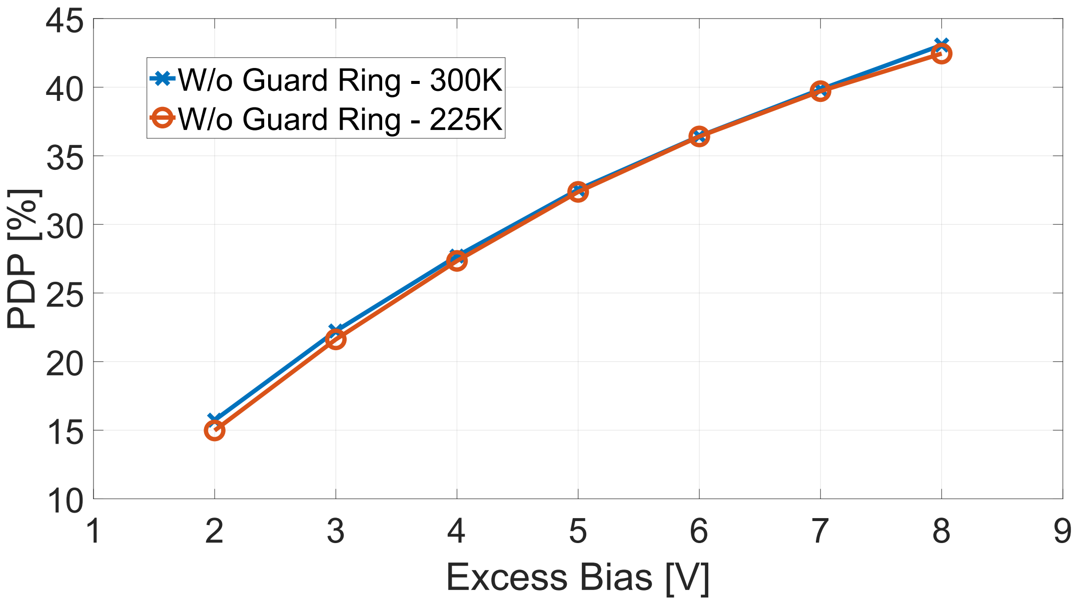


Figure 3.12: PDP measurement results of the front-illuminated device without GR at 225K and 300K (1550 nm).

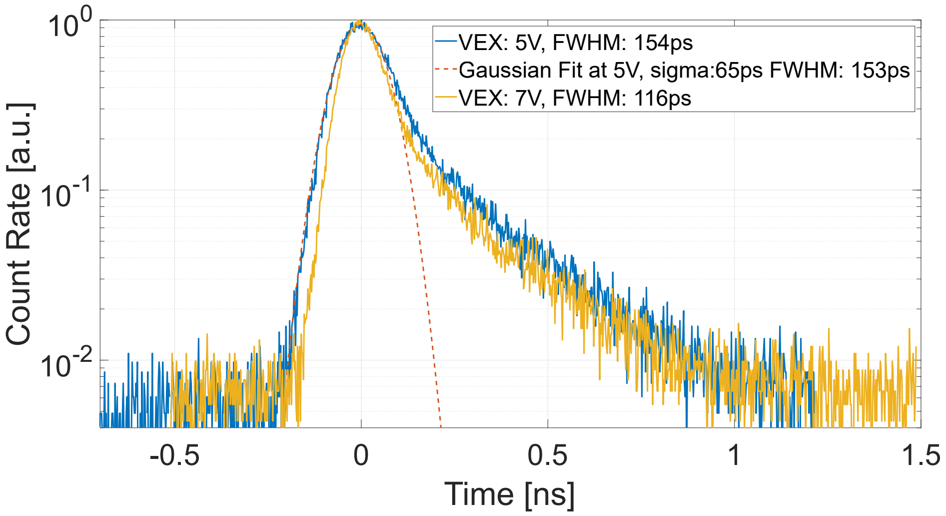


Figure 3.13: Jitter measurement results at 5 V and 7 V excess bias. Measurements were done with a focused 1550 nm picosecond pulsed laser at 300K.

reported in [117]. Furthermore, improving jitter performance can be achieved by employing a high-quality sensing circuit with a low threshold and reduced parasitic capacitance, as highlighted in [64, 135].

3.8 Device Size Shrinking

In the previous section design methods and characterization results of a large active diameter SAG-based InGaAs/InP SPAD have been discussed. This section will focus on reducing the size of the SPADs and its impact on device performance. For fiber-coupled applications such as QKD, reducing device diameter is beneficial since all optical power can be focused on the device while the DCR/noise of the device will be lower. On the other for free space applications and applications that require 2D focal plane arrays such as LiDAR, reducing device diameter serves reducing pixel pitch. This results in a better spatial resolution for the same chip size and improved noise performance.

Understanding how device size shrinking impacts the device performance is critical for achieving optimum pixel pitch and active diameter. Zn diffusion depth and curvature could be affected while device size is reduced. This change could impact breakdown voltage, response uniformity, and noise of the device. Figure 3.14 shows breakdown voltage change with respect to device diameter. 60.3 V 58.9 V, and 57.4 V breakdown voltages are measured for 25 µm, 17 µm, and 6 µm active diameters, respectively. While the opposite trend is reported in the previous work [136], in these devices breakdown voltage slightly decreases for smaller device sizes. Normally, in the same fabrication run, SAG thickness should increase at a smaller mask opening. As a result, Zn diffusion would penetrate less in the InP and the device would have a thicker avalanche region. Finally, this would reflect an increase in the breakdown voltage with reduced active diameter. This inconsistency is attributed to the non-uniform response of the devices from this run. A thinner overall SAG is used for this particular wafer which resulted in enhanced edge response. This is further discussed in the area uniformity part of this section. Figure 3.15 demonstrates the breakdown voltage change with temperature of three devices with 25 µm, 17 µm, and 6 µm Zn diffusion mask opening diameters. For each device, breakdown voltage decreases with temperature due to lower phonon scattering with a slope of approximately 1V/10K.

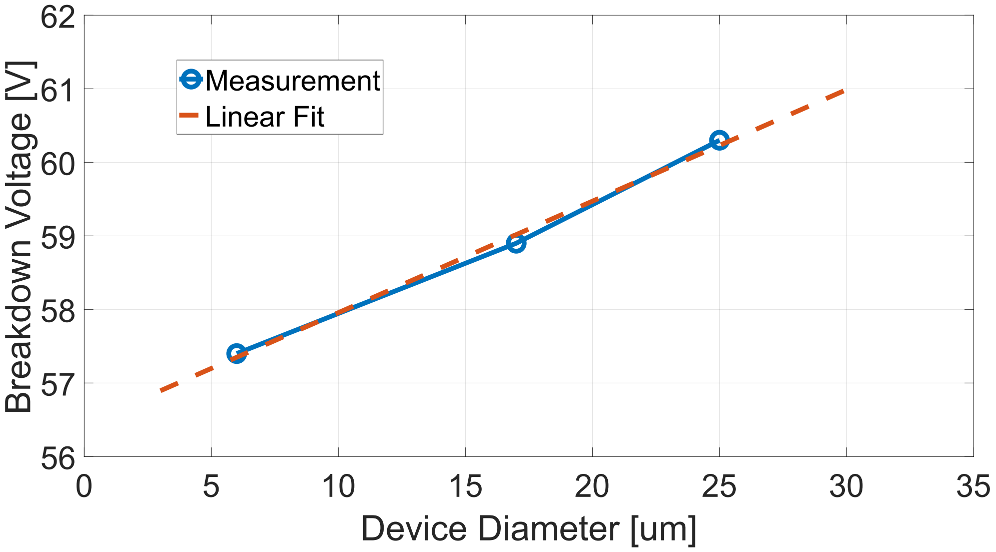


Figure 3.14: Breakdown voltage change with reduced device diameter at 300K

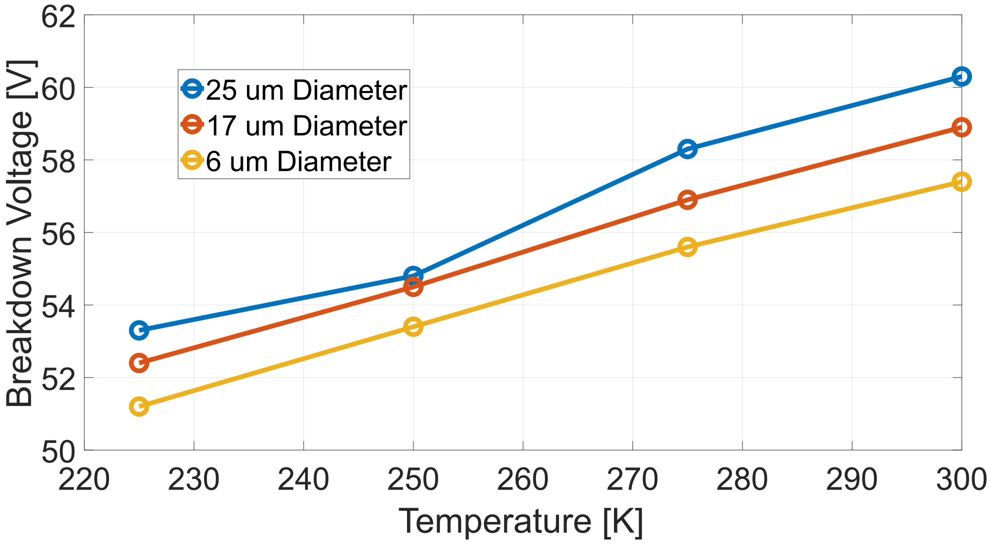
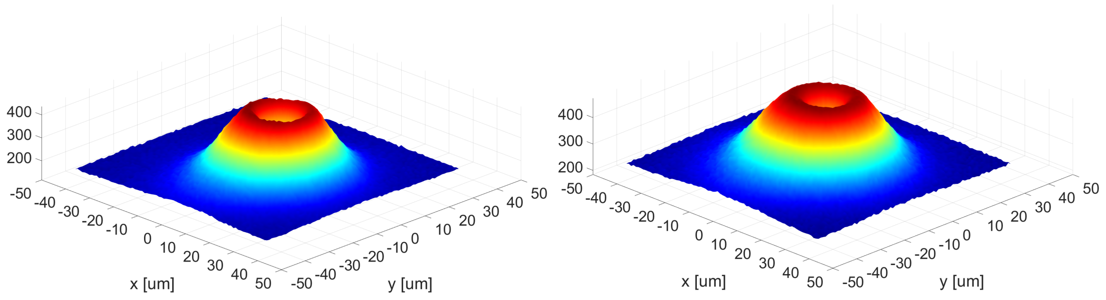


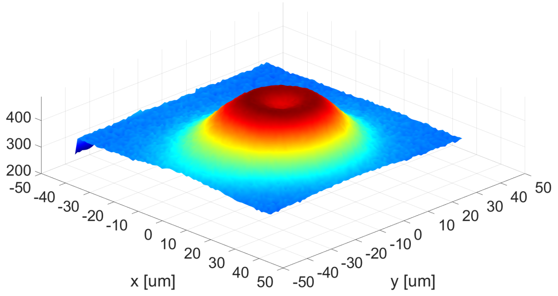
Figure 3.15: Breakdown voltage change with respect temperature for three different device sizes

3.8.1 Active Area Scan

Similar to the previous section the active area of the two devices with active diameters of 17 µm and 6 µm are scanned using the setup shown in Figure 3.11 with a beam size of 1.7 µm. These devices are designed in 2 × 5 array configuration with 17 µm one having a 25 µm pixel pitch and 6 µm one with a 15 µm pitch. Unlike the 70 µm device which had donut-like anode contact metal, these devices are fully covered with circular anode metal on the surface. Therefore, scans are performed in back back-side illumination configuration. A specific printed circuit board (PCB) is fabricated for these samples such that the whole die fits in a hole in the PCB and a smaller area underneath the samples is completely drilled. The fully drilled region allows for light to reach the back side of the device. Then the PCB is placed backward to the setup. This configuration is more realistic since InGaAs/InP SPAD arrays will be used in BSI mode once they are flip-chip bonded to the read-out circuit. Moreover, the shading effect of anode contact metal will not be effective in this configuration.



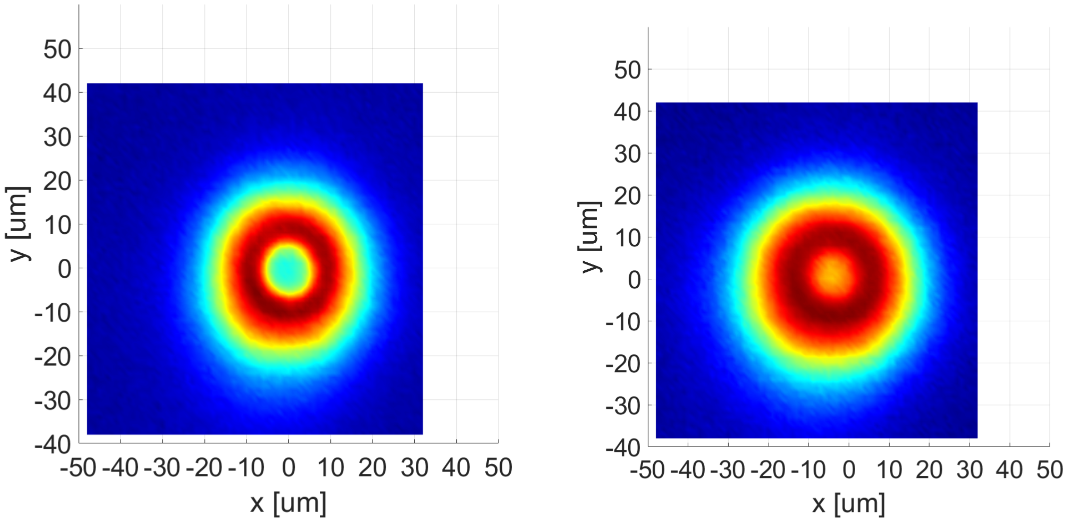
(a) 3 V excess bias (b) 5 V excess bias



(c) 7 V excess bias

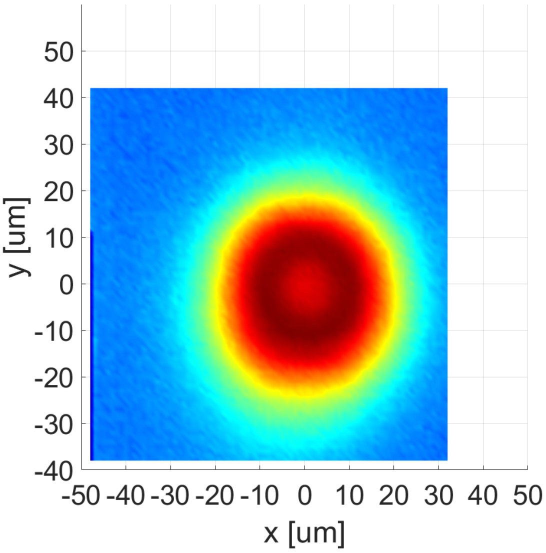
Figure 3.16: Active area scans of the 17 µm device under focused 1550 nm laser light at 300K (Angled View).

Figure 3.16 illustrates the scan results of the 17 µm diameter device at 3, 5, and 7 V excess biases at 300K with an angled view. When the laser is far from the active area count rates correspond to the DCR level of the device. Figure 3.17 demonstrates the same results from top view. The photoresponse of the device is highest at around 8.5 µm distance from the center which corresponds to the mask edge of the Zn diffusion. Photocarriers generated at a distance of 30 µm away from the highest response point could still contribute to the count rate of the device. At 3 V excess bias, the device primarily responds at the periphery while at 7 V excess bias catches up and device response becomes more uniform at 7 V excess bias.



(a) 3 V excess bias

(b) 5 V excess bias

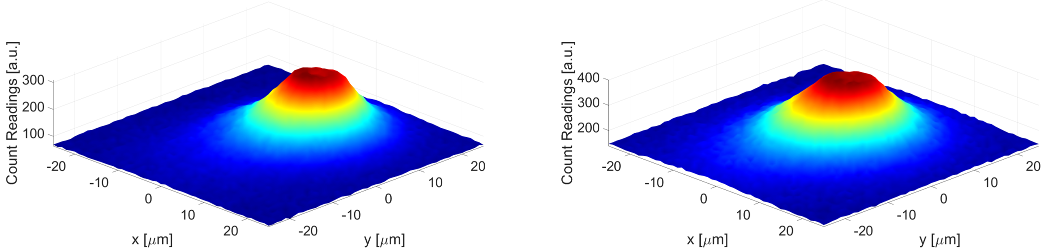


(c) 7 V excess bias

Figure 3.17: Active area scans of the 17 μm active diameter device under focused 1550 nm laser light at 300K (Top View).

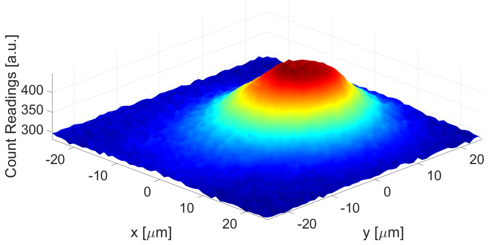
Figure 3.18 and Figure 3.19 demonstrate the scan results from the 6 μm active diameter device. We observe that as the edges of the devices get closer with reduced active diameter, peak responses occurring at the device periphery merge. Consequently, demonstrating a quite uniform response. At 3 V excess bias, there is still a slightly lower response at the device center. Similar to the 17 μm active diameter device, the total response diameter is quite wider compared to the Zn diffusion diameter. At 5V excess bias, the FWHM of device response is ∼

21 µm diameter corresponds to an extension of 7.5 µm radius. Considering these devices have a pixel pitch of 15 µm, the ability to collect photons that are generated 13.5 µm away from the device center could cause significant crosstalk. Nevertheless, forming trenches between the pixels by etching could reduce the crosstalk [137, 138].



(a) 3 V excess bias

(b) 5 V excess bias



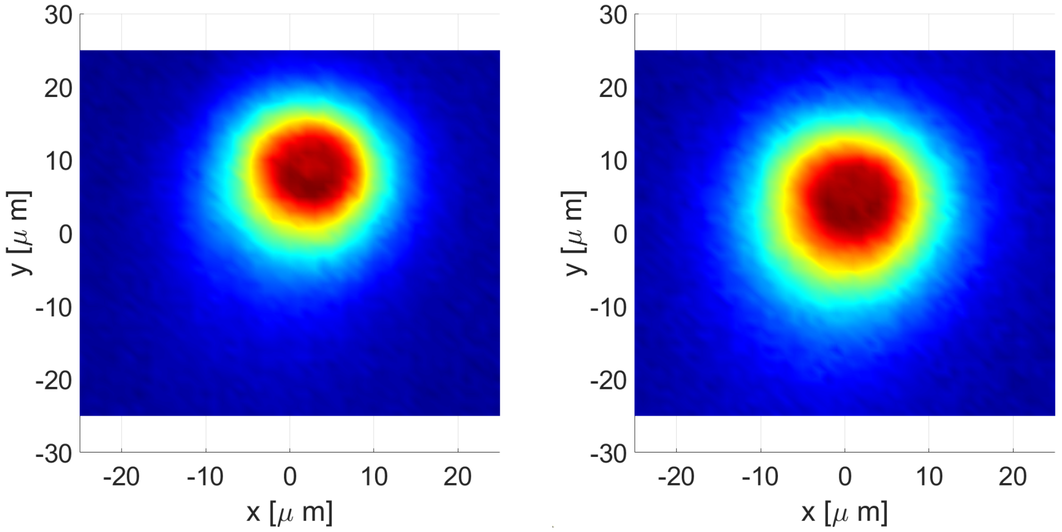
(c) 7 V excess bias

Figure 3.18: Active area scans of the 6 µm device under focused 1550 nm laser light at 300K (Angled View).

3.8.2 DCR

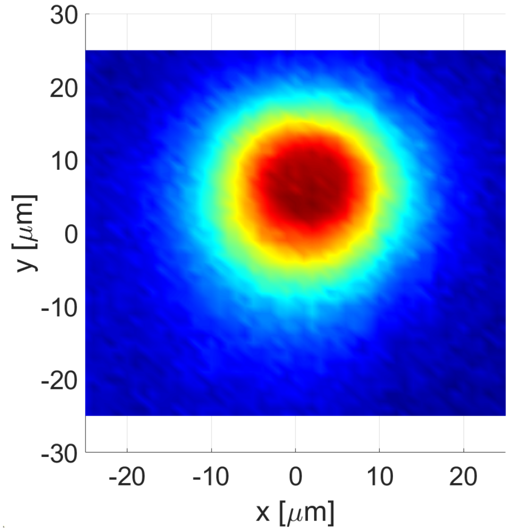
All DCR measurements presented in this section are performed with 200 µs gate-off time and 100 ns gate-on time. Figure 3.20 shows the DCR change with excess bias at 225K, 250K, 275K, and 300K for three devices with changing active diameter. 6 µm active diameter devices exhibit 6 kcps, 29.3 kcps, 180 kcps, and 1.73 Mcps DCR from 225 K to 300 K at 5 V excess bias. Each device shows one order of magnitude DCR decrease for a 25K decrease in temperature down to 250K due to reduced thermal generation. DCR decreases from 250K to 225K is slightly lower since below a certain temperature thermal generation becomes comparable to TAT generation which does not depend on temperature significantly.

Figure 3.21a presents the DCR change with excess bias for 4, 6, 17, and 25 µm active diameter devices together with a 7 µm active diameter device with floating guard ring. The latter is formed in an array format of 2 x 5 with a pixel pitch of 25 µm. Outside the 7 µm active area, a ~ 7 µm thick donut-shaped area is left undoped and all of the rest is Zn doped. This results in an array with all outer regions are doped with Zn similar to a floating GR implementation. DCR increases by a factor of two when the 6 µm active area device is compared to the 7 µm



(a) 3 V excess bias

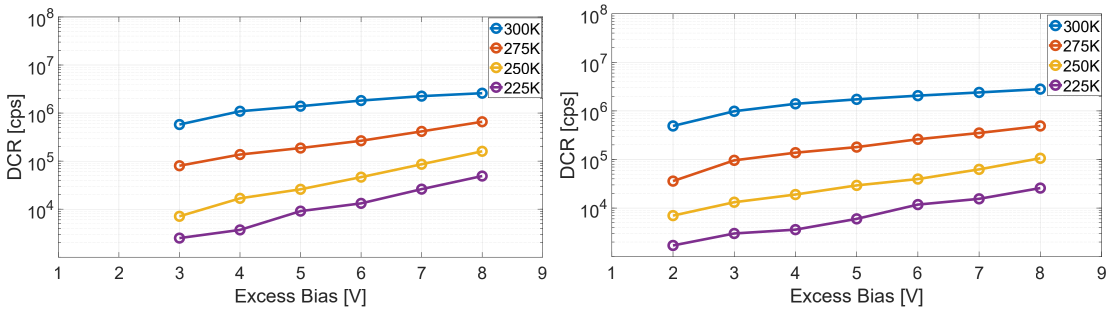
(b) 5 V excess bias



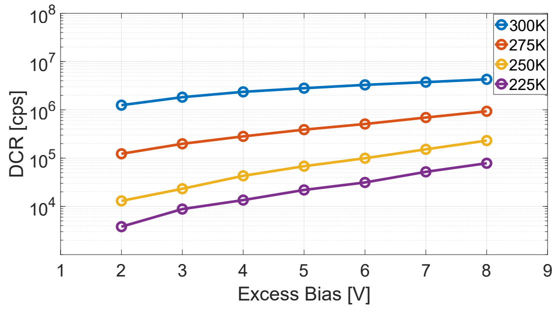
(c) 7 V excess bias

Figure 3.19: Active area scans of the 6 µm active diameter device under focused 1550 nm laser light at 300K (Top View).

active area with floating GR. Similar to the 70 um diameter device, floating GR enhances the electric field around the device periphery causing a higher TAT generation in InP. When the DCRs of the devices without a floating GR are compared, one can notice a decreasing trend in DCR with reduced active diameter. For example, 17, 6, and 4 µm diameter devices show 2.81 Mcps, 1.73 Mcps, and 1.39 Mcps at 5 V excess bias respectively. Ideally, one would expect DCR to decrease proportionally to the active area since the volume of the dark carrier generation rate is expected to change with the active area. However, the 17 µm device has a



(a) 4 µm active diameter (b) 6 µm active diameter



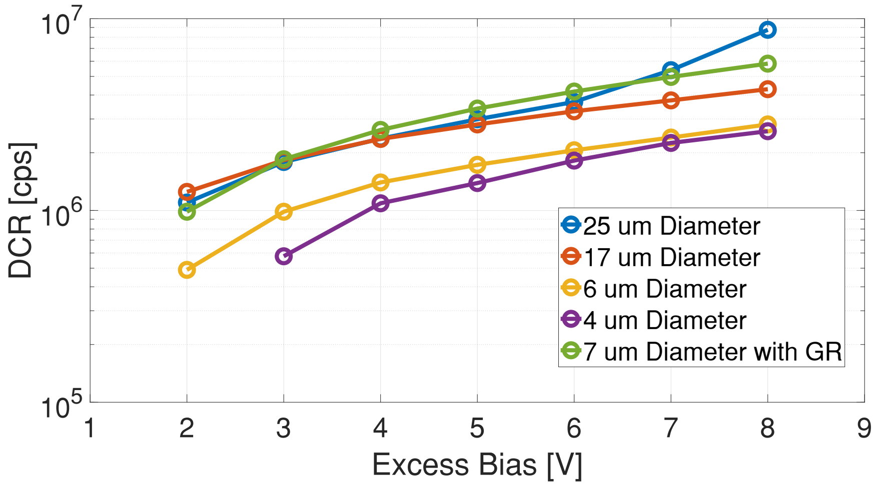
(c) 17 µm active diameter

Figure 3.20: DCR change with excess bias at different temperatures. Measurements are performed with a 200 µs gate-off time and 100 ns gate-on time.

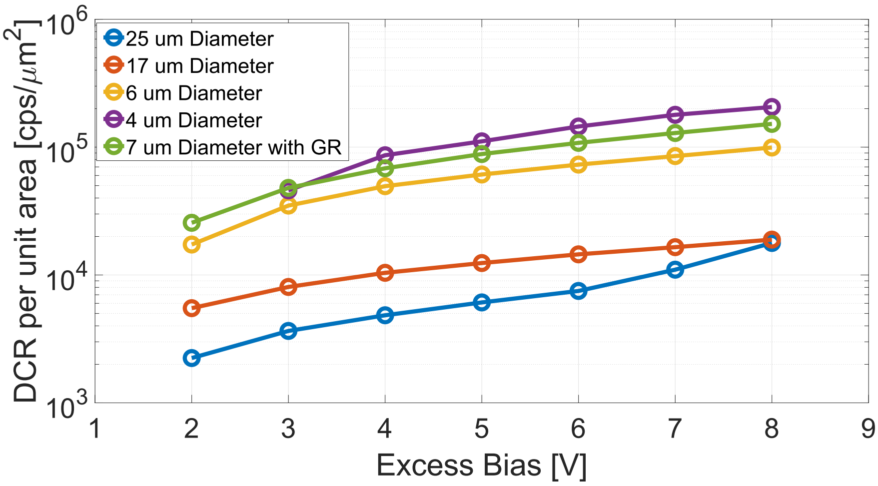
quite non-uniform electric field distribution so it is not fair to assume that it has the same generation rate compared to the 6 µm one. We observed that as the active diameter reaches 6 µm, the device response gets uniform but it is ∼ 15 µm wider than the Zn diffusion diameter. If we assume the same extension applies to the 4 µm diameter device, its response area should have a diameter of 19 µm. This could explain why DCR decrease from 6 µm to 4 µm diameter is not proportional to $(4/6)^2$. Figure 3.21b shows the DCR per unit area change with excess bias. Excluding the device with GR, DCR per unit area decreases with a higher device diameter. Optimizing the Zn-diffusion profile through adjustments in the SAG thickness could allow for a change in the electric field within the device such that it becomes similar to the one of the 70 µm diameter device. As a result, the device response would be uniform. Consequently, achieving a sharper decrease with smaller device sizes.

3.9 Discussion

This chapter presented a new InGaAs/InP SPAD design utilizing a SAG layer for diffusion. This method achieves a uniform electric field in the avalanche region with just a single diffusion step. The devices showcased here are notable for their high efficiency in the SWIR range. They also demonstrate a uniform pixel photoresponse without significant edge field enhancement. The absence of a GR in this structure points to future high-density arrays with a higher pixel fill factor. The reproducibility and scalability to large arrays make these SAG-based SPADs promising for LiDAR cameras. Future developments will focus on reducing pixel sizes and optimizing Zn diffusion and SAG thickness for peak performance, aiming to create SPAD arrays with minimal pixel pitch suitable for LiDAR applications. Currently, an array of 96×96 SAG-based InGaAs/InP SPAD arrays is under development. The SPAD array and readout integrated circuit (ROIC) will be integrated by flip-chip bonding and each pixel will connect to the circuit through indium bumps. A performance comparison with the state-of-the-art InGaAs/InP SPADs is presented at the end of Chapter 5.



(a) DCR comparison at 300K.



(b) DCR per unit area comparison at 300K.

Figure 3.21: DCR and DCR per unit area comparisons of devices with various active diameters. Measurements are performed with a 200 µs gate-off time and 100 ns gate-on time.

4 InGaAs/InP SPAD Simulation Environment

Chapter 3 briefly introduced the electric field simulations of InGaAs/InP SPADs. This chapter aims to expand upon and delve deeper into the simulation results and related aspects. Design, fabrication, and characterization of these SPADs involve extended cycles, demanding a reliable simulation environment capable of estimating crucial device performance metrics. Such an environment is essential for minimizing the number of optimization cycles required. Furthermore, it allows for the optimization of the device for various applications, where higher PDP might be needed to compensate for higher noise, or vice versa.

Several groups have developed in-house simulation environments, such as those demonstrated in [101] and [139], enabling one-dimensional simulations for estimating DCR. However, one-dimensional simulations lack information on the edge curvature of dopants, which can significantly impact the simulated results. Moreover, [140] showcased two-dimensional PDP simulations of InGaAs/InP by post-processing commercial TCAD data with custom code, without simulating DCR.

This chapter primarily focuses on simulations of PDP, DCR, breakdown voltage, and dark current for a 70 μm diameter SAG-based InGaAs/InP SPAD across a temperature range of 225K to 300K. All simulations are carried out entirely in TCAD for ease of use and data management. As Sentaurus TCAD offers capabilities for process and circuit simulations, this work paves the way for a streamlined design and simulation process, covering the entire span from Epi-layer fabrication to the electrical output of an integrated complementary metal–oxide–semiconductor (CMOS) circuit. Ultimately, to assess the accuracy of the model parameters and physical models utilized in this work, the simulation results are compared with the measurement presented in Chapter 3. This comparison ensures the reliability and validity of the proposed simulation framework.

The work presented in this chapter is published in [141].

The author contributed to this work by performing initial simulations and guiding the master project of Vladimir Pešić.

4.1 Simulation Workflow

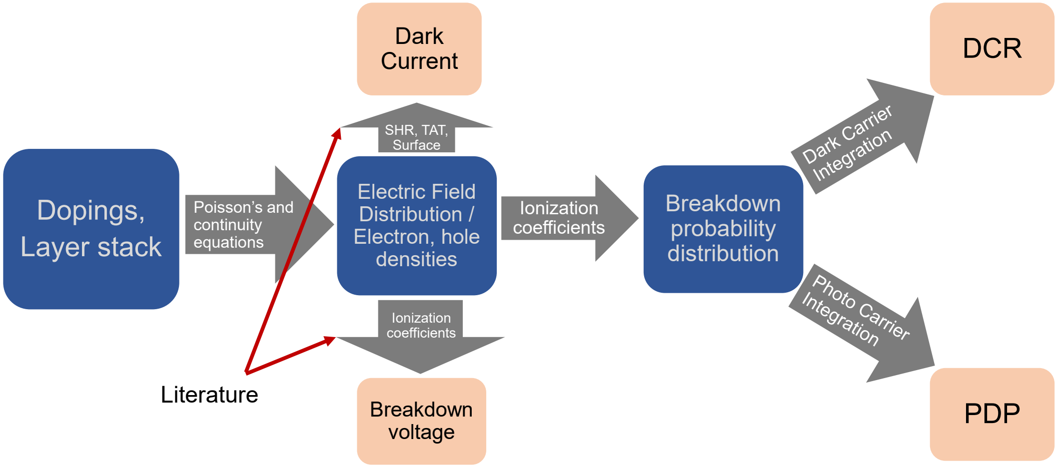


Figure 4.1: Simulation workflow.

The simulation workflow depicted in Figure 4.1 represents a structured process to model and analyze the performance of a SAG-based InGaAs/InP SPAD device using Sentaurus TCAD. The process begins by establishing the physical structure of the device, including doping profiles and layer stack, which form the fundamental basis for the subsequent simulation steps.

Once the structural parameters are defined, the simulation proceeds to solve Poisson's equation along with the drift/diffusion, and continuity equations for electrons and holes as shown in Equations 4.1, 4.2, 4.3, 4.4, and 4.5 [142]. These equations are essential for determining the electric field distribution, the carrier, and current densities within the device.

$$\varepsilon_s \nabla \cdot E = q(p - n + N_D - N_A). \tag{4.1}$$

$$J_n = qn\mu_n E + qD_n \frac{dn}{dx}. \tag{4.2}$$

$$J_p = qp\mu_p E - qD_p \frac{dp}{dx}. \tag{4.3}$$

$$\frac{\partial n}{\partial t} = \frac{1}{q}\nabla \cdot J_n + G - R. \tag{4.4}$$

$$\frac{\partial p}{\partial t} = \frac{1}{q} \nabla \cdot J_p + G - R. \tag{4.5}$$

In these equations, J_n and J_p are the current densities for electrons and holes respectively; n and p are the electron and hole concentrations; μ_n and μ_p are the mobilities for electrons and holes; E is the modulus of the electric field; D_n and D_p are the diffusion coefficients of electrons and holes; G and R stands for generation and recombination rates (impact ionization, SRH, TAT); N_D and N_A are the doping concentrations of donors and acceptors; ε_s is the semiconductor permittivity. By using the finite element method (FEM) the device structure is discretized into a mesh of elements, which transforms the continuous domain of the device into a discrete set. This process results in a system of equations that approximate the original partial differential equations. These equations can then be solved using e.g. iterative methods. In Sentaurus TCAD they are solved with The Bank-Rose method [143]. A doping gradient-dependent meshing is used to define a fine mesh around the junctions and estimate the electric field more accurately. In simulations of heterojunction material systems, the electric field may exhibit discontinuities at the material interfaces due to differences in material properties, such as permittivity and band alignment. To accurately model carrier transport across these heterointerfaces, the thermionic emission model is used [144]. This model considers the thermal emission of carriers over a potential barrier, which is common at a hetero-junction due to band offsets.

In order to simulate dark current and breakdown voltage of the device, material-dependent model parameters are obtained from the literature. They are discussed in detail in the next section. Simulations are run until the break condition of 1 μA is reached. Breakdown voltage is extracted from the corresponding voltage value. Once breakdown voltage is identified, impact ionization generation is deactivated to simulate DCR and PDP at the voltages above breakdown. Carrier densities and electric field are solved as if there is no avalanche multiplication. From the calculated 2D electric profile, values of impact ionization coefficients are found and used to estimate breakdown probability distribution by solving the Oldham model equations along the field lines. The final stage of the simulation involves integrating the breakdown probabilities with dark carrier generation to determine the DCR and with photo-generated carriers to calculate the PDP. The latter requires normalizing the input photons to accurately represent the efficiency of the device in detecting incident light.

4.2 Model Parameters

4.2.1 Ionization Coefficients

Impact ionization coefficients are crucial parameters in the study of semiconductors, particularly in understanding and modeling the avalanche multiplication process. These coefficients represent the rates at which electron-hole pairs are generated due to impact ionization. Also, they define the breakdown voltage of the device. There are various impact ionization models for InP but a quasi-physical model Equation 4.6 from [101] is used since it agrees well with the measurements presented in the paper.

$$\alpha_{h,e}(T,E) = \frac{E}{E_{th-h,e}} \exp\left(- \frac{(E_{th-h,e})^2}{\frac{(E\lambda_{h,e})^2}{3E_{p-h,e}} + E\lambda_{h,e} + kT} \right), \tag{4.6}$$

where E is the field in volts per centimeter, k is Boltzmann's constant in electron volts, T is the temperature in Kelvin, $E_{th-h} = 1.45$ eV, $E_{th-e} = 1.6$ eV, $\lambda_h = 5.38 \times 10^{-7}$ cm × $\tanh\left(\frac{E_{p0-h}}{2kT}\right)$, $\lambda_e = 4.73 \times 10^{-7}$ cm × $\tanh\left(\frac{E_{p0-e}}{2kT}\right)$, $E_{p-h,e} = E_{p0-h,e} \tanh\left(\frac{E_{p0-h,e}}{2kT}\right)$, $E_{p0-h} = 36$ meV, and $E_{p0-e} =$ 46 meV.

However, Equation 4.6 is not available in TCAD. Therefore, the Okuto-Crowell model Equation 4.7 is used. Its coefficients are fitted to Equation 4.6 from 180K to 300K. This is achieved by least square fitting and adjusting the parameters a,b,c,d, γ and δ. A comparison of Equation 4.6 and Equation 4.7 after fitting is demonstrated in Figure 4.2. Parameters c and d of Equation 4.7 control the temperature dependence of the ionization coefficients. Therefore, first other parameters are fitted at room temperature before adjusting c and d.

$$\alpha_{e,h}(F) = \alpha_{e,h}(1 + c_{e,h}(T - T_0))F\gamma_{e,h} \exp\left[- \left(\frac{b_{e,h}(1 + d_{e,h}(T - T_0))}{F} \right)^{\delta_{e,h}} \right], \tag{4.7}$$

where F is the field in volts per centimeter, T is the temperature in Kelvin, $T_0 = 300$K, and a, b, c, d, γ and δ are used as fitting parameters.

4.2.2 Dark carrier generation

Three mechanisms are considered when simulation dark carrier generation; SRH, TAT, and surface traps. SRH is reported to be dominant at room temperature for both InP and InGaAs and TAT in InP is significant at low temperatures or high excess bias [101]. Band-to-band tunneling, Poole Frenkel, or TAT in InGaAs are not used since they become significant at very high electric fields, which is not the case for this device.

SRH is simulated based on the following equation [98]:

$$G_{srh} = \frac{n \cdot p - n_i^2}{\tau_p \cdot \left(n + n_i \cdot e^{\frac{E_{trap}}{kT}}\right) + \tau_n \cdot \left(p + n_i \cdot e^{\frac{-E_{trap}}{kT}}\right)}, \tag{4.8}$$

where n, p are the carrier concentrations, n_i is the intrinsic carrier concentration, τ_n, τ_p are the carrier lifetimes, k is the Boltzmann constant, T is the temperature, E_{trap} is the energy difference between trap and Fermi level, and G_{srh} is the dark carrier generation rate. Carrier

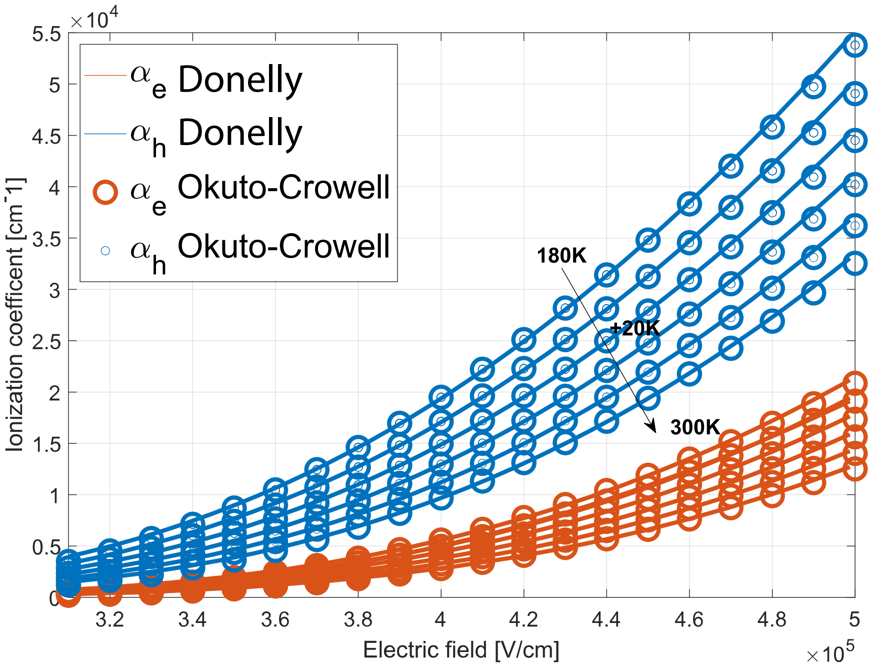


Figure 4.2: Ionization coefficients used in the simulation and proposed in the literature [101].

lifetimes are a function of capture cross-section, thermal velocity, and trap concentration[145]. However, in this work, they are regarded as material quality indicating free fitting parameters since trap concentration could vary from wafer to wafer. TAT generation is a phenomenon where electrons can tunnel through energy levels associated with traps or defects within the semiconductor material. This tunneling happens from the valence band to the conduction band or vice versa. TAT generation becomes greater as the electric field in the device is increased. This is modeled by an electric field-dependent lifetime model [100]. Equation 4.8 is used also to incorporate TAT generation such that τ_n, τ_p are lower at the regions of a higher electric field.

$$\tau_{n,p} = \frac{\tau_{dop}}{1 + g_c(F)}. \tag{4.9}$$

Equation 4.9 shows how this is implemented. $g_c(F)$ is a function of the electric field and defines the relation between the electric field and the lifetime. τ_{dop} stands for doping dependence of the carrier lifetimes.

$$\tau_{dop} = \tau_{min} + \frac{\tau_{max} - \tau_{min}}{1 + \left(\frac{N_i}{N_{ref}}\right)^{\gamma}}, \tag{4.10}$$

where N_{ref} represents a controllable parameter denoting the reference concentration of traps.

N_i signifies the doping concentration, which is individually defined for distinct regions. τ_{max} and τ_{min} represent the maximum and minimum lifetimes specific to the material, and these values can be customized for each material. γ serves as a fitting parameter. In all conducted simulations,τ_{min} was intentionally excluded and set to zero, while $N_{ref} = 10^{16}$ and γ were maintained at their default values of $N_{ref} = 10^{16}$ and $\gamma = 1$.

Finally, traps are introduced at the silicon nitride (SiN) and InP interface to simulate surface leakage. SiN is the hard mask that is used to pattern Zn diffusion. Since the InP crystal will have dangling bonds at SiN interface, trap states will occur. Generation-recombination process due to these traps will not contribute to the DCR since they cannot reach the junction to trigger avalanches. However, they contribute to the dark current generation. Trap positions were presumed to be situated at the midgap position within the semiconductor material. The trap concentration and cross-section were then adjusted as fitting parameters to achieve a close match with the experimental I-V characteristics.

4.2.3 Photo carrier generation

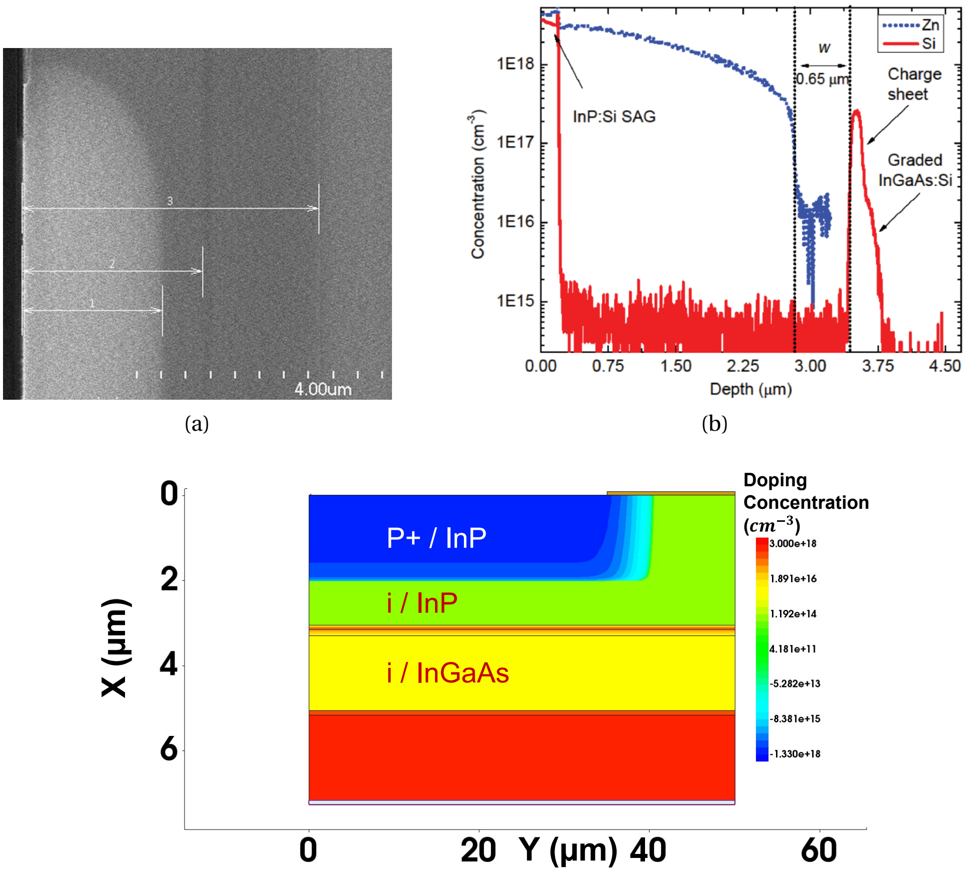
Photocarrier generation is simulated by using a wavelength-dependent absorption coefficient α. Based on the complex refractive index model, α is defined by Equations 4.11 and 4.12.

$$\alpha[\mathrm{cm}^{-1}] = \frac{4\pi k}{\lambda}. \tag{4.11}$$

$$\tilde{n} = n + ik, \tag{4.12}$$

where n is the real and k is the imaginary part of the complex refractive index, λ is the wavelength of light.

In order to simulate the PDP, it is important to accurately compute the absorption across the varying layer stack and the reflections at the interfaces of various materials. This necessitates the calibration of the complex refractive indices for all materials, similar to the methodology described in the reference [140]. The complex refractive index of InP at 300K has been sourced from reference [146], with its associated imaginary component derived from reference [147] at 225 K. For InGaAs, both the real and imaginary parts of its complex refractive index at the stated temperatures are extracted from references [148] and [149], respectively. The refractive index within the grading InGaAsP region is obtained through interpolation from the indices of InGaAs and InP.



(a) (b)

(c)

Figure 4.3: (a) SEM picture of Zn diffusion into InP, acquired from [125]. Markers 1,2 and 3 show the depths of Zn diffusion, InP/InGaAs, and InP/buffer interface respectively. (b) SIMS data acquired from [136] (c) 2D doping profile implemented in TCAD.

4.2.4 Doping Profile

Ideally, Zn diffusion into InP could be simulated with the aid of process simulations in TCAD. However, diffusion parameters of this process are not available in TCAD so the doping profile has to be defined by analytical terms. To this end, the SEM picture is shown in Figure 4.3a is referenced and the same profile is recreated in TCAD. The magnitude of the Zn diffusion is acquired from SIMS measurement (Figure 4.3b). Values of SIMS data are rescaled for different diffusion depths. Charge layer doping is tuned within the accuracy limits of the SIMS tool such that the simulated breakdown voltage matches the measured one. The 2D Doping profile of the simulated device is displayed in Figure 4.3c.

4.2.5 Breakdown Probability

Modeling of the junction in a single dimension enables quick evaluation of a SPAD's detection efficiency [150, 151]. Oldham *et al.* described the voltage-dependent triggering probability with (4.13), (4.14) [112], where P_e and P_h are the probabilities that an electron and hole generated at location x will initiate an avalanche, respectively. Total breakdown probability (P_{total}) is calculated by 4.15 assuming P_e and P_h are independent. The ionization coefficients are denoted by α and β, respectively.

$$\frac{dP_e}{dx} = -(1-P_e)\alpha_e(P_e + P_h - P_e \cdot P_h). \tag{4.13}$$

$$\frac{dP_h}{dx} = (1-P_h)\alpha_h(P_e + P_h - P_e \cdot P_h). \tag{4.14}$$

$$P_{total} = P_h + P_e - P_h \cdot P_e \tag{4.15}$$

Figure 4.4 depicts the breakdown probabilities of electrons and holes together with their joint breakdown probability. Data is obtained from a vertical cutline around the device center at 5 V excess bias. InP and InGaAs absorber and avalanche regions are highlighted. In this configuration where p+ InP is placed between 0 - 2 μm vertical position, the electric field pushes holes from the InGaAs absorber to the avalanche region. This design favors hole initiated process since in InP holes have a higher impact ionization coefficient. This is also evident from Figure 4.4 where holes exhibit higher breakdown probability than electrons under the same electric field. In the avalanche region, holes move towards the left. A hole generated close to the left edge of the avalanche region has a lower probability of triggering a breakdown since it has a shorter path in the high field. Therefore, the breakdown probability of holes increases towards the InP/InGaAs interface. Since the punch-trough voltage is lower than the breakdown voltage, the InGaAs absorber is fully depleted in Geiger mode. The depletion field inside InGaAs pushes holes toward the InP avalanche region. Holes generated inside the absorber have the maximum breakdown probability since they will go through the full path of the high-field avalanche region. Figure 4.5 shows the total breakdown probability changes with excess bias along a cutline placed around the device center. An increase of excess bias results in a higher electric field in the avalanche region, eventually increasing the impact ionization rate. As a result total breakdown probability becomes greater. As it reaches the probability of 1, the increased rate of breakdown probability with the excess bias decreases. This causes the PDP to converge slowly, also called PDP compression.

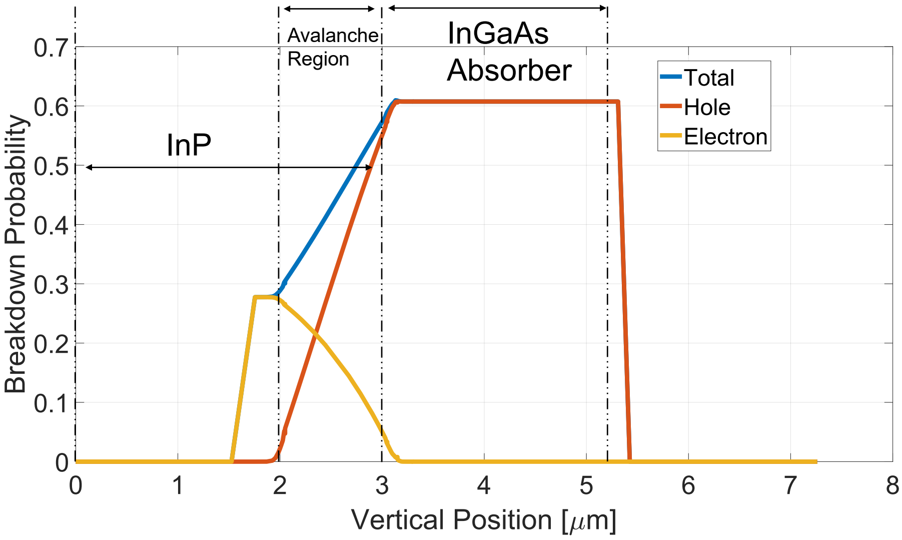


Figure 4.4: Hole, electron breakdown probabilities and their joint probability at 5 V excess bias and at 300K. Data is obtained from a vertical cutline around the device center.

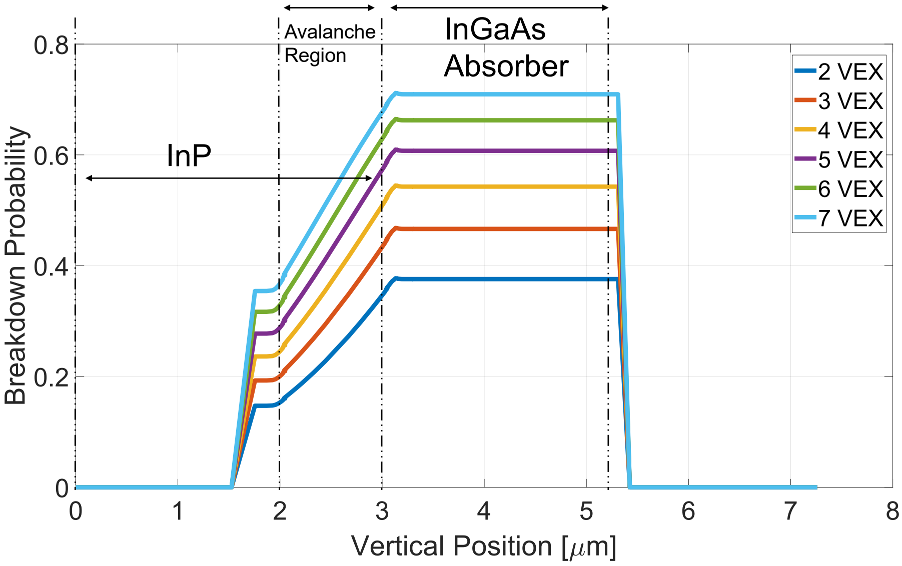


Figure 4.5: Total breakdown probability change with the excess bias at 300K. Data is obtained from a vertical cutline around the device center.

4.3 Simulation Results

4.3.1 Breakdown Voltage

The device described in Chapter 3 is simulated from 225K to 300K with a 25K step. It has a 35 μm radius. The breakdown voltage measurements and simulation results are shown in

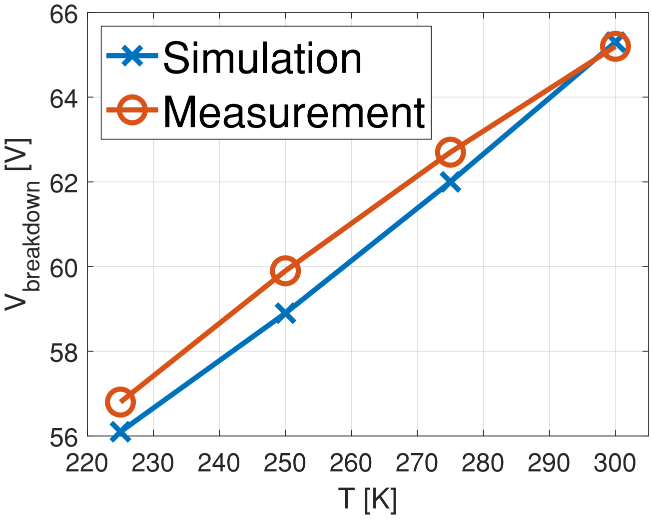


Figure 4.6: Measured vs. simulated breakdown voltages for various temperatures.

Figure 4.6. Breakdown voltage decreases at lower temperatures due to higher impact ionization coefficients. This is illustrated in Figure 4.2. Its slope is 1V / 10K. The breakdown voltage difference between the simulated and measured is smaller than 1 V.

4.3.2 DCR

DCR characteristics of both devices are simulated from 225K to 300K. Dominant noise mechanisms in InGaAs/InP SPADs are SRH in all materials and TAT in InP [101]. Other sources such as band-to-band tunneling or TAT in InGaAs become significant only at a very high electric field, which is not present in any part of the simulated device, so they can be considered negligible. The Scharfetter model is used to simulate SRH generation and TAT is simulated through the electric field-dependent lifetime model (Schenk [100]). DCR is calculated from generation rates using the following equation:

$$DCR = \int_V P_b * G_{srh}\, dV, \tag{4.16}$$

where P_b is breakdown probability, G_{srh} is SRH generation rate and the integral is over volume. Since TAT is modeled by the Schenk model, G_{srh} stands for both SRH and TAT generation.

To ensure that our simulations are reflective of actual device behavior, it is necessary to calibrate the generation rate models accurately. This requires a careful adjustment of the SRH lifetimes for all materials in the model and a precise definition of the TAT parameters specifically for InP. Initially, to establish the TAT parameters, a 1D simulation was conducted on an all-InP device as described in [101]. DCR fitting for this device involved fine-tuning the TAT parameters while employing an SRH lifetime similar to that reported in the referenced paper. The energy level of the trap within the bandgap was set at 1 eV, equivalent to 0.75 times the bandgap of InP, aligning with the understanding that this trap is likely due to a phosphorus vacancy in InP [152]. Once satisfactory DCR fits were achieved for the all-InP device, these TAT parameters were then consistently applied in subsequent simulations.

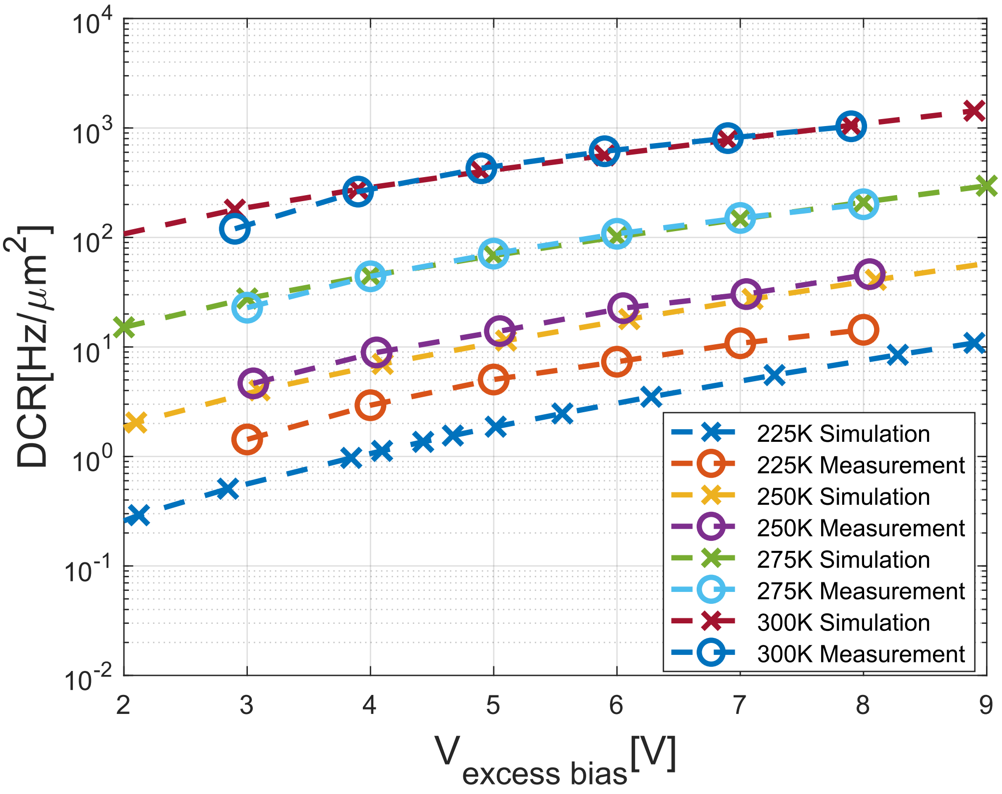


Figure 4.7: Measured vs. simulated DCR of device 1 from 225K to 300K.

For materials like InGaAsP and InGaAs, the traps below the conduction band were positioned at a quarter of their respective bandgaps, similar to the approach used for InP. The next step involved adjusting the SRH lifetimes in InP and InGaAs. These lifetimes were varied to ensure that the simulated DCR at 300K closely matched the experimentally measured DCR. For the InGaAsP and InGaAs materials, defect energy levels were set at one-quarter of their bandgap energies, respectively. The SRH lifetimes for these materials were then adjusted to align the simulated DCR at 300K with the DCR measurements reported in Figure 3.6. These calibrated lifetimes were applied to model the DCR at reduced temperatures, and the results are depicted in Figure 4.7. The simulation matched the measured DCR well at 275K and 250K, but discrepancies are observed at 225K. At this temperature, the observed DCR is marginally higher than our simulation predicts. This discrepancy suggests that the TAT model needs refinement, considering that the TAT is the primary generation mechanism in the InP region at this temperature. Table 4.1 and 4.2 present the values for SRH lifetimes and TAT parameters within our model.

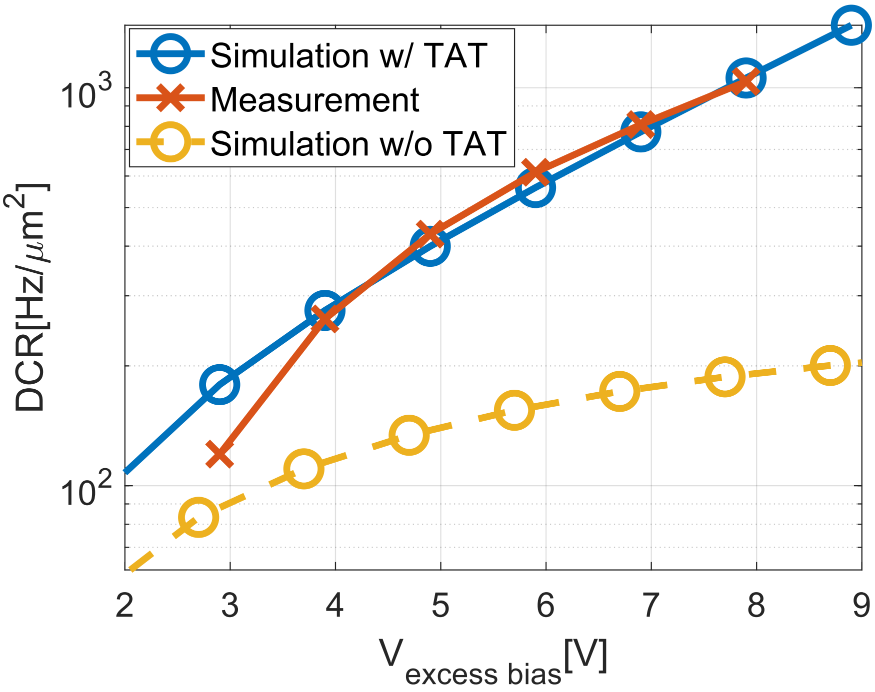


Figure 4.8: DCR simulation at 300K with and without TAT contribution.

Figure 4.8 illustrates the impact of TAT on the DCR at 300K, showing simulations with and without considering TAT effects. The graph underscores TAT's role in accurately simulating DCR, revealing that beyond 2 volts of excess bias, TAT's influence is significant even at 300K. Despite InP having a nominally higher lifetime compared to InGaAs, the intense electric field in the avalanche region increases the TAT generation rate in InP, leading to its enhanced impact on DCR. The lifetime values applied in our simulations for both InP and InGaAs are consistent with published literature references [127][128], indicating that our model is well-aligned with recognized research in this domain.

Table 4.1: Parameters of SRH generation.

Parameter	InP	InGaAs	InGaAsP
Et [eV]	0.34	0.18	0.18
Lifetime [s]	1.6×10^{-6}	9.6×10^{-6}	2.5×10^{-6}

Table 4.2: Parameters of TAT generation.

Parameter	InP
S	0.5
$\hbar\omega$ [eV]	0.1
$m\theta$	0.09

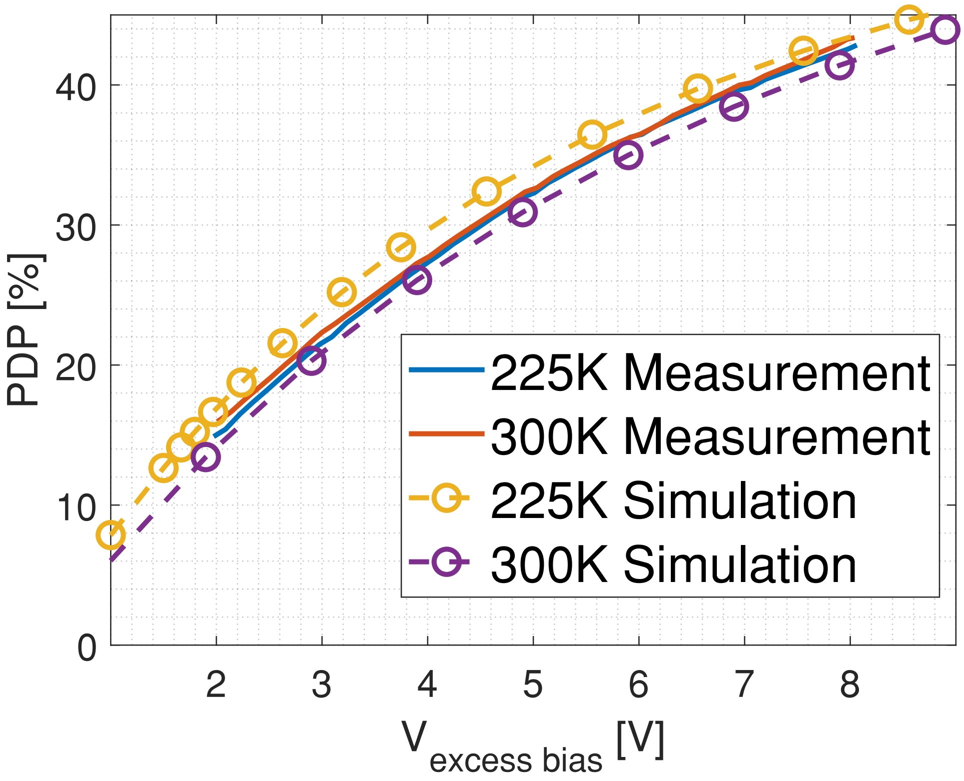


Figure 4.9: Measured vs. simulated PDP under different excess biases and temperatures.

4.3.3 Dark Current

For the I-V curve simulations, we applied the same SRH lifetimes and TAT parameters as those utilized in the DCR modeling. Additionally, it was crucial to factor in the enhanced surface recombination at the interface of InP and SiN passivation.

The surface recombination does not significantly affect the DCR since the carriers generated near the surface have almost zero breakdown probability. However, these carriers are quite influential when it comes to the dark current. For this particular device, the surface generation phenomenon is actually the dominant factor. To simulate the surface recombination, we introduced traps at the interface between the InP and the SiN passivation layers. Trap energy levels are set to the midgap energy level of InP. Both the concentration of these traps and their capture cross-sections were adjusted to ensure that our simulated I-V curve matched the dark current measurements observed at 300K. The results of the comparison between the simulated and measured dark current curves are displayed in Figure 3.3.

4.3.4 Photon Detection Probability

PDP is computed employing equation 4.17. The equation is similar to equation 4.16, but the dark carrier generation rate is changed to the photocarrier generation rate and it is normalized

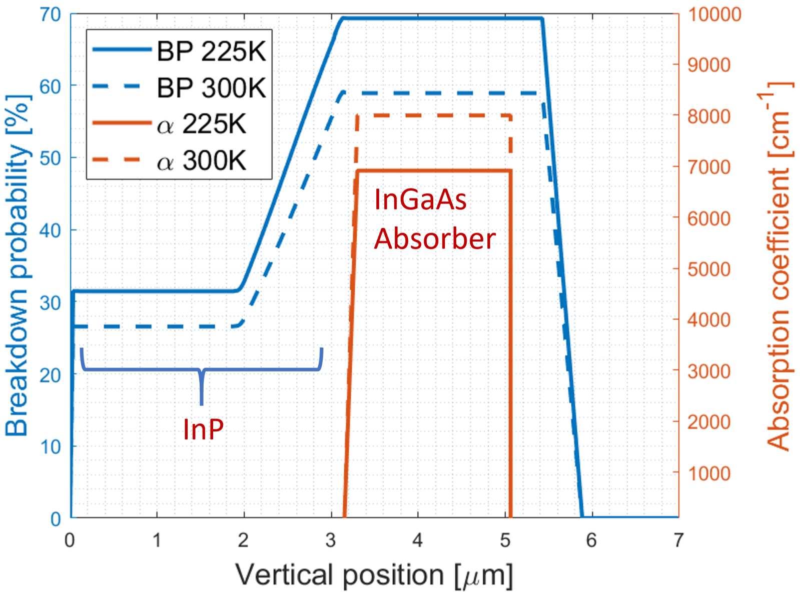


Figure 4.10: Breakdown probability cutline inside the device and absorption coefficient (at 1550 nm) of InGaAs at 225K and 300K at 5V of excess bias.

by the input photon rate.

$$PDP = \frac{\iint_V P_b \cdot G_{\text{photo}}\, dV}{\text{Photon Rate}}. \tag{4.17}$$

PDP measurements were executed using a focused light beam at a wavelength of 1550 nm; thus, the simulations were correspondingly tailored to a 1 µm illumination aperture. A ray tracing optical solver was utilized to assess reflections and propagation of the 1550 nm light within the device. The simulation data demonstrate a consistent match with experimental measurements of PDP at both 300K and 225K, as depicted in Figure 4.9. The PDP does not change significantly from 300K to 225K. Figure 4.10 shows that the breakdown probability experiences a 10% enhancement when the temperature is decreased to 225K under the same excess bias conditions. Nevertheless, the absorption coefficient for the 1550 nm wavelength exhibits a decline at this reduced temperature, resulting in the PDP remaining substantially unchanged between 300K and 225K.

4.4 Discussion

This work provides an optimized simulation environment to model the fundamental figure of merits of an InGaAs/InP SPAD. Breakdown voltage, dark current, DCR, and PDP simulations are performed using the TCAD environment without requiring any other computation outside the tool. To validate the simulations, the results are compared with experimental measurements at temperatures ranging from 225K to 300K. The DCR simulation results are found to be in good agreement with the experimental measurements at temperatures above 225K. As a result, the established simulation environment could be utilized for further optimization of DCR in the future. Furthermore, the experimentally measured PDP at 1550 nm wavelength remains relatively stable from 225K to 300K. This phenomenon is analyzed through simulation and underlying physical reasons are explained. By leveraging the other simulation capabilities of Sentaurus TCAD in the future, complete performance estimations could be done by simulating fabrication processes and circuit couple behavior.

5 Double-diffusion-based InGaAs/InP SPAD

This chapter discusses the design, simulation, fabrication, and characterization of the double-diffusion planar InGaAs/InP SPADs. This method utilizes two distinct diffusion processes. The deep diffusion defines the active area and multiplication thickness. The shallow diffusion acts as a guard ring to reduce the electric field at the device periphery and prevent premature edge breakdown. Electric-field and breakdown probability simulations are performed on Sentaurus TCAD. Details of the simulation environment are discussed in Chapter 4. Growth of epitaxial wafers and double Zn diffusions are conducted by Visual Photonics Epitaxy Co, Ltd. (VPEC). Fabrication of diffusion masks, back-end metallization, and packaging are performed in the Center of MicroNanoTechnology (CMi) of EPFL. Figure 5.1 displays the target cross-section of the device.

The author contributed to this work by performing device design, simulations, backend metal fabrication, and characterization.

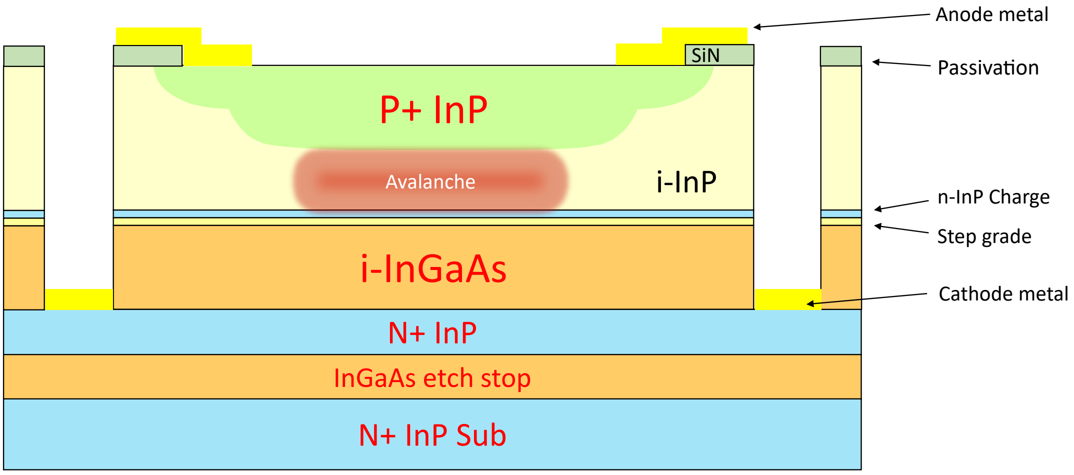


Figure 5.1: Cross-section of the double-diffusion based InGaAs/InP SPAD. Anode metal contact is placed at the top of the device. Cathode contact is reached by etching. The figure is not to scale.

5.1 Device Design & Fabrication

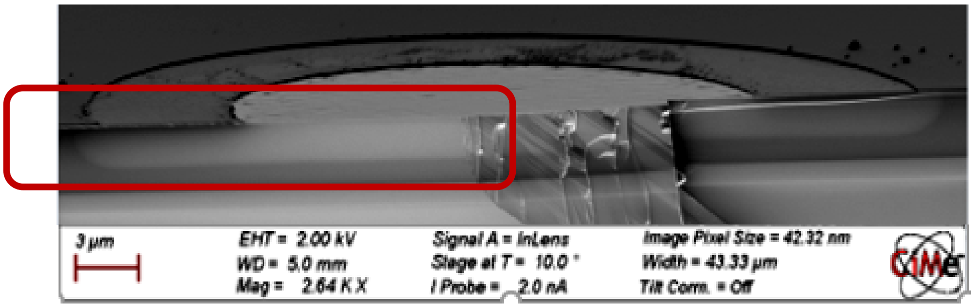
The wafer's epitaxial structure follows the conventional InGaAs/InP SPAD configuration known as SAGCM. With a base of Si-doped InP substrate, approximately 600 μm in thickness, and a doping concentration of roughly ~ 5x10^{18} cm^{-3}, the layers are grown constructed via metal organic chemical vapor deposition (MOCVD). The construction begins with a 100 nm $In_{0.53}Ga_{0.47}As$ layer, serving as an etch stop to precisely manage the wet chemical etching when fabricating back-side illuminated devices. Successive layers include a 100 nm n-InP for cathode contact, a 1 μm $In_{0.53}Ga_{0.47}As$, followed by a 100 nm InGaAsP grading layer, another 100 nm n-InP charge layer, and finally, a 3 μm of undoped InP. The uppermost intrinsic InP layer is deliberately thick to allow for gradual edge curvature during Zn diffusion. The specific doping levels and layer thicknesses are detailed in Table 5.1.

Table 5.1: The layer stack of double-diffusion InGaAs/InP SPAD.

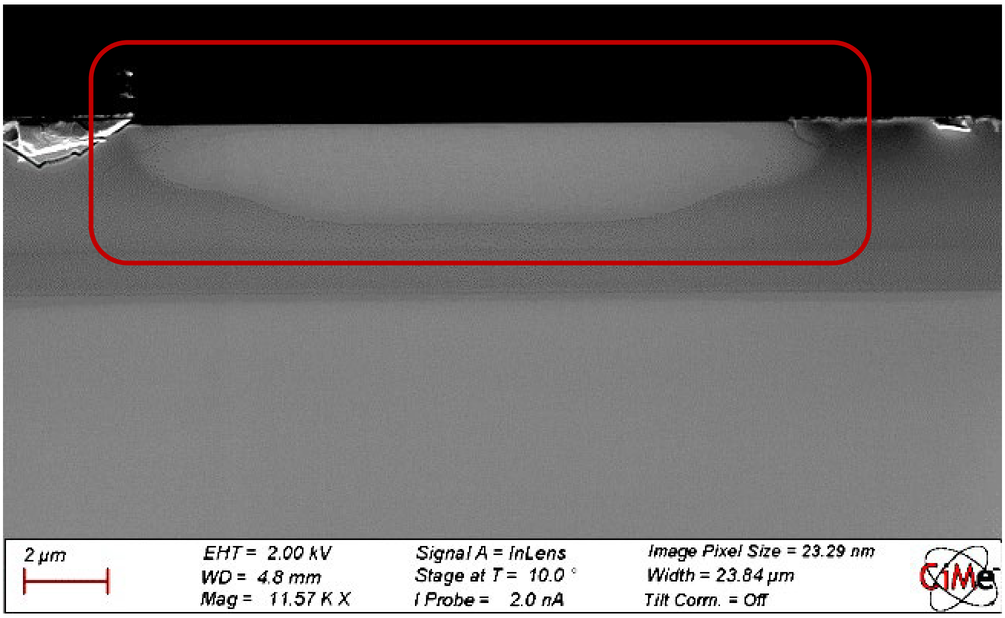
Layer Number	Material	Thickness [μm]	Doping [cm^{-3}]
6	i-InP	3	6.8×10^{14}
5	n-InP	0.1	2.2×10^{17}
4	i-InGaAsP	0.1	-
3	i-InGaAs	1	-
2	n-InP	0.1	8×10^{17}
1	n-InGaAs	0.1	6.3×10^{17}
0	n-InP	600	5×10^{18}

For the fabrication of diffusion masks, CMi employs the VPG200 from Heidelberg Instruments to create a mask on two 5” × 5” chromium-coated glass substrates. These substrates undergo standard photolithography and etching to prepare them for use in patterning the SiN mask on the epitaxial wafers, thereby controlling the Zn-diffused areas. Across two separate fabrication sequences, we alternated the diffusion order; in the first, shallow diffusion preceded deep diffusion, which resulted in the shallow regions extending further during the subsequent deep diffusion, effectively negating the depth variation between layers as depicted in Figure 5.2a. The initial logic was to perform deep diffusion as the last step to achieve better control over its depth. This method has been shown to be successful in achieving guard ring implementation in InGaAs/InP APDs [153]. However, even though the surface of the guard ring area was masked, Zn atoms were able to diffuse deeper during the second diffusion. As a result, devices fabricated from this run exhibited very high DCR because of the strong electric field at the device edge. In the second fabrication run, shallow diffusion is done after the deep diffusion.

The diffusion outcomes within the MOCVD chamber are influenced by the diffusion time, the temperature, and the size of the exposed regions. Higher temperatures facilitate deeper Zn diffusion, particularly through masked sections, as supported by references [154, 155]. Consistent with findings from multiple studies [155–158], the Zn diffusion depth is directly proportional to the square root of the diffusion time. Target diffusion depths are achieved



(a) Cross section of the device fabricated in the first run.



(b) Cross section of the calibration device.

Figure 5.2: SEM pictures of the device cross sections of the first fabricated device and the device from calibration wafer. Zn diffused regions are highlighted.

Table 5.2: The diffusion parameters of InGaAs/InP SPAD.

Wafer Number	Avalanche Region Thickness [μm]	Deep/Shallow Diffusion Depth Difference [μm]
W1	1.5	0.5
W2	1.5	0.25
W3	1.2	0.25

by fabricating a calibration wafer and performing all diffusions at 520 °C via VPEC. Blanket diffusion calibrates the initial wafer's diffusion times, after which we examine the diffusion depths in a 10 μm masked region by fracturing the sample and measuring the depths via SEM. This process reveals a 20% increase in depth for the 10 μm regions relative to the blanket calibrated depths.

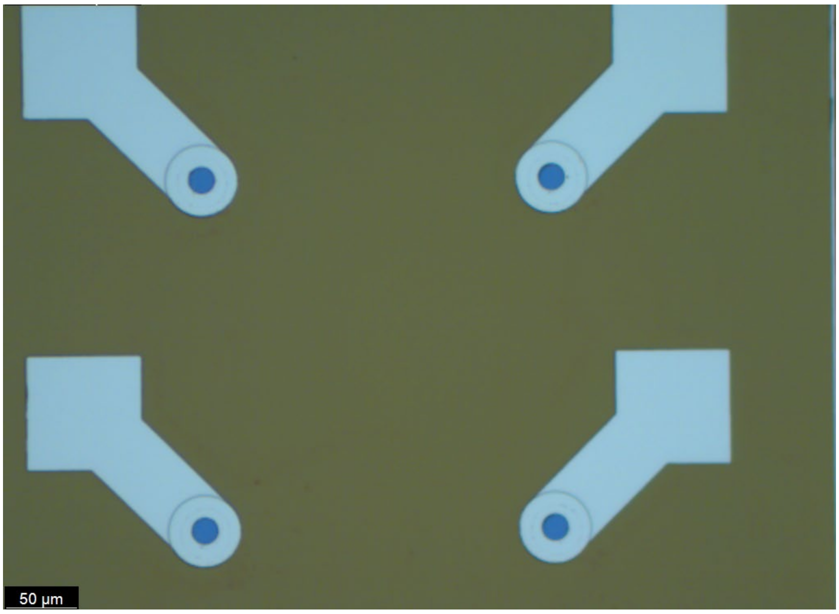
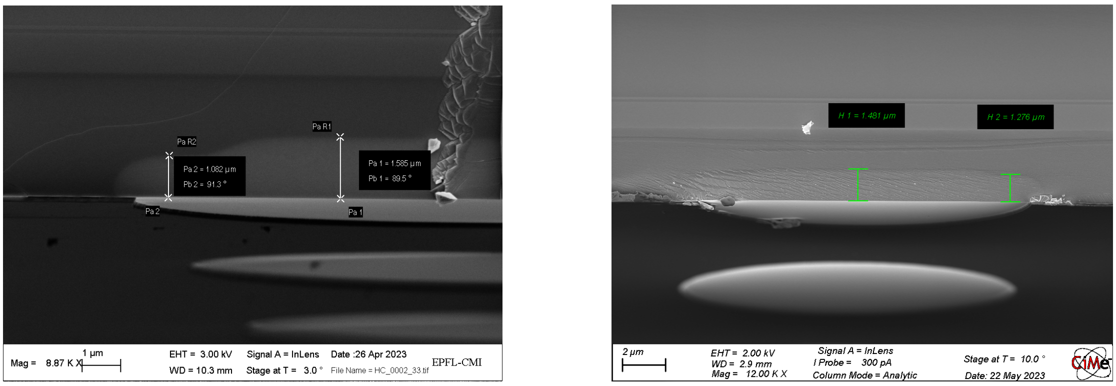


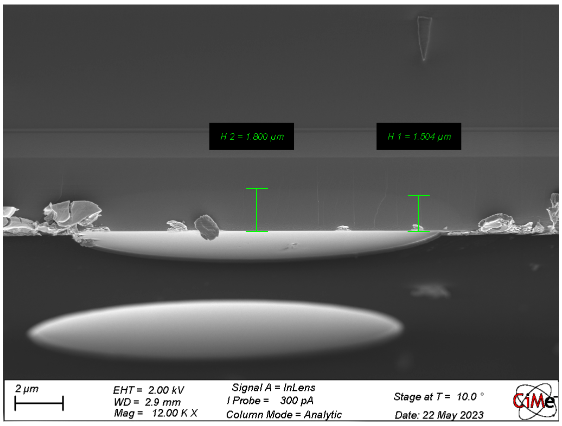
Figure 5.3: Microscope image of the final device.

Informed by these findings, we determined the diffusion times for our target wafers. Figure 5.2b graphically presents the cross-section of 10 μm wide device from the calibration wafer. By examining this device and referring to the principle that diffusion depths scale with the square root of diffusion time, we have precisely calibrated the diffusion times needed for the target wafers. The Zn diffusion depths for the three wafers designed are recorded in Table 5.2. In order to verify the final diffusion depths and profiles of the target wafers, images of the Zn distribution in InP are acquired (Figure 5.4). An array of SPADs with 10 μm deep diffusion diameter and 20 μm shallow diffusion diameter is cleaved through its crystal direction by a diamond cutter. Then, the cross-sections of the broken pieces are observed with an SEM. Subsequent to these steps, the metal contacts are deposited at the CMi. The anode metal is deposited atop the device, adopting a donut-shaped configuration to allow front-side illumination. For both anode and cathode metal contacts, 250 nm / 1000 nm Ti/Au are deposited with electron-beam evaporation. The fabrication of the device's cathode metal contact involves a careful etching process to remove all material layers situated above the N+ InP. This etching process starts with the use of dry etching to remove the SiN passivation layer. Then, a selective wet etching method is employed for the InGaAs, InGaAsP grade, and InP layers. InP layer is etched with a mixture of phosphoric acid (H_3PO_4) and hydrochloric acid (HCl) while for InGaAs a mixture of deionized water, hydrogen peroxide (H_2O_2), and H_3PO_4 is preferred. Etchant of InGaAs does not attack InP and vice-versa so it is relatively straightforward to adjust the etch duration. Figure 5.3 shows the microscope image of the final device. They are arranged with a separation of 250 μm for electrical isolation.



(a) Inverted cross-section of W1

(b) Inverted cross-section of W2

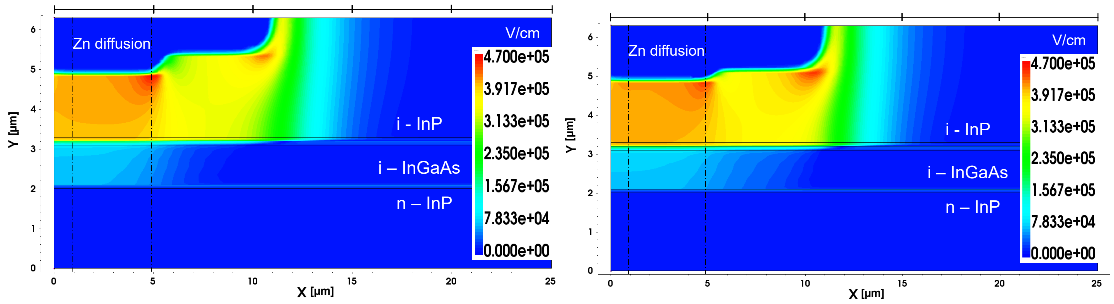


(c) Inverted cross-section of W3

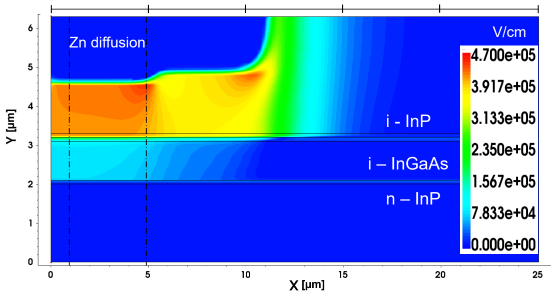
Figure 5.4: Inverted SEM cross-sections of three double-diffusion InGaAs/InP SPADs. Measured diffusion depths of shallow and deep regions are shown.

5.2 Device Simulation

Device simulations are conducted in the simulation environment explained in Chapter 4. 10 μm wide devices from three wafers are simulated for their electric field and dark IV curve. Zn diffusion and charge layer dopings from Figure 4.3b are also used in double diffusion InGaAs/InP SPAD simulations. The diffusion doping peak values are scaled down for shallower depths. In order to minimize the computation time, half devices are simulated. Figure 5.5 depicts 2D simulated electric fields from three wafers. 1D electric field data is acquired at the highlighted cutline positions and displayed in Figure 5.6. The region from 6.3 μm to 3.3 μm is intrinsic InP. Charge and grading layers are placed between 3.3 - 3.1 μm segment, and the InGaAs absorber covers the 3.1 - 2.1 μm. It is clear that the high field causing the avalanche multiplication is confined in InP and the field inside the InGaAs absorber is kept below 1x10^5 V/cm at the breakdown voltage. Since the charge dopings and thicknesses in three wafers are the same, the electric field drop from the avalanche region to the InGaAs absorber is the same as well. Avalanche region becomes narrower when W2 and W3 are compared. As a result, the

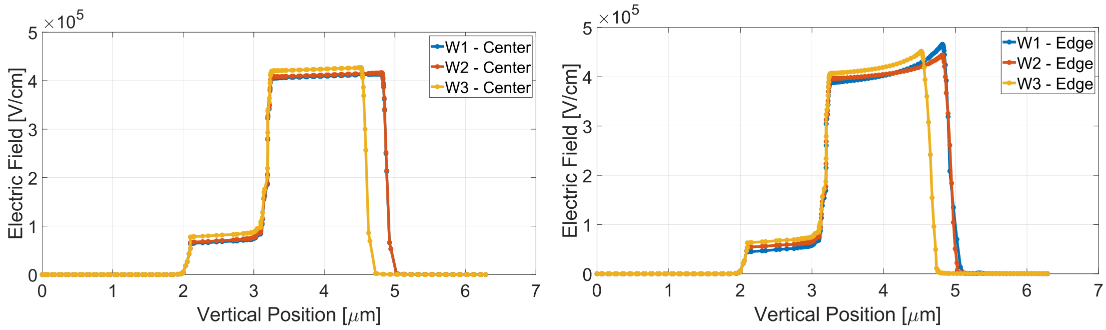


(a) W1 (b) W2



(c) W3

Figure 5.5: 2D electric field distributions of three wafers at the breakdown voltage. Cutlines are placed at the shown locations to investigate the 1D electric field. One cutline is at the center, while the other is at the edge of the deep diffusion.



(a) Cutline at the center. (b) Cutline at the edge.

Figure 5.6: 1D electric field distributions of three wafers at breakdown voltage.

breakdown condition from the ionization integral is reached at a higher electric field in InP and lower reverse bias voltage for W3. Looking at the cutline from the edge, it is evident that some field enhancement occurs due to the curvature of the doping profile. Unlike W2 and W3, W1 has a 500 nm depth difference between shallow and deep diffusion. This results in W1 having a more steep doping transition at the deep diffusion edge. Therefore, W1 demonstrates the highest field enhancement at the cutline from the edge.

Figure 5.7 shows simulated dark current results from three wafers. Surface current is not included in these simulations. The voltage at which the InGaAs absorber region becomes depleted is called punch-through voltage (VP). If the voltage is below VP, carriers generated in the InGaAs region do not drift towards the avalanche region. Therefore, it is designed to be lower than the breakdown voltage (VB), so that in Geiger mode photocarriers generated in the InGaAs could trigger avalanche pulses. On the other end, if the VP is too low, the electric field developed inside the InGaAs can become too large. Eventually causing high noise due to TAT generation in InGaAs. VP values from simulations are 39 V, 39 V, 30 V, and VB values are 76.5 V, 77 V, 68 V respectively for W1 to W3. Compared to other wafers, W1 has a larger discrepancy in VP and VB results when they are compared to the measured ones. Simulation overestimates VP and VB by ~ 6-9 V. This suggests that the electric field at the edge of the device is higher than the simulated result. It is possible that when the depth difference between deep and shallow diffusion is 500 nm, the doping curvature of deep Zn diffusion is not accurately represented in the simulation environment. If in reality curvature is steeper around the edge, the device could exhibit a higher-field and lower VB. VP and VB match quite well for W2 and W3 except with a 3V mismatch on the VP of W2. These wafers are also simulated in Geiger mode from 1 to 5 V excess bias by turning off the impact ionization generation. 1D electric field data from a cutline around the device center is displayed in Figure 5.8. The electric field in the InGaAs absorber region is higher for W3. This is a result of a narrower avalanche region thickness.

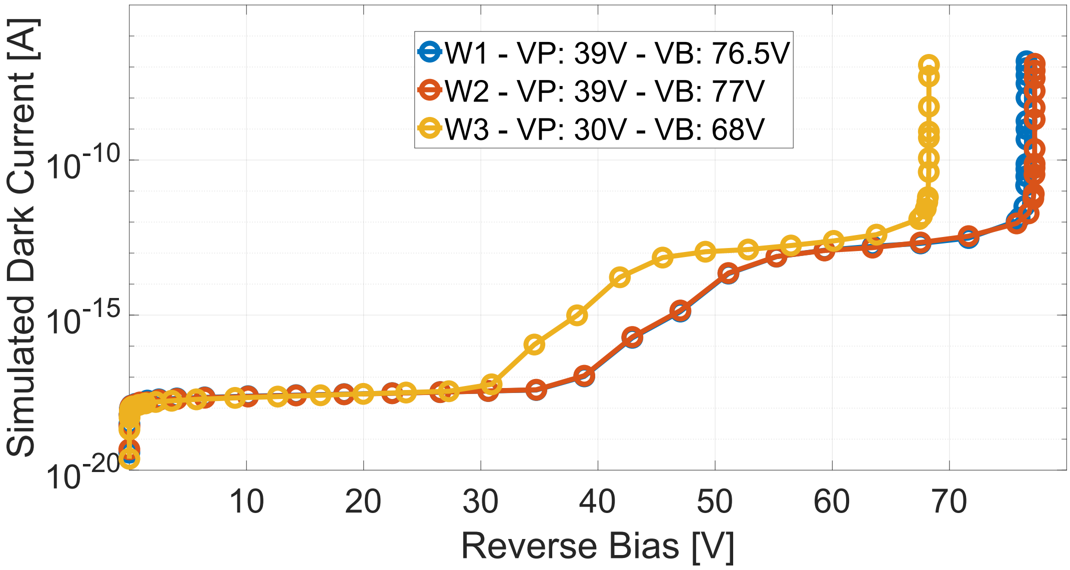
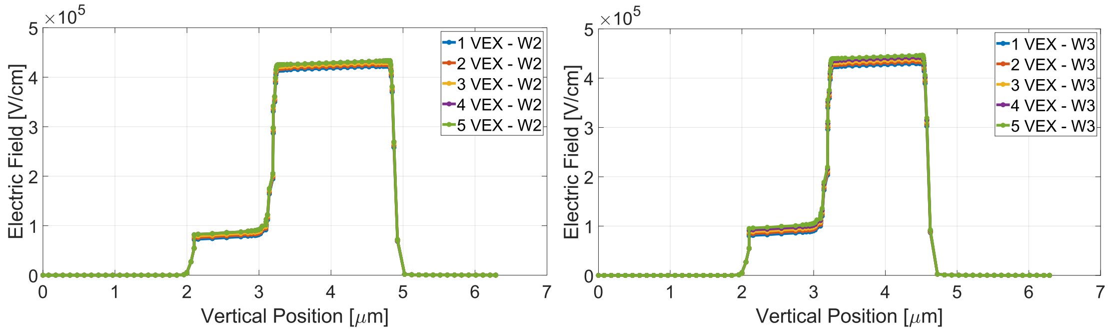


Figure 5.7: Simulated dark current curves of three wafers. The first current increase voltage corresponds to punch-through voltage.



(a) W2 (b) W3

Figure 5.8: 1D electric field distributions of three wafers at the device center in Geiger-mode.

5.3 Measurement Results

The measured IV characteristics are depicted in Figure 5.9. For all three wafers, VP voltage is lower than VB ensuring full depletion of the InGaAs absorber. The measured VB of W1 is lower than the simulated one. This indicates devices from W1 are suffering from premature edge breakdown possibly due to deep diffusion edge curvature.

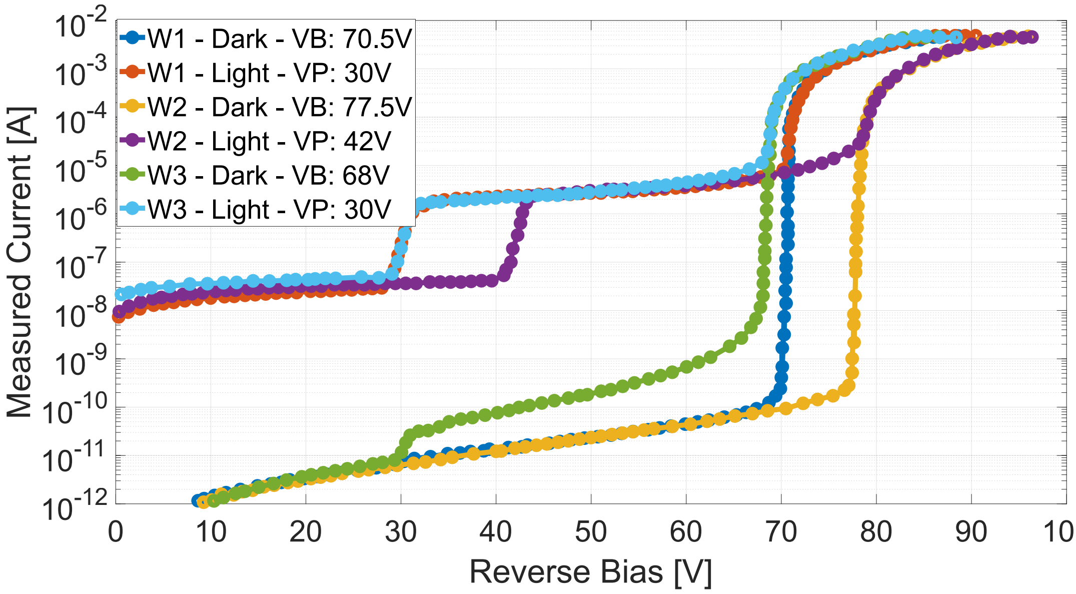


Figure 5.9: Measured SPAD currents in dark and under illumination.

Similarly to the noise measurements of SAG-based InGaAs/InP SPAD, DCR measurements are conducted in gated mode. SPAD on time is defined by 100 ns long voltage pulses with a 10 kHz frequency. DC voltage level is set to one volt below the VB. Figure 5.10 shows the DCR measurements with respect to excess bias from 10 μm diameter devices from three wafers. These measurements are done at 225K since the devices from W1 and W3 are saturated by high DCR at 300K. A DCR of 6.5 Mcps is measured at 5V excess bias from w1, 43 kcps from w2 and 7.2 Mcps from W3 respectively. The lowest DCR is observed from W2 while other wafers exhibited three orders of magnitude higher DCR. Ideally, at 225K DCR is dominated by TAT generation in InP since the SRH generation in InGaAs should be eliminated at lower temperatures. However, W1 and W3 demonstrate very high DCR even at 225K. Supported by the VB mismatch between simulation and measurement, W1 could be suffering from a higher electric field at the InP region around the deep diffusion periphery. This high field could enhance TAT generation in InP eventually increasing the DCR. On the other hand, for W3 the source of high DCR is most likely TAT generation in InGaAs. Looking again at Figure 5.8, the field inside InGaAs becomes very close to 1x10^5 V/cm. This is a result of lower-than-optimum charge layer doping and could be avoided in future fabrication runs. The DCR of the 10 μm diameter device from W2 is also measured at higher temperatures (Figure 5.11). DCR at 5V excess bias increases from 43 kcps at 225K to 625 kcps, 2.8 Mcps, 8.6 Mcps at 250K, 275K, and 300K respectively. This increase is due to higher dark carrier generation in InGaAs (SRH) at higher temperatures.

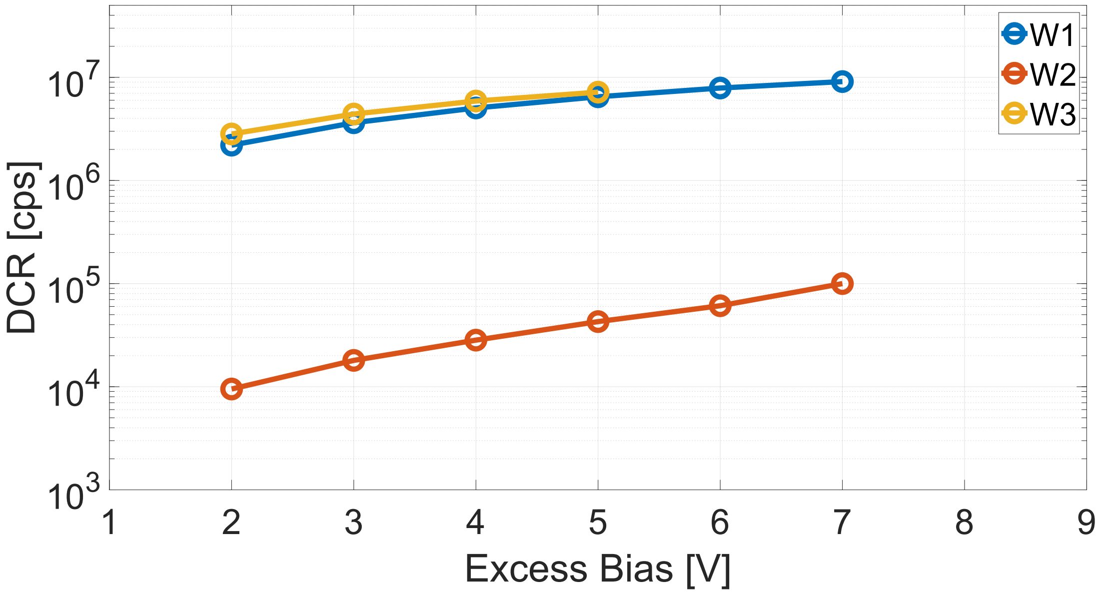


Figure 5.10: DCR measurements with respect to the excess bias from three wafers with 100 ns gate pulses at 225K. The gate frequency is 10 kHz. W1 has a larger depth difference between shallow and deep diffusions compared to W2. W3 has narrower avalanche regions compared to other wafers.

Active area scans, PDP, and jitter of the device from W3 are measured using the setup in Figure 3.11. Active area scans are quite useful to understand the uniformity of the device response. Scan results depicted in Figure 5.12 are acquired by focusing the 1550 nm picosecond pulsed laser beam to a ~ 2.5 μm spot size. During the measurement, the laser beam position is kept constant while the device is moved with a 1 μm step over a 30 μm by 30 μm window. At each position, the count rate of the SPAD is recorded. Measurements are done in gated mode

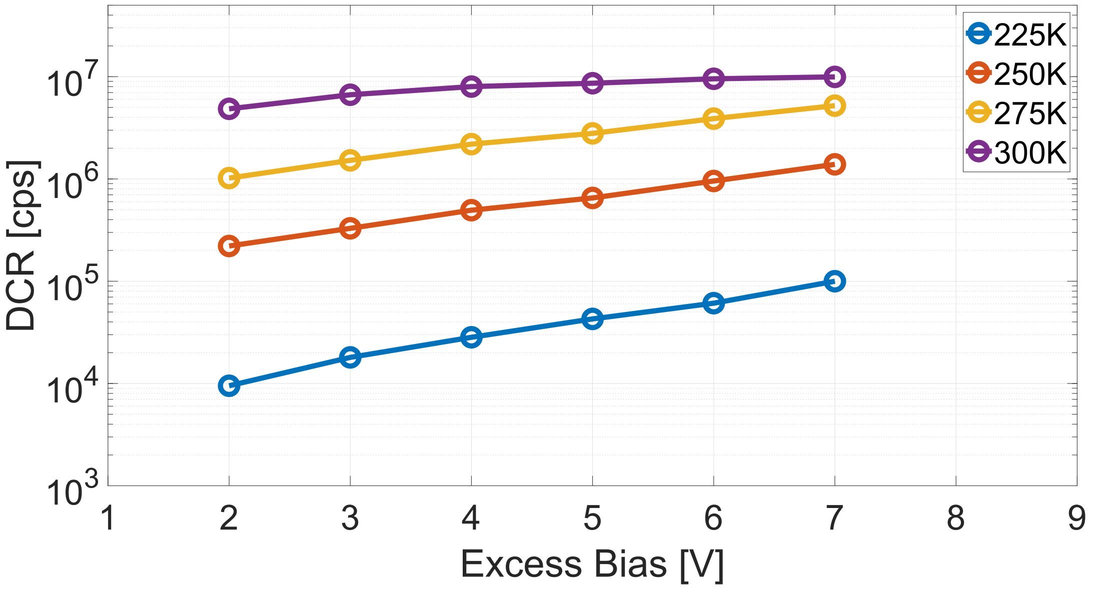
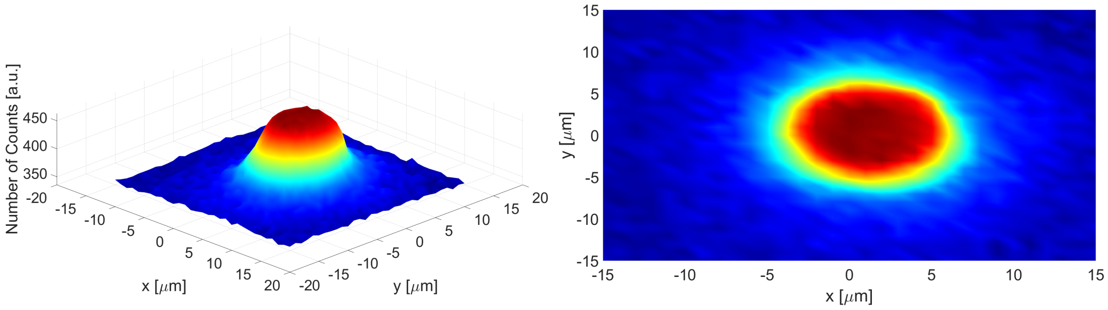


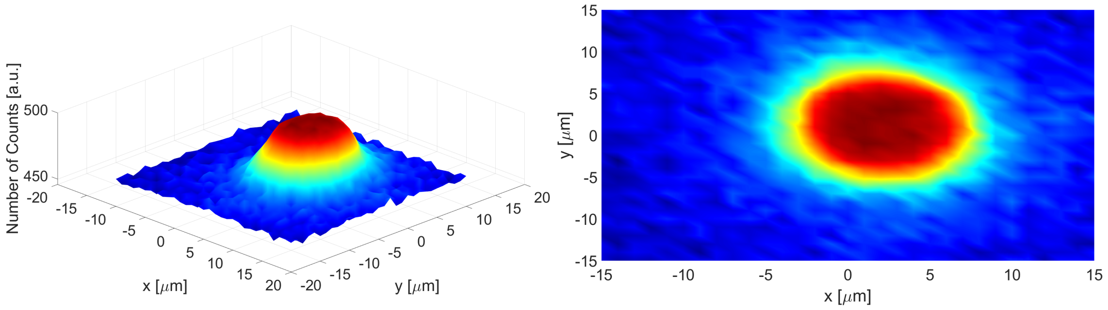
Figure 5.11: DCR measurements with respect to the excess bias of a 10 μm diameter device from W2 at a temperature range of 225K to 300K. The gate frequency is 10 kHz and on time is 100 ns.

to suppress afterpulsing. Excess bias voltage is generated from a waveform generator which is also synchronized to the laser driver. The phase difference between the 100 ns long excess bias pulse and the photon pulse is adjusted, such that photons impinge at the beginning of the gate. The device is illuminated from the backside through the InP substrate since InP is almost transparent to 1550 nm photons. This configuration also helps to eliminate the metal shading on the scan results. Device response is quite uniform and with a diameter of 10 μm as defined by the deep diffusion diameter. When the excess bias is increased from 3 V to 5 V, the response area becomes slightly wider, and the peak response increases. This is expected since the PDP is higher at higher excess biases.

As shown in the Figure 3.11, the beam splitter's 99% output was connected to a calibrated power meter (Thorlabs PM20A) for measuring the input laser power. A calibrated InGaAs photodiode (Thorlabs FGA21) calculated the mean number of photons per optical pulse, accounting for the power reduction at the 1% beam splitter port, the attenuator, and the confocal microscope's objective. The measurements were ultimately carried out with an average of 0.15 photons per optical pulse. The PDP with respect to excess bias is measured when the laser beam is positioned at the center of the device. Its calculation is done by the Equation 3.2. The PDP is ~ 5 % at 3 V excess bias and rises to 15 % at 5 V excess bias (Figure 5.13). Compared to the SAG-based device, the double-diffusion-based InGaAs/InP SPADs exhibit lower PDP due to thinner absorption layer thickness. 1.8 μm thick InGaAs absorber thickness is reduced to 1 μm for these devices to reduce the DCR. PDP is also slightly lower than the device reported in [159] which incorporates a similar layer stack. This might be due to a lack of anti-reflection (AR) coating in our implementation. AR coating could reduce the reflections at the air/InP interface resulting in a higher number of photons reaching the InGaAs absorber region.



(a) W2, 3 V excess bias, angled view (b) W2, 3 V excess bias, top view



(c) W2, 5 V excess bias, angled view (d) W2, 5 V excess bias, top view

Figure 5.12: Active area SCAN results of double-diffusion InGaAs/InP SPAD.

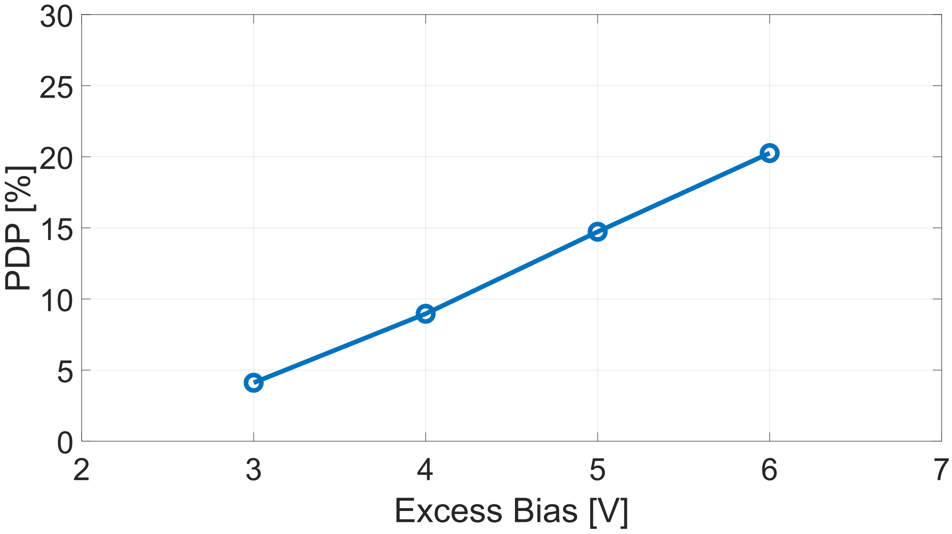


Figure 5.13: PDP measurement results of the 10 μm double-diffusion based InGaAs/InP SPAD from W2 at 1550 nm and 300K.

The timing accuracy of the device is obtained by constructing the histogram of the time difference between the photon emission and its detection by the SPAD in the TCSPC regime. The laser trigger signal is used as a mark of photon emission. The time difference between the avalanche pulse and the laser trigger signal is measured by an oscilloscope (Teledyne LeCroy WaveMaster 813 Zi-B). The sharp increase of the excess bias gate signal results in a parasitic response in the quenching resistor [160]. In order to minimize their contribution to the overall timing jitter, the SPAD is first kept 15 V below the breakdown voltage. Then the waveform of the parasitic response is captured and later subtracted from the avalanche pulse to obtain a clean signal. The measurement is performed while the laser is focused at the center of the device. Flood illumination substantially causes a higher jitter due to the lateral diffusion of carriers. Figure 5.14 shows the measured timing histogram together with a Gaussian fit at the center and an exponential fit at the tail at 5 V excess bias, at 1550 nm wavelength and 300K. A FWHM of 131 ps is measured including the contribution of the 40 ps laser pulse width. After deconvolving the laser width, the timing jitter of the device is 125 ps. Compared to the SAG-based approach, the timing jitter of the double-diffusion version is improved by 25 ps. This improvement is attributed to the choice of a thinner absorption region which reduces the PDP of the device. As a result, a trade-off between maximum attainable PDP and timing jitter exists when deciding the thickness of the InGaAs absorber. Finally, the time constant of the diffusion tail is calculated as 82 ps.

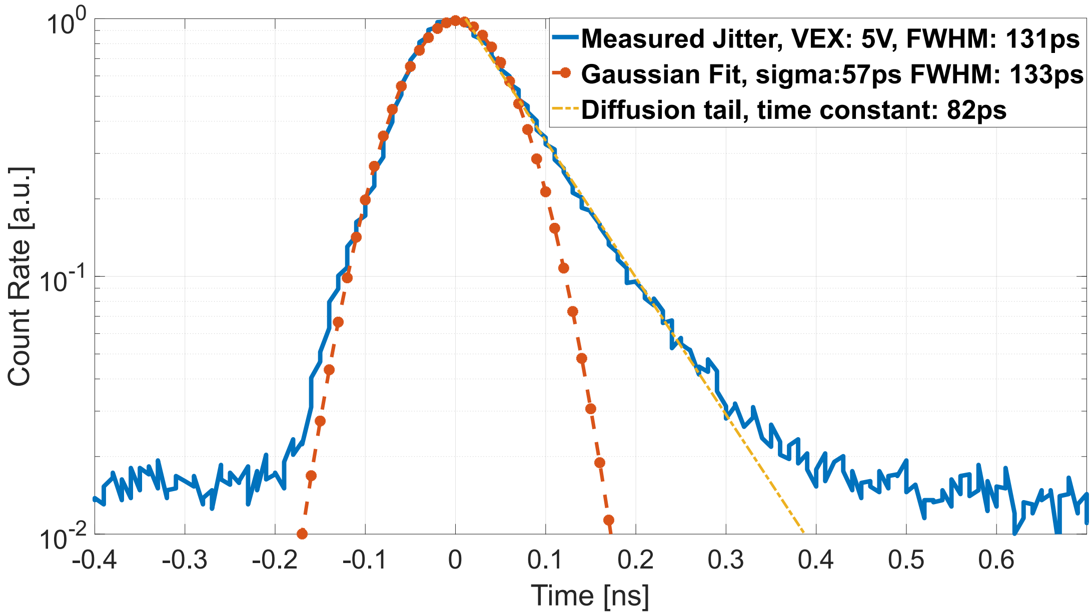


Figure 5.14: Timing jitter result from the W2 at 5 V excess bias at 1500 nm wavelength and 300K.

Finally, afterpulsing probability is measured in gated mode at 10 μs gate-off time, 100 ns gate-on time, and 5 V excess bias. Measurements are done from 225K to 275K with inter-arrival time histogram method. The time difference between the consecutive avalanche pulses is measured with an Oscilloscope and a histogram is formed based on the measurements. The obtained histogram at 225K is displayed in Figure 5.15. In order to extract the value of afterpulsing probability from the inter-arrival time histogram, an exponential fit is done at the tail of the curve. Primary DCR events are independent of each other. Therefore, they form a Poissonian distribution. On the other hand, afterpulses are correlated to the preceding avalanche event such that they cause disturbances in the inter-arrival timing histogram around the small inter-arrival times. The area between the measured curve and the exponential fit corresponds to counted pulses due to afterpulsing. The ratio between the area formed by afterpulses to the total measurement area gives the afterpulsing probability. Results of afterpulsing probability are demonstrated in Figure 5.16. 6.2 %, 4.4 %, and 3.3 % afterpulsing probabilities are measured at 225K, 250K, and 275K, respectively. At lower temperatures carriers that are trapped during the initial avalanche pulse have less energy to escape. This enhances the trap lifetime and the afterpulsing at low temperatures.

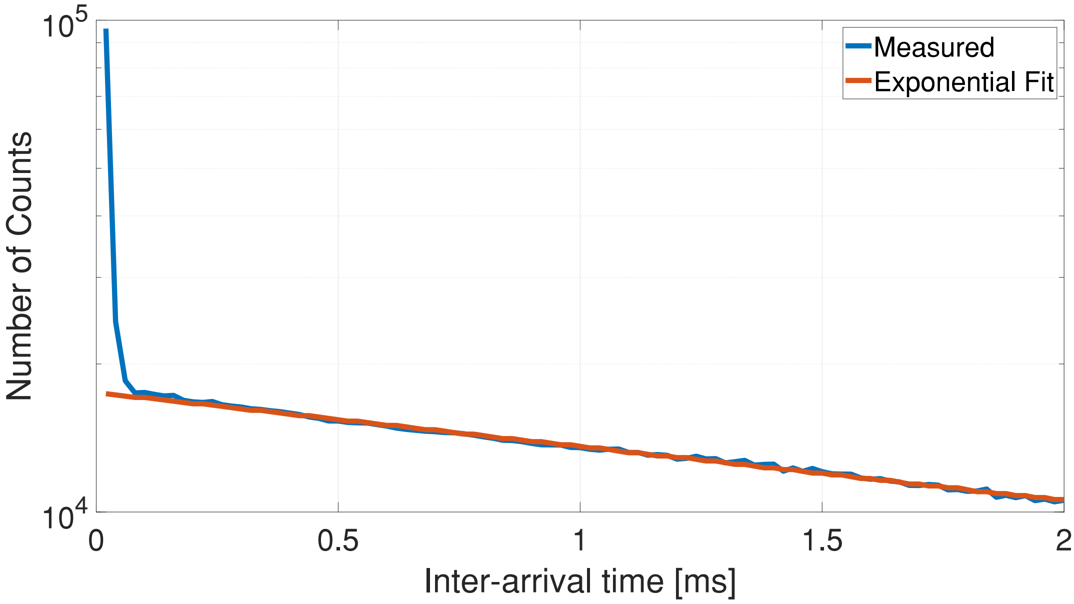


Figure 5.15: Inter-avalanche timing histogram from the W2 at 10 μs gate-off time, 100 ns gate-on time, and 5 V excess bias and 225K.

5.4 Discussion and Comparison with the State-of-the-Art

This chapter presented the performance results from three design variants based on double-diffusion type InGaAs/InP SPADs. The first parameter investigated is the shallow diffusion depth change for the same deep diffusion depth. Measurement results show that when the depth difference between shallow and deep diffusions is increased to 500 nm from 250 nm, DCR increases significantly. Moreover, the breakdown voltage is reduced by 7 V in the case of a 500 nm depth difference. This is an unexpected change since the breakdown voltage should be established by the depth of deep diffusion and since it stayed the same, the breakdown

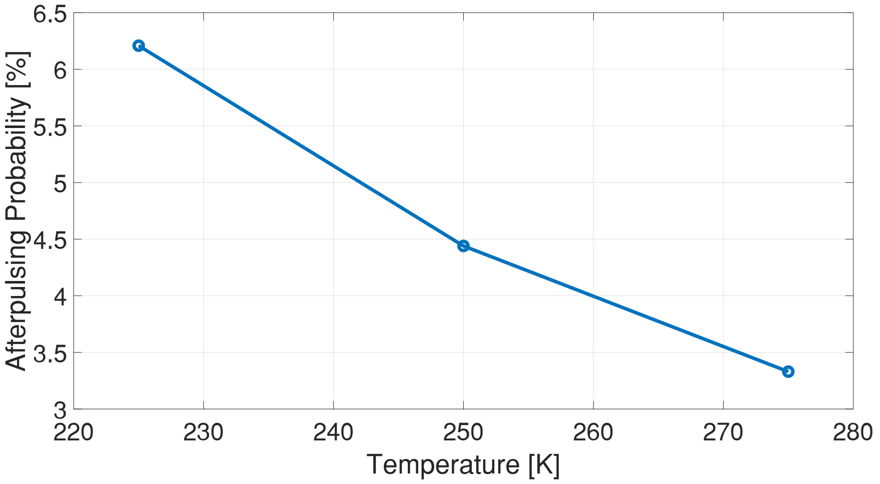


Figure 5.16: Afterpulsing probability at different temperatures from the W2 at 10 μs gate-off time, 100 ns gate-on time, and 5 V excess bias.

voltage should stay the same as well. The shift to a lower breakdown voltage is attributed to the enhanced curvature of the deep diffusion. Although simulation results cannot capture the lower breakdown voltage, they do show the electric field is higher around the deep diffusion periphery inside InP. The enhanced field results in higher TAT generation in InP and higher DCR even at low temperatures.

The second design aspect involves adjusting the depth of the deep diffusion, maintaining a fixed 250 nm difference from the shallow diffusion. Deeper active diffusion reduces the thickness of the avalanche region, leading to a lower breakdown voltage. This change results in a higher overall electric field, thus affecting the InGaAs absorber. Simulations at breakdown voltage reveal that the electric field in the InGaAs absorber stays below $1\text{x}10^5$ V/cm for both designs, minimizing TAT in InGaAs. However, with a 1.2 μm avalanche region, the field approaches the critical level in Geiger mode, increasing TAT in InGaAs and leading to a higher DCR in devices with this avalanche region thickness.

The lowest DCR is obtained from the device fabricated from W2. Although the electric field inside the InGaAs absorber is lower compared to W3, it could be further decreased to minimize the TAT contribution from this region. It could be implemented by utilizing a higher charge doping in future fabrications. The characterized device from W2 has a 10 μm deep diffusion diameter and 20 μm shallow diffusion diameter. The active area of this device is quite uniform and confined within the diameter defined by deep diffusion, unlike the SAG-based version. For the latter, the active area is extended by ~ 15 µm compared to the area defined by the layout. The SAG device is promising for achieving a better fill factor since it does not have an inactive area due to shallow diffusion. However, optimizing its diffusion profile is challenging compared to the double-diffusion approach.

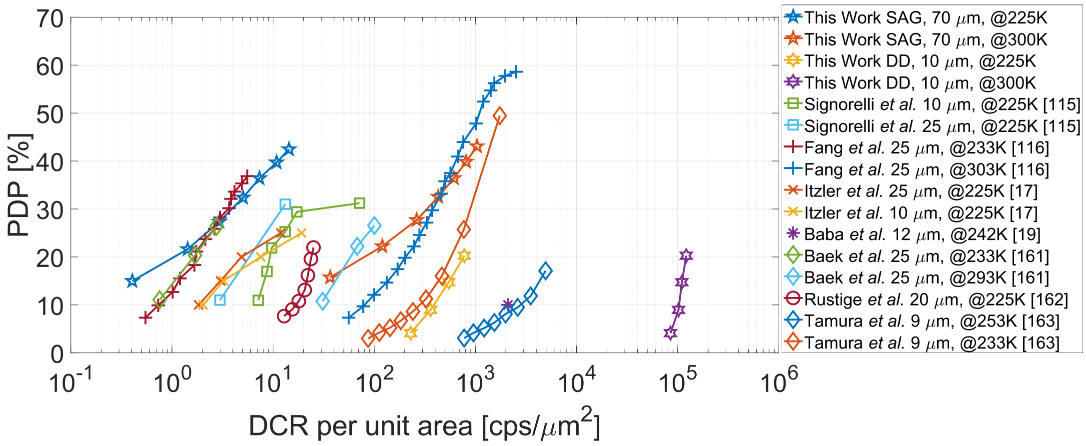


Figure 5.17: Comparison of DCR per unit area and PDP of the SAG-based and double-diffusion-based InGaAs/InP SPADs with the state-of-the-art.

Figure 5.17 and Table 5.3 compare the PDP and DCR per unit area of the SAG-based and double-diffusion-based devices demonstrated in this work together with other state-of-the-art devices from the literature [17, 19, 115, 116, 161–163]. The SAG-based device with a 70 μm demonstrated the lowest DCR per unit area at 225 K. However, DCR did not scale down with the active area as much as expected due to the changes in the diffusion profile at small-diameter devices. The diffusion profile in the SAG-based approach strongly depends on the surface pattern of the selective-area grown layer since Zn diffuses through this layer. This surface pattern changes as the device size shrinks. As a result, the electric field within the smaller diameter devices is not exactly the same as the 70 μm one. The double-diffusion-based device exhibited a higher DCR per unit area compared to the SAG-based one and the devices from the literature. It is expected to improve with higher charge doping and a lower electric field inside the InGaAs absorber. In terms of PDP, the SAG-based device demonstrated a quite high peak PDP at 1550 nm. This improvement is attributed to the thicker InGaAs absorber (1.8 μm). For the same number of photons, a thicker absorber region could capture more photons and achieve better efficiency. The PDP of the double-diffusion device is comparable to the other implementations in the literature. It is slightly lower than optimum due to the lack of back reflector metal. In most of the designs, the device surface or the back side, depending on the illumination, is covered with a metal. This causes the photons to reflect and do a second round-trip through the absorber region. In double-diffusion devices, a donut-shaped metal is preferred to allow for FSI testing as well.

Table 5.3: Summary and Comparison of InGaAs/InP SPAD Performance

Reference	Temperature [K]	Active Diam. [μ]	PDP [%] @1550 [nm]	DCR/area[d] [cps/ μm^2]	DCR[d] [cps]	Jitter [ps] @ PDP [%]
This Work SAG	300	70	15 - 43	36.7 - 1 k	141 k - 4 M	149 - 109 @32 - 43
	225	70	15 - 43	0.4 - 14.4	1.6 k - 55 k @15 - 43	N.A.
	300	21[c]	N.A.	1.4 k - 8.1 k	490 k - 2.81 M	N.A.
	225	21[c]	N.A.	4.9 - 74.5	1.7 k - 25.8 k	N.A.
This Work DD	300	10	4.1 - 20.3	85 k - 120 k	6.7 M - 9.5 M	125 @14.7
	225	10	4.1 - 20.3	230 - 777	18 k - 61 k	N.A.
Itzler *et al.* [17]	225	10	10 - 25	2 - 19	160 - 1.5 k	N.A.
	225	25	10 - 25	1.8 - 12	900 - 5.8 k	N.A.
Baba *et al.* [19]	242	12	10	2.1 k	234 k	260[a]
Signorelli *et al.* [115]	225	10	11 - 30	7 - 71.5	560 - 1.4 k	225 - 84 @11 - 30
	225	25	11 - 30	2.98 - 13.2	1460 - 6.5 k	340 - 119 @11 - 30
Fang *et al.* [116]	303	25	7 - 60	55.9 - 2.5 k	27 k - 1.24 M	N.A.
	233	25	7 - 37	0.5 - 5.5	260 - 2.5 k	N.A.
Baek *et al.* [161]	293	25	11 - 27	31 - 100	15 k - 50 k[b]	N.A.
	233	25	11 - 27	0.8 - 2.8	370 - 1.4 k[b]	N.A.
Rustige *et al.*[162]	225	20	13 - 25.1	7.6 - 22	4 k - 8k	349 - 129 @13 - 25.1
Tamura *et al.* [163]	253	9	1.7 - 49.5	1.2 k - 4.9 k	443 k - 2.8 M	430 @17
	233	9	5.2 - 49.5	143 - 1.7 k	50 k - 1 M	400 @49

[a] Performed at @ 1.6 V_{EX}. [b] Calculated from dark count probability and 2 ns pulse width. [c] Active area value obtained from the scan result. [d] DCR values are at the excess biases that correspond to shown PDP values.

6 NIR Silicon SPAD

In this chapter, four types of SPADs implemented in the GF 55 nm BCD process are analyzed. Implementing SPADs in higher technology nodes makes them available for a broad range of designers thus accelerating the development of SPAD-based technologies. The main challenge in working with foundry processes instead of custom designs is the limited availability of doping options. A SPAD device designer has to work with available doping levels to come up with an optimum junction configuration to optimize device performance. TCAD simulations are crucial tools for this process making it possible to estimate the electric field and the breakdown probability distributions of a design. Consequently, the failure rate is significantly reduced and the number of fabrication runs to achieve optimum device performance is curtailed.

SPADs are investigated in two families characterized by deep and shallow junctions. Deep junctions exhibited a higher breakdown voltage due to their wide depletion region. This also enables them to reach higher PDP values at reasonable excess bias. The main achievement concerning deep SPAD work is the insertion of p-well (PW) layer substantially enlarging the photo collection region. As a result, the sensitivity of the device is significantly improved. The other achievement is the insertion of shallow p-well (SPW) layer to extend the drift region in shallow junction SPADs. This enables the device to operate at higher excess bias and to reach improved sensitivity. Employing shallow junction SPADs simplifies the design process for circuit designers due to their lower breakdown voltage. This property minimizes the risks associated with high-voltage handling, making them a safer alternative if compared to their deep junction counterparts. Throughout the chapter, SPAD performance metrics are detailed by explaining the measurement methods and presenting the obtained results. The reasons for performance improvements are explained with the support of TCAD simulations.

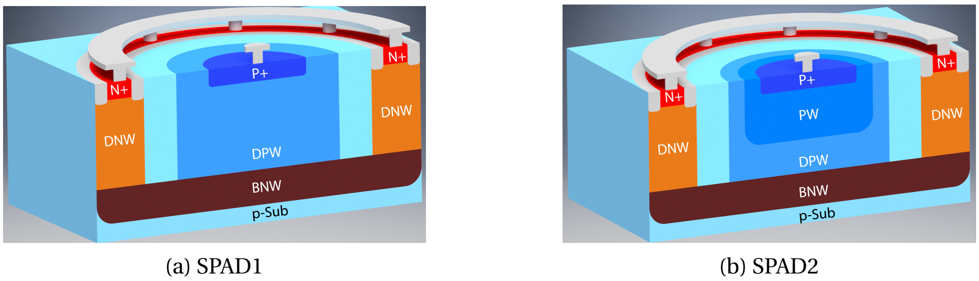
This work has been published in [164, 165]

The author contributed to this work by performing device simulations and device characterization.

6.1 Deep Junction SPADs

6.1.1 Device Structure

The general structure of the proposed SPADs is shown in Figure 6.1. The only difference between the variants is the additional PW implant in SPAD2.



(a) SPAD1 (b) SPAD2

Figure 6.1: Cross-sections of proposed 55 nm BCD Deep SPADs. A single implant (PW) highlights the difference between structures.

Due to the higher ionization coefficients of electrons compared to holes for silicon, it is desirable to allow the drift of electrons from the depletion region to the multiplication region [112]. A crucial and intuitive approach to implement this principle is by embedding the

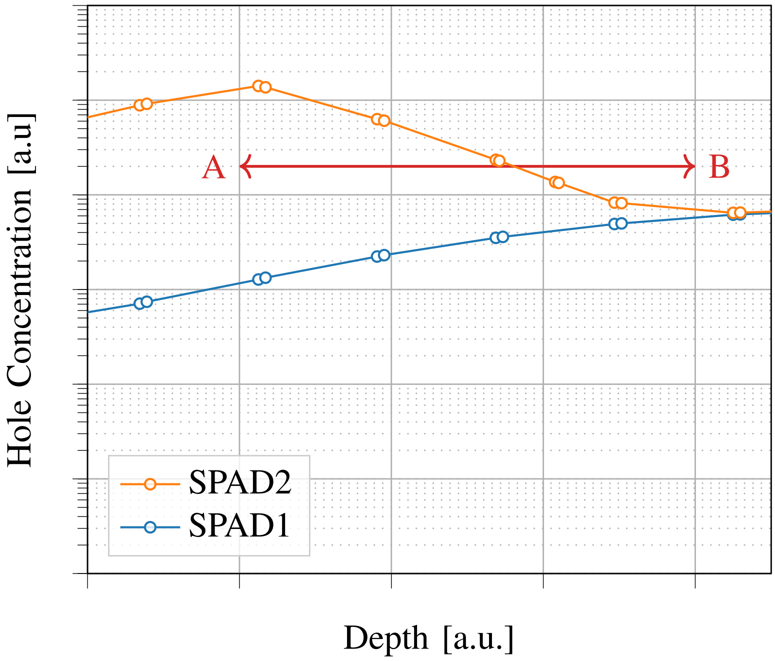


Figure 6.2: Doping comparison showing decrease of hole concentration for SPAD2 in contrast with increase shown for SPAD1 level. Net relative doping levels after all implants are placed is displayed at 5 V_{EX}.

junction deep inside the silicon, thereby enabling the minority electrons to get swept into the high-electric field where an avalanche can be initiated. A widely-used method for achieving a practical implementation of this deep junction involves the implantation of high-energy Boron ions, [87, 166, 167]. This method enhances the detection efficiency of low-energy red/NIR photons. Nonetheless, while the depth of the junction is vital for SPAD performance, even slight variations in the design of the quasi-neutral (photo-collector) region, i.e. the net implant profile approaching the junction, can produce significantly disparate results in terms of PDP. In an ideal scenario, the designer would manipulate the net doping to completely deplete the depth range where photons are absorbed. This enables the drift of generated charge carriers to the multiplication region, resulting in incremented breakdown probability and maximum PDP. However, the net doping combined displays variations due to the limited number of available masks in a standard process. In order to examine the impact of these variations on detection efficiency, it is possible to observe the band diagram in a TCAD simulation and, subsequently, simulate the corresponding breakdown probability.

6.1.2 Device Physics and TCAD Simulation

The concentration of majority carriers (holes) in two deep SPADs for a section of the surface near the transition from P+ to PW implants is displayed inFigure 6.2. The conduction band diagram is also presented in Figure 6.3. It is apparent that while SPAD2 displays a monotonic decrease in hole concentration towards the junction after a certain depth (as highlighted by point A), SPAD1 shows a concentration increase in a specific region. The rise of holes is due to

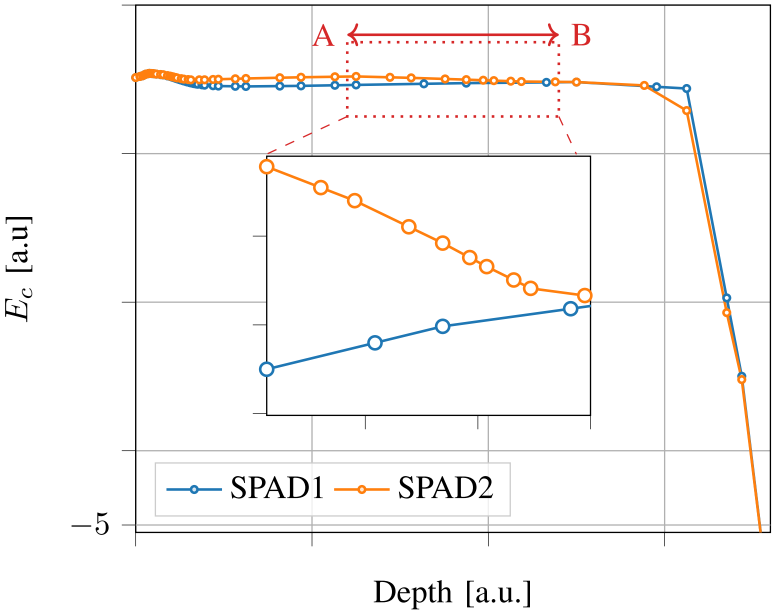
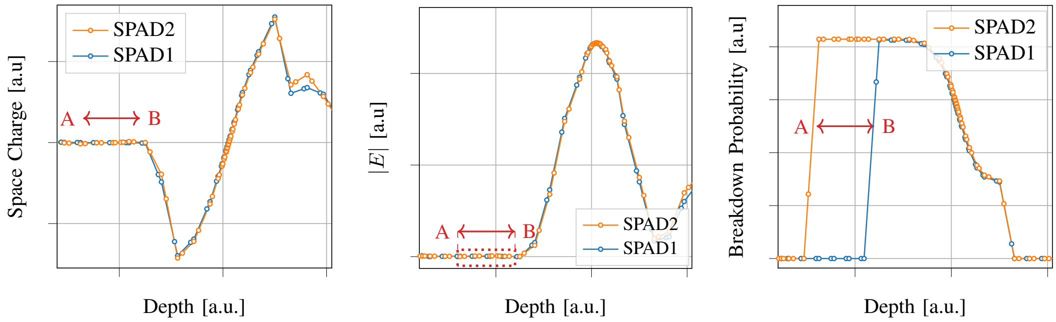


Figure 6.3: Simulated conduction band diagram of SPADs. Inset illustrates the drift barrier resulting from the net concentration difference between the proposed SPADs.



(a) Space charge illustrating nearly identical multiplication regions.

(b) Electric field magnitude, $|E|$, showing depth range of interest is outside the multiplication region.

(c) Simulated breakdown probability for each SPAD which culminates in the dramatic difference in performance between the two devices.

Figure 6.4: Comparison between the simulated junction parameters of SPAD1 and SPAD2 at breakdown. Space charge and E-field highlight nearly identical junction parameters. Normalized values are shown at 5 V_{EX}.

the overall concentration resulting from the profile of all implants employed. It is important to highlight that this depth range (A-B) lies beyond the high-field multiplication region illustrated in Figure 6.4a and Figure 6.4b. Various analytical demonstrations can shed light on the effect of this variation in doping. From the energy band diagram shown in Figure 6.3, it is clear that a barrier is formed, hindering the diffusion of electrons towards the multiplication region due to photons being absorbed near the surface. However, in the case of SPAD2, the effective photo-collector area widens, allowing photo-generated electrons in the charge-neutral region to move to the multiplication region. The outcome is a significantly enhanced breakdown probability. The principle is confirmed by the combined breakdown probability of both holes and electrons in simulations, as displayed in Figure 6.4c. The carriers that are introduced into SPAD2 have a greater chance of triggering an avalanche breakdown over a wider range of depths, resulting in a wider spectral response. The simulation clearly demonstrates that the space charge regions and electric field magnitudes of both devices are almost identical in the multiplication area. Hence, it is reasonable to deduce that variations in measured PDP would, to a large extent, be due to differences in doping in the photo-collector region.

6.1.3 I-V Characteristics

I-V curves were obtained with the Keysight B1500A semiconductor analyzer. The resulting dark currents and breakdown voltages (~32 V) agree well with the simulation conducted in TCAD and are illustrated in Figure 6.5. Notably, as anticipated by TCAD, SPAD2 exhibits a greater photo-current near the breakdown voltage. Initially, light emission testing was performed to qualitatively ascertain the absence of edge breakdown in both devices. Micrographs of SPAD1 and SPAD2, as well as the light emission test (LET) image, demonstrate uniformity around an active diameter of ~ 4.4 μm in Figure 6.6. The LET images are acquired using a microscope.

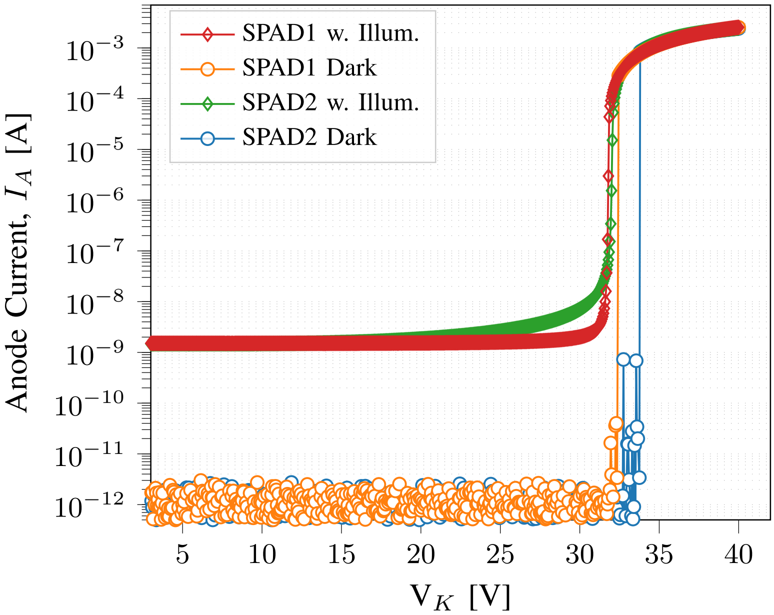
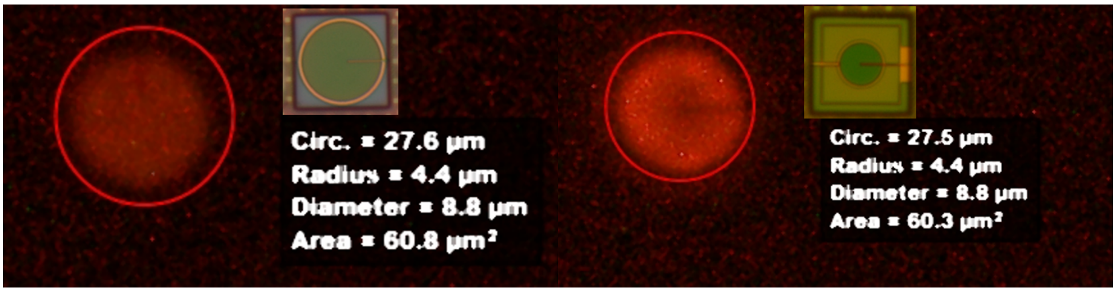


Figure 6.5: I-V curves for proposed SPADs in dark and illuminated conditions at room temperature.



(a) SPAD1 (b) SPAD2

Figure 6.6: Light emission test images with micrographs of SPAD1 (left) and SPAD2 (right) taken at $V_{EX} = 3$ V

The microscope illumination is turned off and SPAD is kept above breakdown voltage without a quenching resistor. The current is limited to 5 mA in order to avoid damaging the SPAD. Images are captured by a Silicon CCD camera with a long enough integration time.

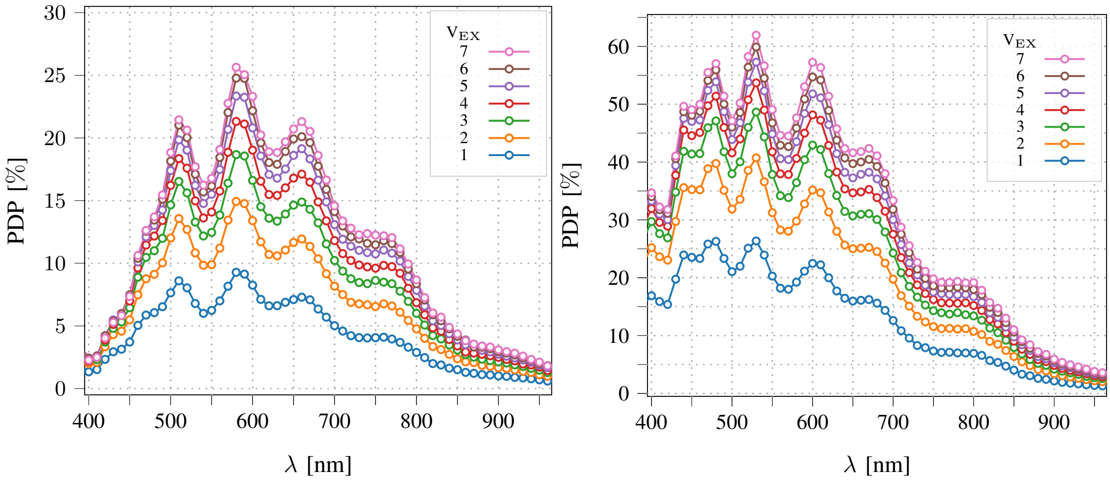
6.1.4 Photon Detection Probability

The results of the PDP measurements for the two suggested SPADs at excess bias voltages ranging from 1-7 V are presented in Figure 6.7a and Figure 6.7b. The measurements were conducted at room temperature with 10 nm intervals utilizing the continuous light technique [168]. Our setup employs a monochromator and integrating sphere to select a temporally stable and spatially uniform distribution. A reference photodiode (Hamamatsu s2281) and device under test (DUT) are placed at a calibrated distance from the light output. The SPAD's proper response is verified by monitoring the output using an oscilloscope and a universal counter (Keysight 53230A). SPAD2's sensitivity has been overall improved, as suggested by the simulation. Additionally, the low detection efficiency of SPAD1 for blue wavelength photons, which has a peak PDP of 26% (580 nm), reveals a significant performance difference. In contrast, SPAD2 shows a superior performance of 62% (530 nm) with an excess bias of 7 V and sustains > 19 % PDP up to 800 nm. To make a more precise comparison of both SPADs across excess bias and wavelength, we refer to Figure 6.8a and Figure 6.8b, which illustrate the devices' results on the same plot for reference excess biases and wavelengths. The graphs presented depict the convergence of performance between the two SPADs as longer wavelengths are used since lower energy photons penetrate deeper into the silicon and trigger avalanches with comparable probabilities. The standing wave pattern apparent across the visible spectrum arises from an optical stack unsuited for silicon.

6.1.5 Noise Performance

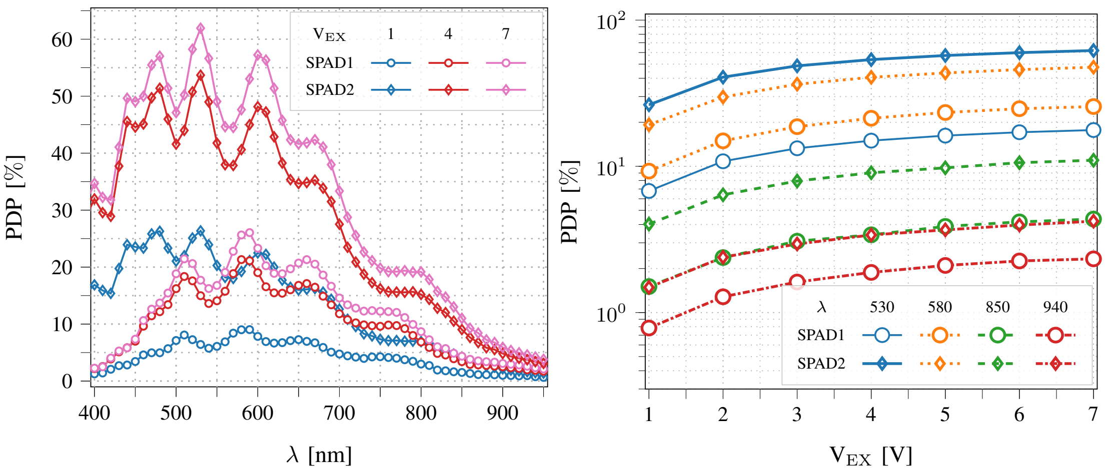
DCR

The primary and secondary pulses of a single-photon avalanche diode characterize its noise, which, when combined, constitute the DCR in the dark. The major contributors to primary pulses are thermally generated carriers and band-to-band tunneling. When an avalanche or thermal generation is initiated, carriers can become trapped in "deep" levels due to defects in the silicon process. These carriers subsequently release and cause an avalanche, leading to the formation of secondary pulses known as afterpulses. The DCR was analyzed using an oscilloscope, specifically the Teledyne LeCroy WaveMaster 813 Zi-B. Figure 6.9 displays the



(a) SPAD1 (b) SPAD2

Figure 6.7: PDP measurements for several excess bias voltages, measured with a 220 kΩ quenching resistor at 20°C for both SPAD designs presented in this work.



(a) Measurements at 1, 4, 7 V excess biases.

(b) Measurements at select wavelengths.

Figure 6.8: PDP comparison between the two proposed SPADs highlighting variation across wavelength and excess bias.

median result across 10 dies for both SPADs at an excess bias voltage range of 1-7 V. The results obtained from SPAD2 indicate a count rate of approximately 1 cps/μm^2 at 4 V excess bias and around 2.6 cps/μm^2 at 7 V excess bias. Due to the superior PDP of SPAD2 as compared to the other SPADs, the main focus was on the characterization of SPAD2, as both SPADs exhibit relatively low DCR. In order to conduct measurements, a single SPAD2 die was tested across a range of temperatures from -60°C to 60°C, and the outcome is depicted in Figure 6.10. The findings demonstrate promising performance even at high temperatures and with excess bias, with a value of <10 kcps at 60°C and V_{EX} = 7 V. The noise is mainly due to trap-assisted thermal generation, as evidenced by the strong temperature dependence of the DCR.

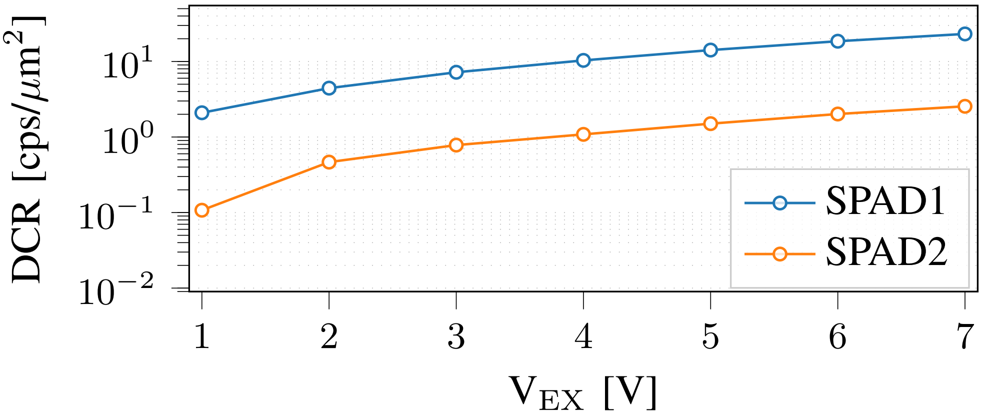
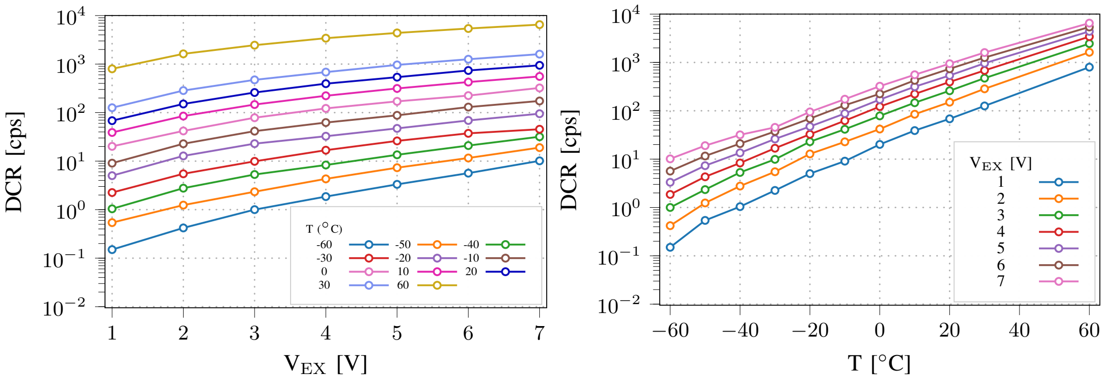


Figure 6.9: Normalized DCR comparison at room temperature across excess bias values. Median value shown for 10 dies of each SPAD



(a) Sweep of temperature from -60 to 60 °C.

(b) Measured at 1-7 excess bias.

Figure 6.10: Temperature dependence of DCR for a single randomly selected SPAD2 die illustrating strong temperature dependence of DCR.

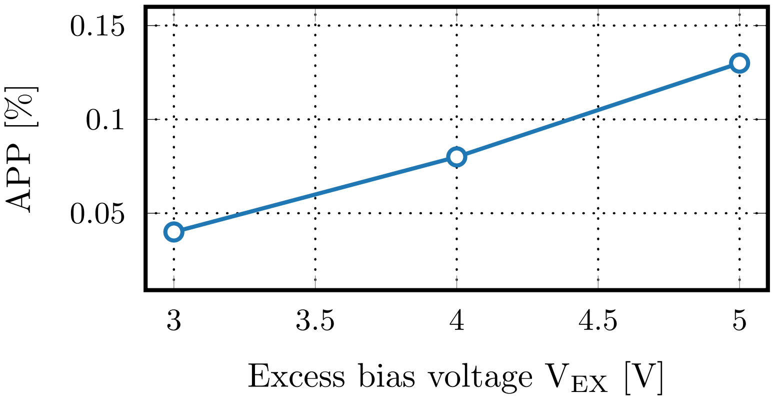


Figure 6.11: Afterpulsing probability of SPAD2 at room temperature at 3-5 V excess bias with integrated PQAR circuit and a dead time of ~1.5 ns.

Afterpulsing

The afterpulsing of silicon SPADs is usually analyzed by creating a histogram of the inter-arrival time between dark counts. It can also be estimated indirectly using the time-correlated carriers counting (TCCC) technique [103]. This technique is useful for determining the lifetime and density of traps. The method of inter-arrival histogramming was applied, using a passive quench-active recharge (PQAR) circuit, integrated on the same die at 3 V_{EX} to 5 V_{EX} and 50% range of signal swing. The design of the circuit is similar to the previous work represented in [64, 169]. The setup featured a dead time of approximately 1.5ns at the 50% level. The resulting afterpulsing probabilities with respect to excess bias are shown in Figure 6.11. This low afterpulsing probability is in line with the overall low DCR rate, indicating a relatively low concentration of defects.

Given that thermal generation is the primary factor in determining DCR, activation energies

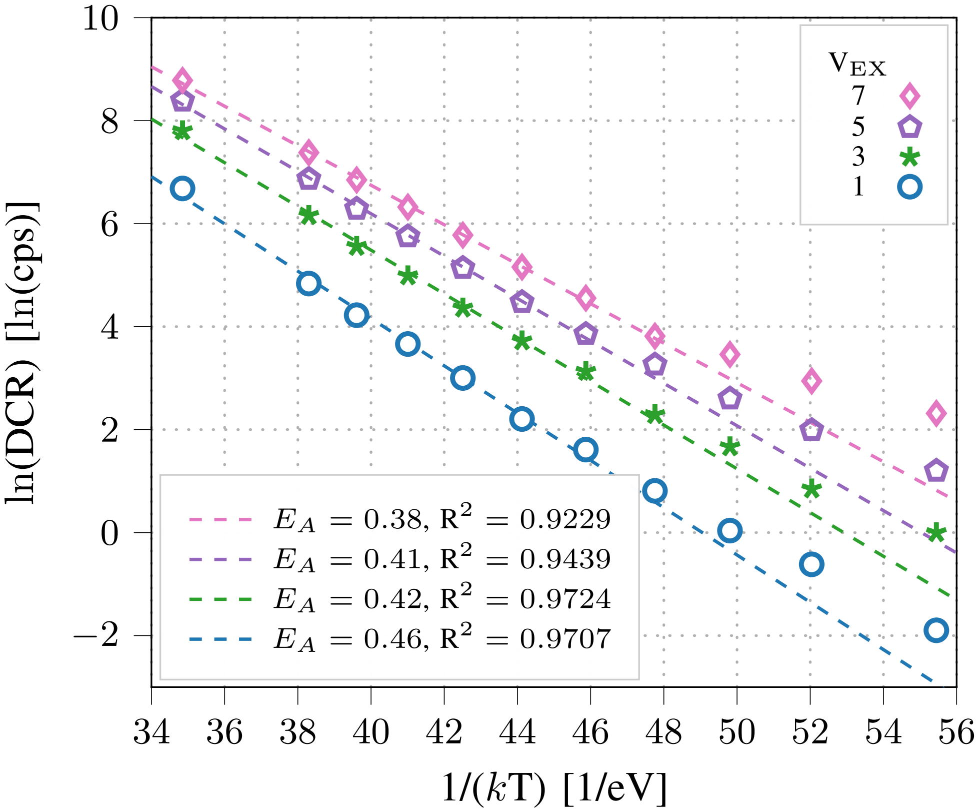


Figure 6.12: Arrhenius plot of the DCR for SPAD2 at the excess bias voltages of 1, 3, 5, and 7 V.

can be deduced from its temperature sensitivity [170]. This premise is justified in our case because of the low relative doping of the deep junction, which results in a lower electric field at breakdown (<1 MV/cm). Additionally, the strong relationship between DCR and temperature was previously highlighted in Figure 6.10. The Arrhenius diagram is illustrated in Figure 6.12. At elevated temperatures and with low excess bias, where SRH effects are more prominent, this qualitative analysis is employed to derive the activation energies of traps. The 0.46 eV value depicted at 1 V_{EX} can be attributed to phosphorus-vacancy defects [171], with the alteration in slope indicating an increased prevalence of tunneling effects at higher excess bias.

6.1.6 Timing Jitter

The timing performance of SPADs can be a crucial parameter for applications requiring precise sensing. The jitter results of SPAD2 are presented in Figure 6.13. To gather data, the time-correlated single-photon counting (TCSPC) acquisition method was employed, using the same oscilloscope used for the afterpulsing measurement, which operates as a time-to-digital converter (TDC). The characterization setup used is comparable to the one reported in [64]. A 1560 nm laser (NKT Onefive ORIGAMI-08) uses second harmonic generation in order to produce pulses at a 780 nm wavelength with a pulse width of 150 fs, repeated at 50 MHz. A 45

GHz optical receiver was added to act as a trigger on the 1560 nm branch. Additionally, neutral density filters were installed to reduce light intensity and ensure single-photon detection. Using the same PQAR circuitry at 3 V_{EX} to 5 V_{EX}. Additionally, the exponential time constant of the diffusion tail is 76 ps at 5 V_{EX}.

6.2 Shallow Junction SPADs

Figure 6.14 shows the cross-sections of two shallow junction SPADs designed in 55 nm BCD technology. SPAD3 utilizes a junction between p+ / DNW while for SPAD4 a slightly lower-doped SPW is inserted in between. Compared to the junction layers in deep SPADs, shallow SPADs utilize highly doped layers. As a result, the depletion region becomes narrower and breakdown voltages decrease to 18 V and 20 V for SPAD3 and SPAD4 concerning 32 V breakdown voltage observed in deep SPADs. The main benefit of shallow junction SPADs is to reduce breakdown voltage and simplify system design by avoiding high voltages. Concerning guard ring implementation methods, a donut-shape, low-doped PW layer is placed around the p+ layer to mitigate the electric field around the device periphery for SPAD3. For SPAD4, the SPW layer is extended to a wider size forming a junction between SPW and p-epi.

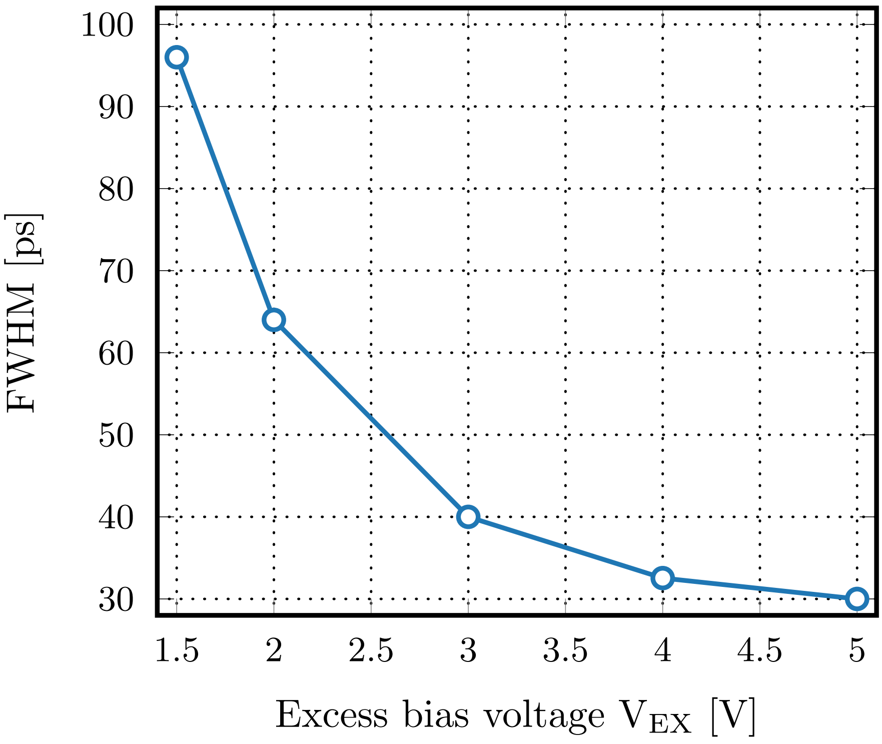
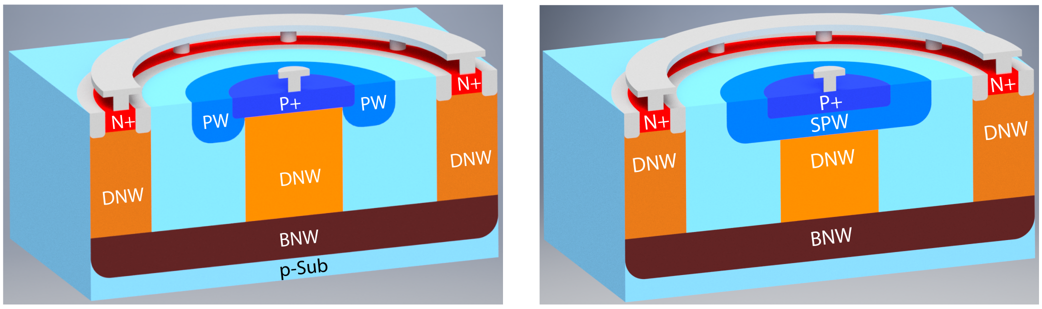


Figure 6.13: Jitter measurement of SPAD2 performed with a 780 nm fempto pulsed laser.



(a) SPAD3 (b) SPAD4

Figure 6.14: Cross-sections of proposed 55 nm BCD Shallow SPADs.

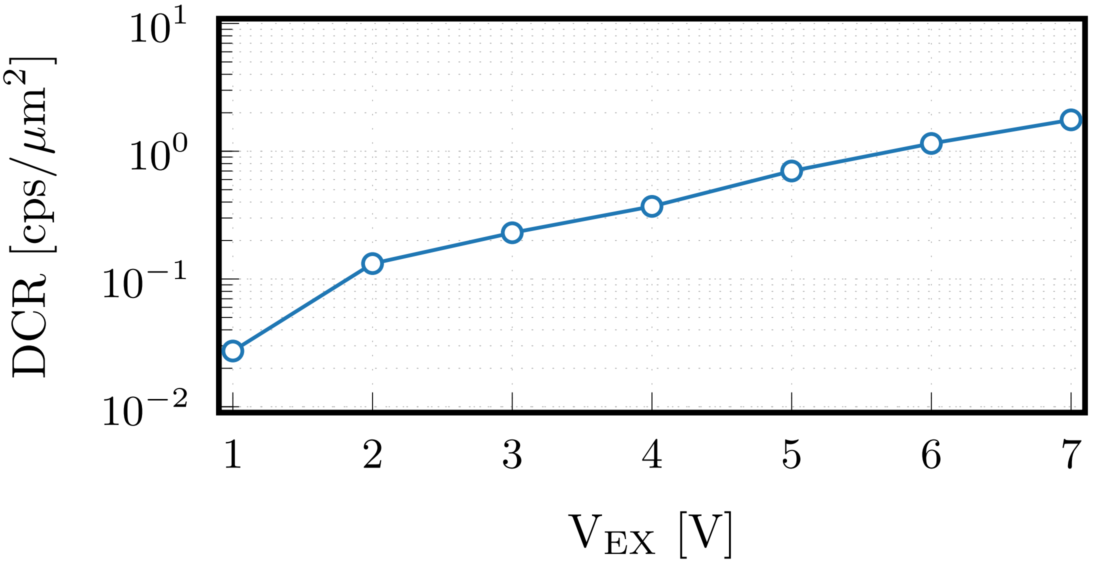
6.2.1 SPAD3

Noise

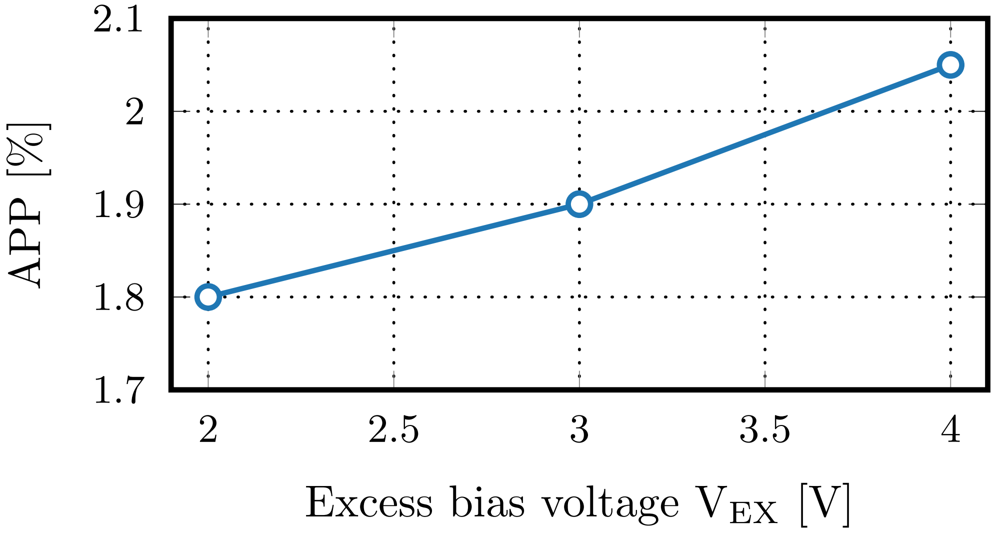
Figure 6.15a showcases a key feature of this design by depicting the normalized DCR over varying excess bias levels of a 4.5 μm active radius device. Notably, the SPAD3 maintains noise levels under 2 cps/μm^2 until reaching 7 volts of excess bias. The afterpulsing behavior of this SPAD is analyzed between 2 and 4 V_{EX}, as illustrated in Figure 6.15b. Compared to the deep junction SPAD previously discussed, this device exhibits a higher afterpulsing rate. However, an afterpulsing rate of 2% at 3 V_{EX} meets the requirements for various applications.

PDP and Jitter

Spectral PDP of the 4.5 μm active radius device is shown in Figure 6.16a. The measurements are done with a monochromator at 10 nm step, at 300 K, and using a 220 Kohm quenching resistor. Silicon has a higher absorption coefficient and lower penetration depth for short wavelengths (figurename 2.5). Therefore, a junction close to the surface is more efficient in collecting photocarriers generated by higher energy photons. This is evident when the spectral PDP curves of deep and shallow junction SPADs are considered. Compared to deep SPADs, the PDP peak shifts to the blue wavelengths for SPAD3 as the junction is close to the surface. However, peak PDP is also lower. This is essentially due to the design of the device. SPAD3 is an isolated SPAD with a junction close to the surface. As a result, the majority of the photocarriers that are collected from the bottom side of the junction will be holes. In Silicon, holes have a significantly lower impact ionization rate compared to electrons [172, 173]. Consequently reducing the overall breakdown probability for SPAD3. On the other hand, SPAD2 collects electrons from the upper part of the junction, achieving a higher overall breakdown probability. Unfortunately, it is not straightforward to obtain higher doping levels at deep locations in Silicon. Thus, it is difficult to implement a shallow isolated deep SPAD based on available foundry layers. A way of tackling this limitation would be to implement a non-isolated shallow junction SPAD. In this case, electrons would be collected by the upper junction resulting in a higher breakdown probability. However, this could also increase the DCR of the device if the



(a) DCR per unit area.

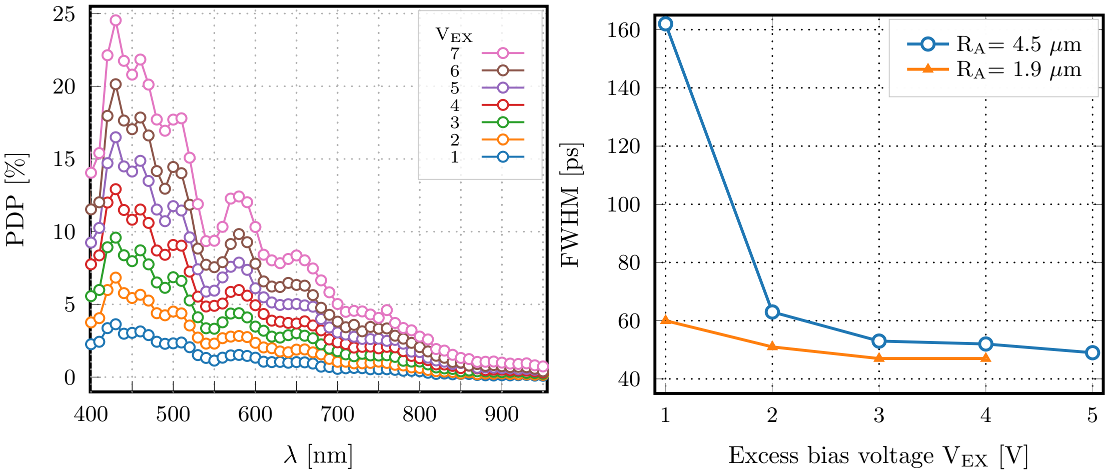


(b) Afterpulsing probability with respect to excess bias.

Figure 6.15: Noise performance of SPAD3. Afterpulsing measurement is performed with the inter-avalanche time histogram method. Both measurements are conducted at 300K.

substrate has a large number of traps.

As outlined in Chapter 2, the timing performance of SPADs is influenced by a variety of factors including doping levels, and the geometric design of the SPAD. The study examined the timing performance of two distinct types of SPAD3, one with an active radius of 1.9 μm and the other with 4.5 μm. Their respective results are presented in Figure 6.16b. The measurements were done with flood illumination at a wavelength of 780 nm under different excess biases. When operating at higher excess bias values, the FWHM of the timing jitter is comparably similar for both SPADs. However, at an excess bias of 1 V, the larger SPAD achieves a timing jitter of 160 picoseconds, while the smaller SPAD exhibits approximately 60 picoseconds. This finding effectively illustrates the impact of both excess bias and SPAD geometry on the timing jitter. Considering the Equation 2.15, jitter contributions from T_{diff} parameter will be different. At 1 V excess bias, the lateral drift speeds of carriers appear to be lower giving rise to a significant difference between the jitters of two devices. At higher excess biases, carriers probably reach a saturation velocity reducing the difference between the jitter results [174]. Nevertheless, the smaller radius version exhibits lower jitter at all excess biases due to the reduced travel distance. Similar results are also observed in reference [175].

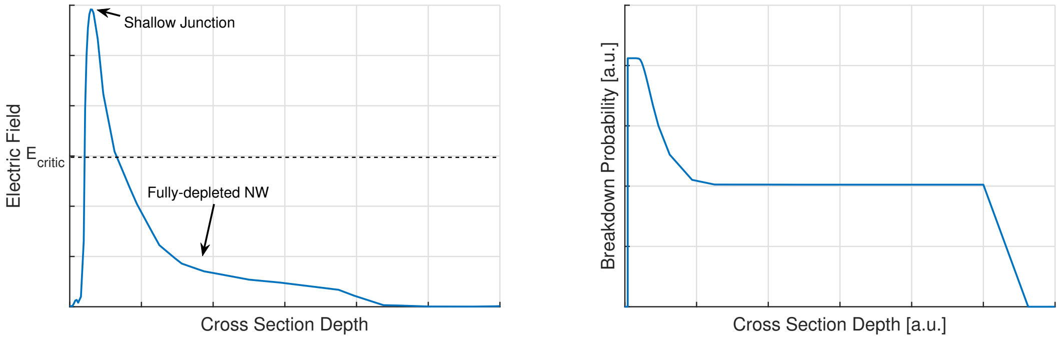


(a) PDP (b) Timing jitter

Figure 6.16: Photon detection probability and timing jitter results of SPAD3.

6.2.2 SPAD4

SPAD4 differs from SPAD3 by having the SPW layer at its junction. Slightly lower doping in SPW results in a higher breakdown voltage of 20 V compared to 18 V of SPAD3. Electric field and breakdown probability simulations are conducted in TCAD and presented in Figure 6.17. Both of the simulation results show that the avalanche region is confined in a shallow region but the drift region extends all the way to the buried n-well (BNW). Another remarkable observation on the breakdown probability curve is that it peaks near the surface, and then decreases as we move down through the avalanche region. Finally, as we pass the high field region, the breakdown probability stays flat. Since the impact ionization rates of electrons are higher in Silicon, probability peaks near the surface where electrons sweep down to the avalanche region. For example, in InGaAs/InP SPADs, we observe an opposite trend since the impact ionization rate of holes is higher. The breakdown probability stays flat in the drift region as all of the holes injected from the bottom of the device will travel through the same length and experience the same amount of high field.



(a) SPAD4 electric field cutline at the center. (b) SPAD4 breakdown probability cutline at the center.

Figure 6.17: Electric field and breakdown probability simulation results of SPAD4.

The SPAD's DCR has been quantified and is depicted in Figure 6.18. There's a consistent upward trend in the DCR as V_{ex} increases, peaking at 350 cps when the voltage reaches 18 V. The DCR of SPAD4 is lower than SPAD3 making it possible to reach much higher excess biases. At the maximum bias point, the noise level is approximately 34 cps/μm^2 considering the 1.8 μm active radius.

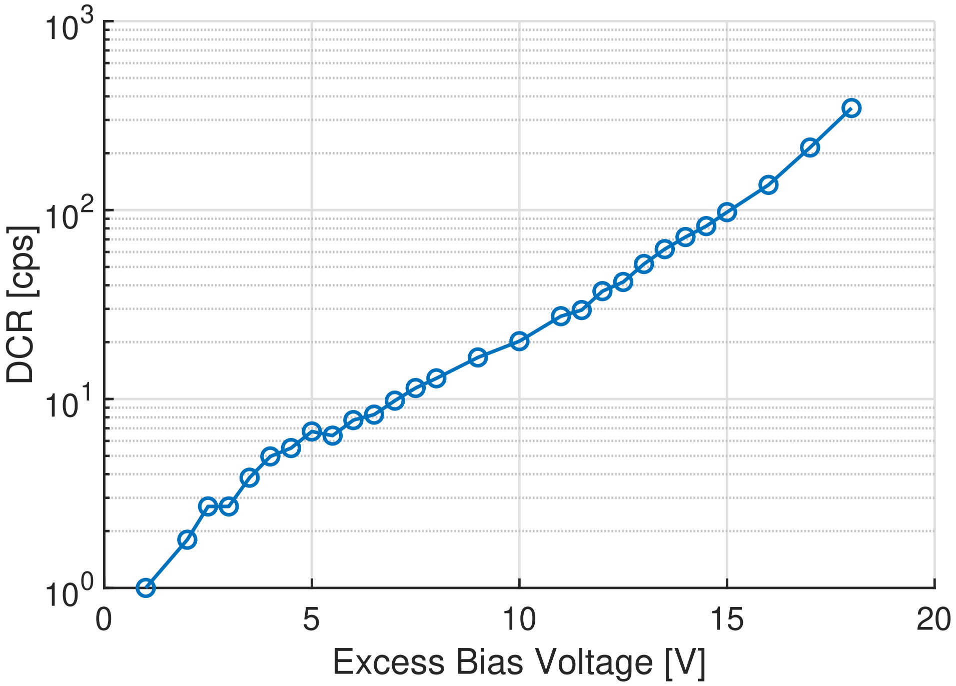
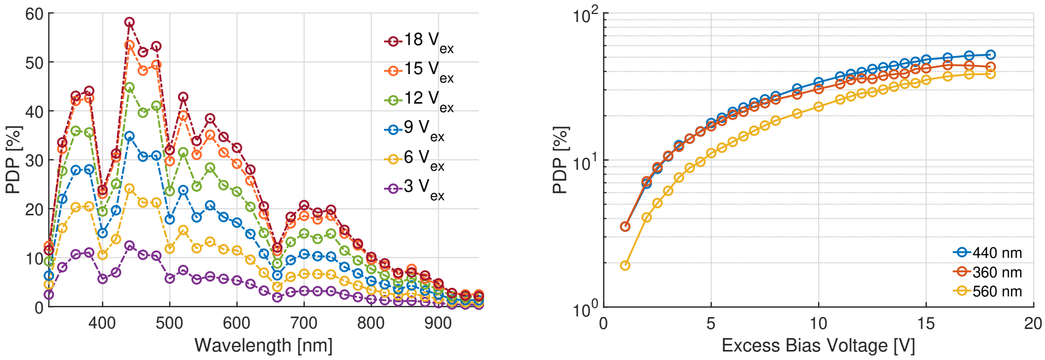


Figure 6.18: DCR result of SPAD4.

Figure 6.19a shows the spectral PDP result of SPAD4, while PDP change with excess bias is depicted in Figure 6.19b at some selected wavelengths. Both SPAD3 and SPAD4 exhibit a similar PDP peak of 10% at 3 V excess bias, but it increases to 60% for SPAD4 at 18 V excess bias. As breakdown probability saturates to 1 at elevated excess biases, PDP compression occurs as seen in the Figure 6.19b.



(a) Spectral PDP of SPAD4.

(b) PDP as a function of Vex for some selected wavelengths.

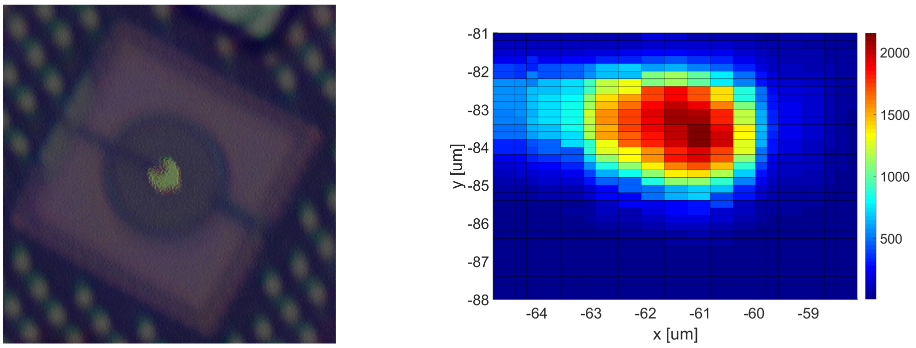
Figure 6.19: Photon detection probability results of SPAD4.

The uniformity of the SPADs response is justified with the light emission test in Figure 6.20a. Active area scans in Figure 6.20 are acquired using the confocal microscope. The scan is performed with a 100 nm step and a 405 nm picosecond pulsed laser at three different excess biases. Obtained scan results are able to capture the 3.6 μm active diameter of the SPAD. There is some distortion in the measurement results due to the very fine step of 100 nm. The motorized stage (PLSXY from Thorlabs) has a minimum repeatable step size of 1 μm which is significantly higher than the step size used in this experiment. Consequently, this creates positional errors in the measurement.

6.3 Discussion and Comparison with the State-of-the-Art

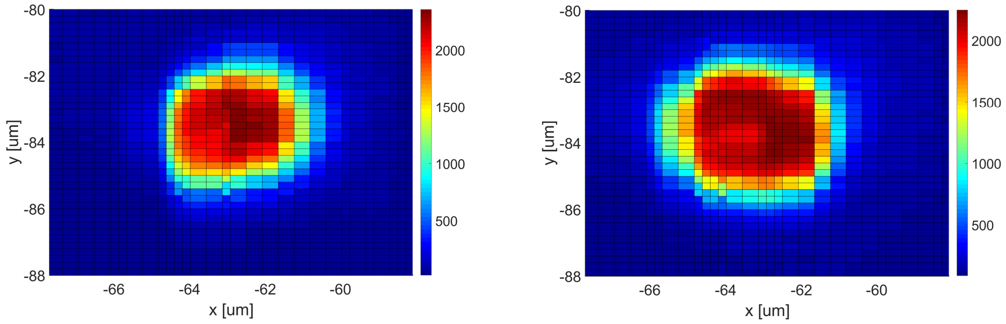
In conclusion, this chapter has provided a comprehensive analysis of the development and performance of NIR SPADs, emphasizing both deep and shallow junction SPADs. Through detailed discussions on device structures, physics, and various performance metrics such as PDP, noise performance, active area uniformity, and timing jitter, the chapter illuminated the significant advancements made in optimizing the efficiency of SPADs in advanced technology nodes.

Table 6.3 summarizes and compares the performance of various FSI SPADs against the state-of-the-art references. This table served as a useful tool for visualizing the evolution of SPAD technology, highlighting key achievements in terms of technological advancements, junction designs, and guard ring implementations. The chapter also sheds light on the challenges and limitations inherent in SPAD design, underscoring the necessity for ongoing research and development to push the boundaries of SPAD technology further.



(a) SPAD4 light emission test.

(b) SPAD4 active area scan at 5 V excess bias.



(c) SPAD4 active area scan at 10 V excess bias.

(d) SPAD4 active area scan at 15 V excess bias.

Figure 6.20: Light emission test and active area scans of SPAD4.

Table 6.1: Performance Summary and Comparison to FSI SPAD

Reference	Year	Tech. [nm]	Junc.	Guard Ring	Active Diam. [μm]	V_{EX} / V_{BD} [V]	Peak PDP [%] @λ [nm]	PDP [%] @940 [nm]	DCR / area [cps / μm^2]^{af}	DCR @V_{EX}^{f} [V]	AP [%]	Jitter [ps]	Jitter λ[nm]	Jitter V_{EX} [V]
Niclass et al. [176]	2007	130	p+/n	pw	10	1-2.7 /10	31-41 @450	1	1300	100k @1.7	N.A.	144	637	1.7
Gersbach et al. [177]	2008	130	p+/nw	STI^m	8.6	1-2 /9	18-30 @480	2	1.5k - 11.5k	670k @2	<1^g	125	637	1
Richardson et al. [178]	2009	130	pw/ DNW	N.A.^{c}	8	0.6-1.4 /14	18-28 @500	2	0.24 - 0.6	30 @1.4	0.02^h	200	815	N.A
Richardson et al.^{y} [167]	2011	130	pw/ DNW^v p+/nw	N.A.^{c}	8^x	0.2-1.2 /12-18	18-33 @450- 560	2-3	0.5-0.97	40-47 @0.8	0.02^i	183 - 237	470	1.2
Webster et al. [179]	2012	90	DNW/ p-epi^k	N.A.^{c}	6.4	2.46 /14.9	44 @690	12	~ 4.6	~150 @1	0.38^l	51	470	2.46
Webster et al. [88]	2012	130	p-epi^k	N.A.^{c}	8	2-12 /20	72 @560	12	0.36	18.0 @2	<4	60	654	12
Leitner et al. [180]	2013	180	p+/nw	N.A.^{c}	10	1-3.3 /21	35-47 @450	N.A.	0.3-1.8	~180 @4	N.A.	N.A.	N.A.	N.A.
Charbon et al. [181]	2013	65	n+/pw	nw	8	0.05- 0.4 /9	2-5.5 @420	0.2	340-15.6k	105 @.05	1^n	235	637	0.4
Villa et al. [182]	2014	350	p+/nw	pw	10-500	2-6 /25	37-55 @420	2	0.05	1.0 @4	1^o	56- 4470	780	6
Veerappan et al. [183]	2014	180	p+/nw	pw	12	2-10 /23.5	24-48 @480	N.A.	0.16-12.8	20 @2	0.03- 0.3^p	86	637	10
Lee et al. [184]	2015	140^d	p+/nw	pw	12	0.5-3 /11	10-25 @500	1-3	0.9-244	30k @3	1.7^p	65	405	3
Veerappan et al. [185]	2015	180	p+/nw	pw	12	1-4 /14	23-47 @480	N.A.	16	2k @4	0.2^q	95	405	4
Veerappan et al. [87]	2016	180	p-epi/ DNW	N.A.^{c}	12	1-12 /25	18-47 @520	N.A.	1.5	200 @11	7.2^q	97	637	11
Takai et al. [186]	2016	180	n/p^r	N.A.^{c}	16	1.5-5 /20.5	62 @210	8	0.5	100 @5	0.35	161	635	5
Pellegrini et al. [187]	2017	130	pw/ DNW	N.A.^{c}	8	0.5 /14.2	43 @480	1.4	1.6	80 @0.5	0.08	100	N.A.	2.4
Xu et al. [188]	2017	150	p+/nw	N.A.^{c}	10	5 / 18.0	31 @450	2	0.4	200 @5	0.85	42	831	4

Continued on next page

Table 6.1 – continued from previous page

Pellegrini et al. [189]	2017	40	pw/ DNW	N.A.[c]	N.A.	1.0 /15.5	45[s] @500	3	N.A.	50 @1.0	0.1	140	850	1
Sanzaro et al.[y] [166]	2018	160[e]	p+/n p/ DNW	N.A.[c]	10-80	3-9/ 25-36	58-71 @450-490	3	0.1-0.2	100 @5	0.02-1.59	28-41	820	5
Moreno-Garcia et al. [190]	2018	110	p+/ LDNW	N.A.[c]	10	8/20	58 @455	N.A.	1.16	450 @6	N.A.	71	831	4
Vornicu et al. [191, 192]	2020	110	p+ / DNW	polySi	5	5/18	67.4 @500	N.A.	2.61	204.5 @5	0.08	92	850	3
This Work (SPAD1)	2021	55[e]	DPW/ BNW	N.A.[c]	8.8	7/ 31.5	27 @530	4.2	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
This Work (SPAD2)	2021	55[e]	DPW/ BNW	N.A.[c]	8.8	7/ 31.5	26-62 @530	4.2	0.1-2.6	6.1-156 @1-7	~0.97[t]	30	780	3
This Work (SPAD3)	2021	55[e]	p+/ DNW	PW	9-3.8	7/18	27 @370	0.8	0.47	30 @8	2	52	780	5
This Work (SPAD4)	2021	55[e]	SPW/ DNW	N.A.[c]	3.6	18/ 20	60 @440	2.6	34	350 @18	N.A.	230	780	10

[a] Taken at max excess bias if not range of excess bias values not specified. [c] Virtual guard ring structure. [d] Silicon-on-insulator. [e] BCD. [f] At 20°C. [g] 180 ns dead time. [h] 200 ns dead time. [i] 50 ns dead time. [k] Not substrate isolated. [l] 15 ns dead time. [m] Shallow trench isolation with passivation implants to create p-type glove structure. [n] At 1 μs dead time. [o] 30 μm diameter. [p] 200 ns dead time. [q] 300 ns dead time. [r] Surface-isolated n-spad/p-spad junction. [s] With microlens. [t] 4.5 ns dead time with 50 % level @ 3 V_{EX}. [v] Two different deep structures were presented one with an epi layer and one with a pw implant. [x] Multiple diameters demonstrated. [y] 3 SPAD structures proposed.

7 Conclusion and Future Work

This thesis focuses on developing high-performance single-photon detectors for QKD and LiDAR applications, utilizing Silicon and InGaAs/InP-based SPAD technologies. The primary goal is to create compact single-photon detectors capable of operating near room temperature, or within the range achievable by thermoelectric cooling. The research includes a comprehensive characterization of SAG-based and double-diffusion-based InGaAs/InP SPADs, as well as Silicon SPADs designed using 55 nm BCD technology. Key performance metrics such as DCR, PDP, timing jitter, and active area uniformity are evaluated. The study also involves analyzing design trade-offs and performance enhancements, backed by TCAD simulations.

Selective-area-growth based InGaAs/InP SPAD

The SAG-based design introduces a novel InGaAs/InP SPAD structure by eliminating the need for shallow diffused regions used as guard rings. This is achieved by initially growing an undoped InP layer, about 300 nm thick, at the device's center using selective area growth. The thickness of this layer increases towards the device's edge due to a higher growth rate. When Zn is diffused over this tapered surface, it creates a gradually transitioning doping profile within the InP, leading to a lower electric field at the device's edges compared to its center. The improvements in DCR and device uniformity are shown by characterizing a 70 µm diameter device. Its performance is compared with a device that utilizes floating guard rings and a DCR improvement of one order of magnitude is presented by eliminating the floating guard rings. The device exhibited a quite uniform active area except for some sidelobes occurring along [100] and [010] crystal directions due to deeper Zn diffusion. In order to further reduce the DCR, the performance of smaller-diameter devices is investigated. 17 µm diameter device demonstrated a strong response around the device periphery because of the not optimized SAG thickness. For the 6 µm diameter device, responses around the edge seem to merge and overall device showed a relatively uniform active area. When its DCR is compared with the 70 µm diameter device which has 20 kcps DCR at 5 V excess bias at 225K, the 6 µm diameter device showed a 6 kcps DCR under the same conditions. A decrease in DCR is observed with reduced device size, but it does not scale directly with the active area. The fundamental reason is that the two devices have a different electric field profile due to the changes in the SAG layer. This is also evident from the breakdown voltage difference between the two. The 70 µm diameter device has a breakdown voltage of 65 V while the smaller one has a breakdown voltage of 57.4 V. This indicates that Zn diffused deeper for the small diameter device and reduced the breakdown voltage. Correcting the SAG thickness at smaller devices could potentially address this issue and improve the DCR.

There are two main research directions for future work. The first is to optimize the small device performance by adjusting the SAG thickness and reaching the DCR per unit area level of the 70 µm one. This small-diameter SPAD will be used for fiber-coupled QKD application. The fiber-coupling requirement sets a constraint on the minimum available device diameter. Considering the minimum distance between the fiber tip and the device, a SPAD diameter of around 10 - 15 µm is targeted. The goal is to reach maximum operation frequency at a given DCR level. Under the same afterpulsing probabilities, the device with a lower primary DCR will allow operation at a higher frequency. Another way of lowering the afterpulsing contribution is to reduce the amount of charge flowing during the avalanche. In future designs, it will be achieved by lowering the parasitic capacitance between the counting circuit and the SPAD. The second research direction is the development of a SAG-based InGaAs/InP SPAD array that is 3D integrated with an in-house developed ROIC. The current challenge in this development is handling the impact of high voltage on the indium bumps that connect two chips.

InGaAs/InP SPAD Simulation Environment

The developed simulation environment enables simulations of PDP, DCR, IV, and breakdown voltage. It differs from previous implementations by allowing 2D simulations and being fully performed with TCAD tools. Material parameters of InGaAs are not as well established as those of Silicon. Therefore, PDP simulations required a lot of material parameter research, especially for real and complex refractive indices. In order to simulate the DCR, initially TAT model parameters of InP are adjusted following the literature. Then, carrier lifetimes are changed so that the simulated DCR at room temperature fits the measured values. The carrier lifetimes depend on the crystal quality and it is difficult to estimate their exact values. The same lifetimes are also used to simulate the DCR at lower temperatures and results matched quite well at 250 and 275 K. At 225 K simulation underestimates the DCR. Considering the fact that at this temperature the TAT generation is the dominant mechanism, TAT parameters could be slightly inaccurate. They are adjusted following a 1D simulation work from the literature. Therefore, it may fail to capture the DCR increase around the device periphery. IV simulations indicate the dark current of the device is dominated by surface leakage but these carriers do not contribute to the DCR. In future work, process simulations will be performed instead of importing doping profiles to the TCAD. TCAD offers growth and diffusion process simulations allowing a more realistic representation of the physical device. Moreover, circuit-coupled simulations could simplify design procedures of quench and recharge circuits.

Double-diffusion-based InGaAs/InP SPAD

Double-diffusion-based InGaAs/InP SPADs incorporate two diffusion processes called shallow and deep processes. Neighboring devices are isolated with an intrinsic InP layer in between. Deep diffusion defines the active area and multiplication thickness of the device while shallow diffusion acts as a guard ring. It surrounds the deep diffusion area and reduces the electric field around the device periphery in order to eliminate premature edge breakdown. The area covered by the shallow diffusion is an inactive region. The photogenerated carriers created underneath cannot trigger avalanche pulses. Therefore, in an array configuration, the area covered by the shallow diffusion reduces the fill factor. However, using two diffusions provides better control over the electric field distribution within the device in contrast to SAG-based implementation. For example, deep and shallow diffusion depths could be adjusted simultaneously to minimize the impact of charge persistence [193].

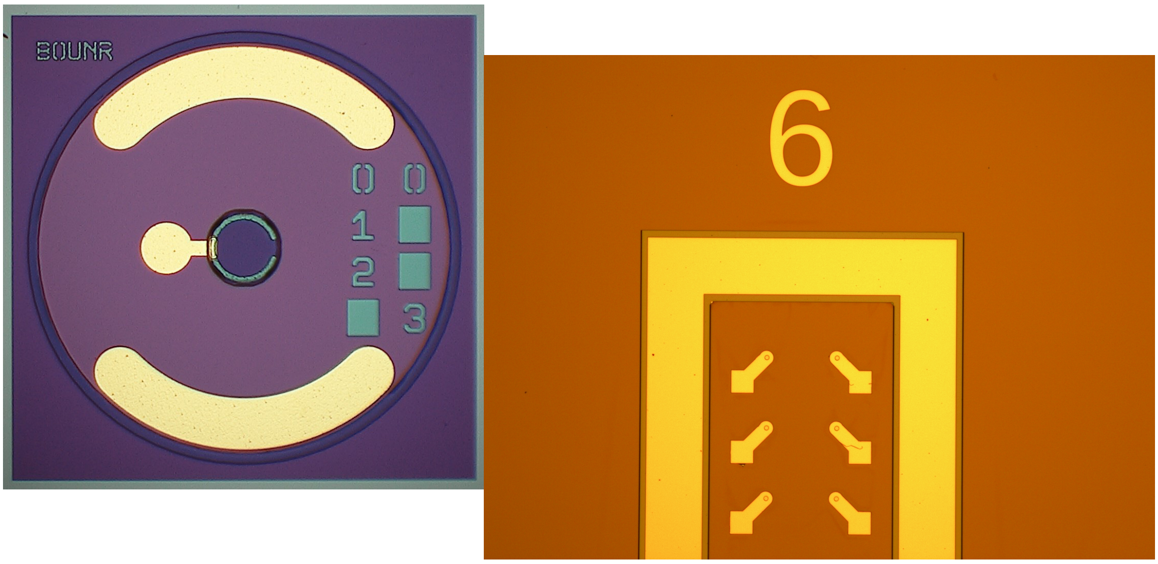
In this work, the effects of depth difference between shallow and deep diffusions and the multiplication thickness on DCR are investigated. Increasing the depth difference from 250 nm to 500 nm resulted in a more steep doping transition around the deep diffusion edge. This caused a higher electric field to emerge around this area and increased the TAT generation in InP. On the other hand, when the depth difference is kept at 250 nm and the depth of deep diffusion is increased to 1.8 μm, a significant DCR increase is observed. The electric field simulations show that for this device the field inside the InGaAs absorber reaches above $1\text{x}10^5$ V/cm in Geiger mode. Thus, TAT generation in InGaAs increases causing a higher DCR. The best performance is achieved with the device having 1.5 μm avalanche thickness and 250 nm depth difference between the deep and the shallow diffusions. This device demonstrated a DCR of 43 kcps at 225K at 5 V excess bias in gated mode which is two orders of magnitude better than the other two variants. Although its DCR is higher than the other implementations in the literature, with a 10 μm active area and 20 μm shallow diffusion diameter, it is one of the smallest InGaAs/InP SPADs available in the literature. Moreover, device performance in terms of PDP and timing jitter is comparable to other state-of-the-art devices.

In future work, the first goal is to further reduce the DCR of best performing SPAD by increasing the charge doping. This device has a breakdown voltage of 77.5 V at room temperature and it decreases to ~ 70 V at 225 K. The punch-through voltage is 42 V with the current charge doping. Increasing the punch-through voltage up to 70 V by using a higher charge doping would still ensure the InGaAs absorber stays depleted even at 225 K. In this configuration, the electric field inside the InGaAs absorber would be minimum eliminating TAT generation in this region completely. Similar to the SAG-based approach the other research direction is the development of SPAD arrays 3D bonded to the in-house developed ROIC. The current target array size is 96 × 96 with a pixel pitch of 25 μm and an active diameter of 10 μm. An early study of under-bump metal, indium bump deposition, and flip-chip bonding has already been made. Under-bump metallization prevents indium atoms from diffusing to the epitaxial layers during the high-temperature flip-chip bonding process. It is composed of 25 nm/35 nm/100 nm Ti/Ni/Au metal stack and deposited by e-beam evaporation prior to the indium deposition. ~ 5 μm thick indium bumps are deposited with thermal evaporation. During flip-chip bonding, both anode and cathode indium bumps must be at the same height to achieve an electrical connection. Therefore, a metal contact path from the chip surface to the buried N-InP has to be formed. This is done by two-step lithography and wet etching processes. The first lithography mask is used to etch the InP and grade layers. The second lithography mask is used to etch the InGaAs absorber and has a wider opening to obtain a staircase-like profile from the surface to the buried N-InP. Some test structures are vent through this process and electrical connection is verified by IV measurements on 3D bonded samples.

NIR Silicon SPAD

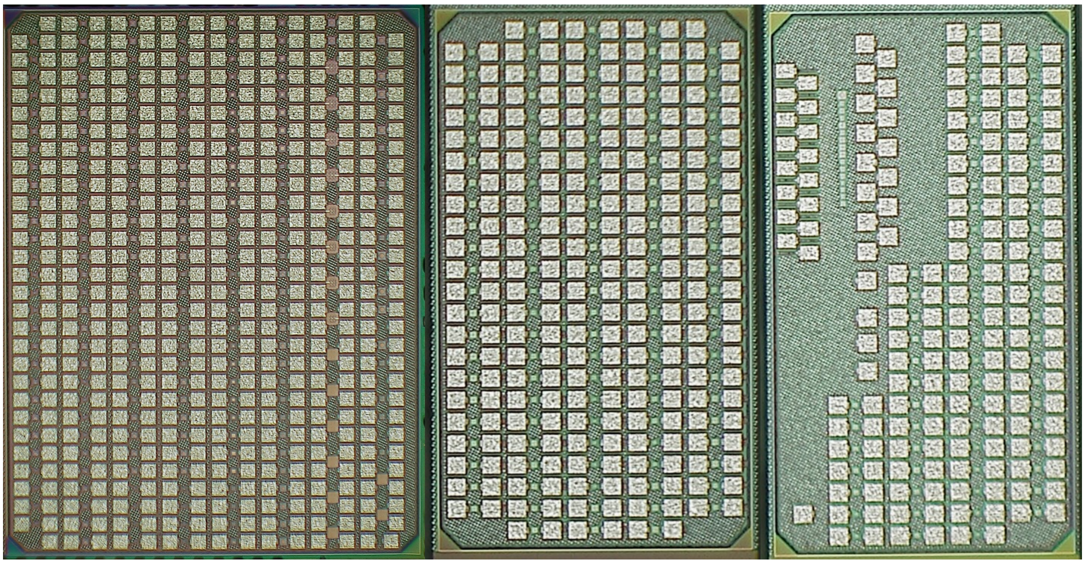
Silicon SPADs are relatively easier to implement thanks to the mature fabrication technologies compared to their InGaAs/InP-based counterparts. They provide advanced back-end techniques and very low defect ratios. Their efficiency in the NIR range is limited due to the large bandgap of Silicon, but allowing monolithic integration, being low cost and widely available makes Silicon SPADs an appealing alternative for LiDAR applications. Custom Silicon SPAD fabrication allows designers to change epi-Silicon thickness, doping depths, and values, or insert trenches with doped sidewalls. These options significantly enhance the design palette and make it possible to implement designs boosting efficiency or jitter at a particular wavelength. On the other hand, standard Silicon technologies allow for application-specific integrated circuit designs using the available SPADs within that particular technology. Moreover, SPADs implemented in standard technologies benefit from the wafer quality improvements pushed by other industries resulting in lower DCR. In this work, four different SPAD designs fabricated in 55 nm BCD technology are presented. Various design methods are introduced to overcome the design limitations imposed by working with standard doping layers. In particular, for the deep SPAD designs two design variations with and without the PW layer are characterized. Results revealed that the device with the PW layer has two times better PDP compared to the one without. The physical reasoning behind this performance improvement is analyzed with the help of TCAD simulations. They show that PW insertion widens the photo collection region resulting in higher PDP. The current research direction based on these monolithic Silicon SPADs is the implementation of BSI versions of them. BSI Silicon SPADs are formed in a 3D configuration similar to InGaAs/InP SPADs such that on one chip quenching, recharge, and counting circuits are implemented while only SPADs are placed on the other chip. This method improves both the fill factor and PDP. Since the circuits are not placed on the same pixel area, the extra area could be used by the SPADs resulting in a better fill factor. PDP improvements stem from the back reflection. Photons reflect from the top metal contact of the SPAD and perform an extra roundtrip on the active area. Consequently, absorbed photon density and PDP increase.

Chip Gallery



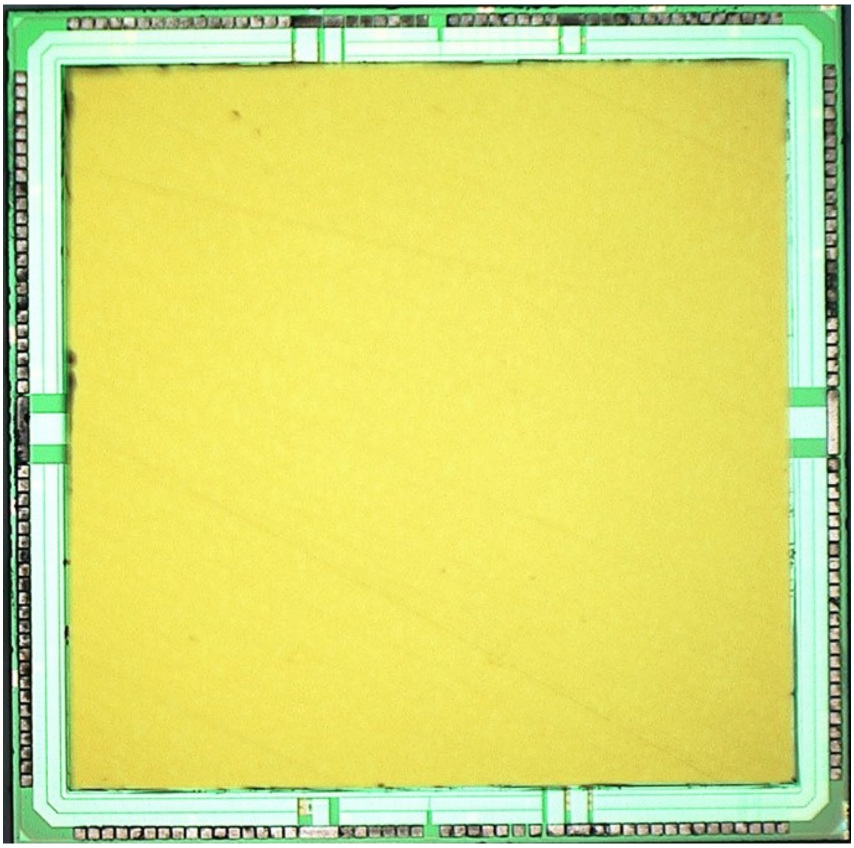
(a) (b)

Figure 7.1: a) SAG-based InGaAs/InP SPAD, b) Double-diffusion-based InGaAs/InP SPADs.



(a) (b) (c)

Figure 7.2: 55 nm BCD SPAD Farms.



(a)

Figure 7.3: Flipchip bonded SAG-based 96x96 InGaAs/InP SPAD array.

Bibliography

[1] Max Planck. “The theory of heat radiation”. In: *Entropie* 144.190 (1900), p. 164.

[2] Albert Einstein. “Über einem die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt”. In: *Annalen der physik* 4 (1905).

[3] Robert H. Hadfield. “Single-photon detectors for optical quantum information applications”. In: *Nature Photonics* 3.12 (Dec. 2009), pp. 696–705.

[4] Peter Zoller et al. “Quantum information processing and communication: Strategic report on current status, visions and goals for research in Europe”. In: *The European Physical Journal D-Atomic, Molecular, Optical and Plasma Physics* 36 (2005), pp. 203–228.

[5] Nicolas Gisin et al. “Quantum cryptography”. In: *Reviews of modern physics* 74.1 (2002), p. 145.

[6] BC Jacobs and JD Franson. “Quantum cryptography in free space”. In: *Optics Letters* 21.22 (1996), pp. 1854–1856.

[7] Akihisa Tomita and Kazuo Nakamura. “Balanced, gated-mode photon detector for quantum-bit discrimination at 1550 nm”. In: *Optics letters* 27.20 (2002), pp. 1827–1829.

[8] JD Franson and H Ilves. “Quantum cryptography using optical fibers”. In: *Applied Optics* 33.14 (1994), pp. 2949–2954.

[9] E Wu et al. “Narrow-band single-photon emission in the near infrared for quantum key distribution”. In: *Optics express* 14.3 (2006), pp. 1296–1303.

[10] Tobias Heindel et al. “Quantum key distribution using quantum dot single-photon emitting diodes in the red and near infrared spectral range”. In: *New Journal of Physics* 14.8 (2012), p. 083001.

[11] NA Peters et al. “Dense wavelength multiplexing of 1550 nm QKD with strong classical channels in reconfigurable networking environments”. In: *New Journal of physics* 11.4 (2009), p. 045012.

[12] Domenico Ribezzo et al. “QKD protocol over 100 km long submarine optical fiber assisted by a system-in-package fast-gated InGaAs single photon detector”. In: *Optical Fiber Communication Conference*. Optica Publishing Group. 2023, pp. M2I–4.

[13] Patrick Eraerds et al. “Photon counting OTDR: Advantages and limitations”. In: *J. Lightwave Technol.* 28.6 (Mar. 2010), pp. 952–964.

[14] Eleni Diamanti et al. “1.5 μm photon-counting optical time-domain reflectometry with a single-photon detector based on upconversion in a periodically poled lithium niobate waveguide”. In: *Optics letters* 31.6 (2006), pp. 727–729.

[15] Matthieu Legré et al. “High resolution optical time domain reflectometer based on 1.55 μm up-conversion photon-counting module”. In: *Optics express* 15.13 (2007), pp. 8237–8242.

[16] Xudong Jiang et al. “InGaAsP/InP Geiger-mode APD-based LiDAR”. In: *Optical Sensing, Imaging, and Photon Counting: From X-Rays to THz*. Vol. 10729. Sept. 2018.

[17] Mark A. Itzler et al. “SWIR Geiger-mode APD detectors and cameras for 3D imaging”. In: *Advanced Photon Counting Techniques VIII*. June 2014, pp. 37–48.

[18] Chao Yu et al. “Fully integrated free-running InGaAs/InP single-photon detector for accurate lidar applications”. In: *Opt. Express* 25.13 (June 2017), pp. 14611–14620.

[19] Takashi Baba et al. “Development of an InGaAs SPAD 2D array for flash LIDAR”. In: *Quantum Sensing and Nano Electronics and Photonics XV*. Vol. 10540. Jan. 2018.

[20] Aongus McCarthy et al. “Kilometer-range depth imaging at 1550 nm wavelength using an InGaAs/InP single-photon avalanche diode detector”. In: *Optics express* 21.19 (2013), pp. 22098–22113.

[21] Federica Villa et al. “CMOS imager with 1024 SPADs and TDCs for single-photon timing and 3-D time-of-flight”. In: *IEEE journal of selected topics in quantum electronics* 20.6 (2014), pp. 364–373.

[22] Alberto Ghezzi et al. “Multispectral compressive fluorescence lifetime imaging microscopy with a SPAD array detector”. In: *Optics Letters* 46.6 (2021), pp. 1353–1356.

[23] David Eric Schwartz, Edoardo Charbon, and Kenneth L Shepard. “A single-photon avalanche diode array for fluorescence lifetime imaging microscopy”. In: *IEEE journal of solid-state circuits* 43.11 (2008), pp. 2546–2557.

[24] Vytautas Zickus et al. “Fluorescence lifetime imaging with a megapixel SPAD camera and neural network lifetime estimation”. In: *Scientific Reports* 10.1 (2020), p. 20986.

[25] FX Kneizys, EP Shettle, and WO Gallery. “Atmospheric transmittance and radiance: The LOWTRAN 5 code”. In: *Atmospheric transmission*. Vol. 277. SPIE. 1981, pp. 116–124.

[26] Matthew H Smith, Richard L Fork, and Spencer T Cole. “Safe delivery of optical power from space”. In: *Optics Express* 8.10 (2001), pp. 537–546.

[27] R. Paschotta. “Infrared Light”. In: (May 2008). URL: https://www.rp-photonics.com/infrared_light.html.

[28] Robert H Hadfield et al. “Single-photon detection for long-range imaging and sensing”. In: *Optica* 10.9 (2023), pp. 1124–1141.

[29] Gordon A Thomas et al. “Towards the clarity limit in optical fibre”. In: *Nature* 404.6775 (2000), pp. 262–264.

[30] Hamamatsu Photonics. *Photomultiplier tubes basics and applications.* URL: https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99_SALES_LIBRARY/etd/PMT_handbook_v4E.pdf.

[31] M Niigaki et al. “Field-assisted photoemission from InP/InGaAsP photocathode with p/n junction”. In: *Applied physics letters* 71.17 (1997), pp. 2493–2495.

[32] Minoru Niigaki, Toru Hirohata, and Hidenori Mimura. “Room temperature photoemission up to a wavelength threshold of 2.3 μm from n+-InAs0. 4P0. 6/p–InAs0. 4P0. 6/p–ln0. 7Ga0. 3As field-assisted photocathode”. In: *Applied Physics Express* 7.11 (2014), p. 112201.

[33] Tomasz Szczesniak et al. “A comparative study of fast photomultipliers for timing experiments and TOF PET”. In: *IEEE Transactions on Nuclear Science* 56.3 (2009), pp. 1017–1023.

[34] Yifang Wang et al. “A new design of large area MCP-PMT for the next generation neutrino experiment”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 695 (2012), pp. 113–117.

[35] Yifang Wang. “Neutrino detectors: Present and future”. In: *Physics Procedia* 37 (2012), pp. 22–33.

[36] GN Gol’Tsman et al. “Picosecond superconducting single-photon optical detector”. In: *Applied physics letters* 79.6 (2001), pp. 705–707.

[37] WeiJun Zhang et al. “NbN superconducting nanowire single photon detector with efficiency over 90% at 1550 nm wavelength operational at compact cryocooler temperature”. In: *Science China Physics, Mechanics & Astronomy* 60 (2017), pp. 1–10.

[38] Peng Hu et al. “Detecting single infrared photons toward optimal system detection efficiency”. In: *Optics Express* 28.24 (2020), pp. 36884–36891.

[39] Iman Esmaeil Zadeh et al. “Single-photon detectors combining high efficiency, high detection rates, and ultra-high timing resolution”. In: *Apl Photonics* 2.11 (2017).

[40] Francesco Marsili et al. “Detecting single infrared photons with 93% system efficiency”. In: *Nature Photonics* 7.3 (2013), pp. 210–214.

[41] Dileep V Reddy et al. “Superconducting nanowire single-photon detectors with 98% system detection efficiency at 1550 nm”. In: *Optica* 7.12 (2020), pp. 1649–1653.

[42] J Chang et al. “Detecting telecom single photons with 99.5- 2.07+ 0.5% system detection efficiency and high time resolution”. In: *APL Photonics* 6.3 (2021).

[43] B Korzh et al. “Wsi superconducting nanowire single photon detector with a temporal resolution below 5 ps”. In: *CLEO: QELS_Fundamental Science*. Optica Publishing Group. 2018, FW3F–3.

[44] Boris Korzh et al. “Demonstration of sub-3 ps temporal resolution with a superconducting nanowire single-photon detector”. In: *Nature Photonics* 14.4 (2020), pp. 250–255.

[45] Iman Esmaeil Zadeh et al. “Superconducting nanowire single-photon detectors: A perspective on evolution, state-of-the-art, future developments, and applications”. In: *Applied Physics Letters* 118.19 (2021).

[46] Emma E Wollman et al. “Kilopixel array of superconducting nanowire single-photon detectors”. In: *Optics express* 27.24 (2019), pp. 35279–35289.

[47] Ling-Dong Kong et al. “Readout-efficient superconducting nanowire single-photon imager with orthogonal time–amplitude multiplexing by hotspot quantization”. In: *Nature Photonics* 17.1 (2023), pp. 65–72.

[48] Adam N McCaughan et al. “The thermally coupled imager: A scalable readout architecture for superconducting nanowire single photon detectors”. In: *Applied Physics Letters* 121.10 (2022).

[49] Bakhrom G Oripov et al. “A superconducting-nanowire single-photon camera with 400,000 pixels”. In: *arXiv preprint arXiv:2306.09473* (2023).

[50] Charles H Bennett and Gilles Brassard. “Quantum cryptography: Public key distribution and coin tossing”. In: *Theoretical computer science* 560 (2014), pp. 7–11.

[51] Shuang Wang et al. “Twin-field quantum key distribution over 830-km fibre”. In: *Nature photonics* 16.2 (2022), pp. 154–161.

[52] Shuang Wang et al. “2 GHz clock quantum key distribution over 260 km of standard telecom fiber”. In: *Optics letters* 37.6 (2012), pp. 1008–1010.

[53] Boris Korzh et al. “Provably secure and practical quantum key distribution over 307 km of optical fibre”. In: *Nature Photonics* 9.3 (2015), pp. 163–168.

[54] Hua-Lei Yin et al. “Measurement-device-independent quantum key distribution over a 404 km optical fiber”. In: *Physical review letters* 117.19 (2016), p. 190501.

[55] Alberto Boaron et al. “Secure quantum key distribution over 421 km of optical fiber”. In: *Physical review letters* 121.19 (2018), p. 190502.

[56] Yu-Ao Chen et al. “An integrated space-to-ground quantum communication network over 4,600 kilometres”. In: *Nature* 589.7841 (2021), pp. 214–219.

[57] Fadri Grünenfelder et al. “Fast single-photon detectors and real-time key distillation enable high secret-key-rate quantum key distribution systems”. In: *Nature Photonics* 17.5 (2023), pp. 422–426.

[58] Wei Li et al. “High-rate quantum key distribution exceeding 110 Mb s–1”. In: *Nature Photonics* 17.5 (2023), pp. 416–421.

[59] LC Comandar et al. “Quantum key distribution without detector vulnerabilities using optically seeded lasers”. In: *Nature Photonics* 10.5 (2016), pp. 312–315.

[60] Philip A Hiskett et al. “Eighty kilometre transmission experiment using an InGaAs/InP SPAD-based quantum cryptography receiver operating at 1.55 μm”. In: *Journal of Modern Optics* 48.13 (2001), pp. 1957–1966.

[61] N Namekata et al. “High-rate quantum key distribution over 100 km using ultra-low-noise, 2-GHz sinusoidally gated InGaAs/InP avalanche photodiodes”. In: *Optics express* 19.11 (2011), pp. 10632–10639.

[62] Nanxi Li et al. “A progress review on solid-state LiDAR and nanophotonics-based LiDAR sensors”. In: *Laser & Photonics Reviews* 16.11 (2022), p. 2100511.

[63] Behnam Behroozpour et al. “Lidar system architectures and circuits”. In: *IEEE Communications Magazine* 55.10 (2017), pp. 135–142.

[64] Francesco Gramuglia et al. “A Low-noise CMOS SPAD Pixel with 12.1 ps SPTR and 3 ns Dead Time”. In: *IEEE Journal of Selected Topics in Quantum Electronics* 28.2 (Mar. 2022), pp. 1–9.

[65] Maik Beer et al. “SPAD-based flash LiDAR sensor with high ambient light rejection for automotive applications”. In: *Quantum Sensing and Nano Electronics and Photonics XV*. Vol. 10540. SPIE. 2018, pp. 320–327.

[66] Preethi Padmanabhan et al. “7.4 A 256× 128 3D-stacked (45nm) SPAD FLASH LiDAR with 7-level coincidence detection and progressive gating for 100m range and 10klux background light”. In: *2021 IEEE International Solid-State Circuits Conference (ISSCC)*. Vol. 64. IEEE. 2021, pp. 111–113.

[67] Xudong Jiang et al. “InGaAsP/InP geiger-mode APD-based LiDAR”. In: *Optical Sensing, Imaging, and Photon Counting: From X-Rays to THz*. Vol. 10729. SPIE. 2018, pp. 33–44.

[68] Agata M Pawlikowska et al. “Long-range 3D single-photon imaging lidar system”. In: *Electro-Optical Remote Sensing, Photonic Technologies, and Applications VIII; and Military Applications in Hyperspectral Imaging and High Spatial Resolution Sensing II*. Vol. 9250. SPIE. 2014, pp. 21–30.

[69] Manijeh Razeghi. *Technology of quantum devices*. Springer, 2010.

[70] PA Wolff. “Theory of electron multiplication in silicon and germanium”. In: *Physical Review* 95.6 (1954), p. 1415.

[71] William Shockley. “Problems related to pn junctions in silicon”. In: *Solid-State Electronics* 2.1 (1961), pp. 35–67.

[72] Gene A Baraff. “Distribution functions and ionization rates for hot electrons in semiconductors”. In: *Physical review* 128.6 (1962), p. 2507.

[73] RJ McIntyre. “Multiplication noise in uniform avalanche diodes”. In: *IEEE Transactions on Electron Devices* 1 (1966), pp. 164–168.

[74] PP Webb. “Properties of avalanche photodiodes”. In: *RCA review* 35 (1974), p. 234.

[75] PP Webb and RJ McIntyre. “Single photon detection with avalanche photodiodes”. In: *Bulletin of the American Physical Society*. Vol. 15. 6. 1970, p. 813.

[76] Alexander W Lightstone et al. “Photon counting modules using RCA silicon avalanche photodiodes”. In: *NASA Laser Light Scattering Advanced Technology Development Workshop. Cleveland Ohio (Sept. 1988)*. 1989, pp. 65–80.

[77] Antonio Rogalski. *Infrared detectors*. CRC press, 2000.

[78] Edoardo Charbon and Matthew W Fishburn. “Monolithic single-photon avalanche diodes: SPADs”. In: *Single-Photon Imaging* (2011), pp. 123–157.

[79] Simone Tisa et al. “Electronics for single photon avalanche diode arrays”. In: *Sensors and Actuators A: Physical* 140.1 (2007), pp. 113–122.

[80] Sergio Cova et al. “Avalanche photodiodes and quenching circuits for single-photon detection”. In: *Applied optics* 35.12 (1996), pp. 1956–1976.

[81] RJ McIntyre. “Theory of microplasma instability in silicon”. In: *Journal of Applied Physics* 32.6 (1961), pp. 983–995.

[82] Francesco Ceccarelli et al. “Recent advances and future perspectives of single-photon avalanche diodes for quantum photonics applications”. In: *Advanced Quantum Technologies* 4.2 (2021), p. 2000102.

[83] Henri Dautet et al. “Photon counting techniques with silicon avalanche photodiodes”. In: *Applied optics* 32.21 (1993), pp. 3894–3900.

[84] Angelo Gulinatti et al. “New silicon SPAD technology for enhanced red-sensitivity, high-resolution timing and system integration”. In: *Journal of Modern Optics* 59.17 (2012), pp. 1489–1499.

[85] K Morimoto et al. “3.2 megapixel 3D-stacked charge focusing SPAD for low-light imaging and depth sensing”. In: *2021 IEEE International Electron Devices Meeting (IEDM)*. IEEE. 2021, pp. 20–2.

[86] Kazuhiro Morimoto et al. “Megapixel time-gated SPAD image sensor for 2D and 3D imaging applications”. In: *Optica* 7.4 (2020), pp. 346–354.

[87] C. Veerappan and E. Charbon. “A Low Dark Count p-i-n Diode Based SPAD in CMOS Technology”. In: 63.1 (Jan. 2016), pp. 65–71.

[88] Eric A. G. Webster, Lindsay A. Grant, and Robert K. Henderson. “A High-Performance Single-Photon Avalanche Diode in 130-nm CMOS Imaging Technology”. In: 33.11 (Nov. 2012), pp. 1589–1591.

[89] Myung-Jae Lee et al. “A 73% Peak PDP Single-Photon Avalanche Diode Implemented in 110 nm CIS Technology with Doping Compensation”. In: *IEEE Journal of Selected Topics in Quantum Electronics* (2023).

[90] Peter Vines et al. “High performance planar germanium-on-silicon single-photon avalanche diode detectors”. In: *Nature Communications* 10.1 (Mar. 2019).

[91] Manijeh Razeghi. *Fundamentals of solid state engineering*. Springer, 2006.

[92] Simon M Sze, Yiming Li, and Kwok K Ng. *Physics of semiconductor devices*. John wiley & sons, 2021.

[93] Sadao Adachi. *Physical properties of III-V semiconductor compounds.* John Wiley & Sons, 1992.

[94] TP Pearsall and M Papuchon. "The Ga0. 47In0. 53As homojunction photodiode—A new avalanche photodetector in the near infrared between 1.0 and 1.6 μm". In: *Applied Physics Letters* 33.7 (1978), pp. 640–642.

[95] Masanori Ito et al. "Tunneling currents in In0. 53Ga0. 47As homojunction diodes and design of InGaAs/InP hetero-structure avalanche photodiodes". In: *Solid-State Electronics* 24.5 (1981), pp. 421–424.

[96] Katsuhiko Nishida, Kenko Taguchi, and Yoshishige Matsumoto. "InGaAsP heterostructure avalanche photodiodes with high avalanche gain". In: *Applied Physics Letters* 35.3 (1979), pp. 251–253.

[97] Jesus A Del Alamo. *Integrated Microelectronic Devices: Physics and Modeling.* Vol. 25. Pearson, 2018.

[98] WTRW Shockley and WT Read Jr. "Statistics of the recombinations of holes and electrons". In: *Physical review* 87.5 (1952), p. 835.

[99] Godefridus Adrianus Maria Hurkx, DBM Klaassen, and MPG Knuvers. "A new recombination model for device simulation including tunneling". In: *IEEE Transactions on electron devices* 39.2 (1992), pp. 331–338.

[100] A. Schenk. "A model for the field and temperature dependence of Shockley-Read-Hall lifetimes in silicon". In: *Solid-State Electronics* 35.11 (1992), pp. 1585–1596. ISSN: 0038-1101.

[101] J.P. Donnelly et al. "Design Considerations for 1.06mum InGaAsP–InP Geiger-Mode Avalanche Photodiodes". In: *IEEE Journal of Quantum Electronics* 42.8 (July 2006), pp. 797–809.

[102] Matthew W Fishburn. *Fundamentals of CMOS single-photon avalanche diodes.* fishburn, 2012.

[103] S. Cova, A. Lacaita, and G. Ripamonti. "Trapping phenomena in avalanche photodiodes on nanosecond scale". In: 12.12 (Dec. 1991), pp. 685–687.

[104] G Vincent, A Chantre, and D Bois. "Electric field effect on the thermal emission of traps in semiconductor junctions". In: *Journal of Applied Physics* 50.8 (1979), pp. 5484–5487.

[105] Massimo Ghioni et al. "Large-area low-jitter silicon single photon avalanche diodes". In: *Quantum Sensing and Nanophotonic Devices V.* Vol. 6900. SPIE. 2008, pp. 267–279.

[106] Michele Anti et al. "Modeling of afterpulsing in single-photon avalanche diodes". In: *Physics and Simulation of Optoelectronic Devices XIX.* Vol. 7933. SPIE. 2011, pp. 371–378.

[107] KE Jensen et al. "Afterpulsing in Geiger-mode avalanche photodiodes for 1.06 μm wavelength". In: *Applied Physics Letters* 88.13 (2006).

[108] Alexei Trifonov et al. “Single photon counting at telecom wavelength and quantum key distribution”. In: *Journal of Modern Optics* 51.9-10 (2004), pp. 1399–1415.

[109] Mingguo Liu et al. “High-performance InGaAs/InP single-photon avalanche photodiode”. In: *IEEE Journal of selected topics in quantum electronics* 13.4 (2007), pp. 887–894.

[110] Mark A Itzler, Xudong Jiang, and Mark Entwistle. “Power law temporal dependence of InGaAs/InP SPAD afterpulsing”. In: *Journal of Modern Optics* 59.17 (2012), pp. 1472–1480.

[111] MS Tyagi and R Van Overstraeten. “Minority carrier recombination in heavily-doped silicon”. In: *Solid-State Electronics* 26.6 (1983), pp. 577–597.

[112] W. G. Oldham, R. R. Samuelson, and P. Antognetti. “Triggering phenomena in avalanche diodes”. In: 19.9 (Sept. 1972), pp. 1056–1060.

[113] Giulia Acconcia et al. “Timing measurements with silicon single photon avalanche diodes: principles and perspectives”. In: *Optics Express* 31.21 (2023), pp. 33963–33999.

[114] A Lacaita et al. “Observation of avalanche propagation by multiplication assisted diffusion in p-n junctions”. In: *Applied physics letters* 57.5 (1990), pp. 489–491.

[115] Fabio Signorelli et al. “Low-noise InGaAs/InP single-photon avalanche diodes for fiber-based and free-space applications”. In: *IEEE Journal of Selected Topics in Quantum Electronics* 28.2 (Mar. 2022), pp. 1–10.

[116] Yu Qiang Fang et al. “InGaAs/InP single-photon detectors with 60% detection efficiency at 1550 nm”. In: *Review of Scientific Instruments* 91.8 (2020), p. 083102.

[117] Alberto Tosi et al. “Low-Noise, Low-Jitter, High Detection Efficiency InGaAs/InP Single-Photon Avalanche Diode”. In: *IEEE Journal of Selected Topics in Quantum Electronics* 20.6 (Nov. 2014), pp. 192–197.

[118] Jun Zhang et al. “Advances in InGaAs/InP single-photon detector systems for quantum communication”. In: *Light: Science & Applications* 4.5 (May 2015), e286.

[119] Kiwon Lee and Kyounghoon Yang. “Analysis of InGaAs/InP Single-Photon Avalanche Diodes With the Multiplication Width Variation”. In: *IEEE Photonics Technology Letters* 26.10 (Mar. 2014), pp. 999–1002.

[120] Kiwon Lee. “Analysis of InP-based single photon avalanche diodes based on a single recess-etching process”. In: *Japanese Journal of Applied Physics* 57.4 (Mar. 2018), p. 042203.

[121] O.J. Pitts et al. “Planar avalanche photodiodes with edge breakdown suppression using a novel selective area growth based process”. In: *Journal of Crystal Growth* 470 (July 2017), pp. 149–153.

[122] Ekin Kizilkan et al. “Guard-Ring-Free InGaAs/InP Single-Photon Avalanche Diode Based on a Novel One-Step Zn-Diffusion Technique”. In: *IEEE Journal of Selected Topics in Quantum Electronics* 28.5 (Oct. 2022), pp. 1–9.

[123] Mark A. Itzler et al. “Advances in InGaAsP-based avalanche diode single photon detectors”. In: *Journal of Modern Optics* 58.3-4 (Jan. 2011), pp. 174–200.

[124] M Gibbon et al. “Selective-area low-pressure MOCVD of GaInAsP and related materials on planar InP substrates”. In: *IOP Publishing* 8.6 (June 1993), pp. 998–1010.

[125] O.J. Pitts et al. “Optimization of MOCVD-diffused p-InP for planar avalanche photodiodes”. In: *Journal of Crystal Growth* 393 (May 2014), pp. 85–88.

[126] Y. Okuto and C.R. Crowell. “Threshold energy effect on avalanche breakdown voltage in semiconductor junctions”. In: *Solid-State Electronics* 18.2 (Feb. 1975), pp. 161–168.

[127] R. K. Ahrenkiel et al. “Recombination lifetime of In0.53Ga0.47As as a function of doping density”. In: *Applied Physics Letters* 72.26 (June 1998), pp. 3470–3472.

[128] N. M. Shmidt et al. *Handbook Series on Semiconductor Parameters.* 1996.

[129] S. Forrest, R. Smith, and O. Kim. “Performance of In0.53Ga0.47As/InP avalanche photodiodes”. In: *IEEE Journal of Quantum Electronics* 18.12 (Dec. 1982), pp. 2040–2048.

[130] M. Sugiyama et al. “Control of abnormal edge growth in selective area MOVPE of InP”. In: *Journal of Crystal Growth* 287.2 (Jan. 2006), pp. 668–672.

[131] Chong Hu et al. “Characterization of an InGaAs/InP-based single-photon avalanche diode with gated-passive quenching with active reset circuit”. In: *Journal of Modern Optics* 58.3-4 (Sept. 2011), pp. 201–209.

[132] Yi Shan Lee et al. “Characteristics of Charge Persistence in InGaAs/InP Single-Photon Avalanche Diode”. In: *IEEE Photonics Technology Letters* 30.22 (Oct. 2018), pp. 1980–1982.

[133] Niccolò Calandri et al. “Charge Persistence in InGaAs/InP Single-Photon Avalanche Diodes”. In: *IEEE Journal of Quantum Electronics* 52.3 (Feb. 2016), pp. 1–7.

[134] Y. Okuto and C. R. Crowell. “Ionization coefficients in semiconductors: A nonlocalized property”. In: *Phys. Rev. B* 10.10 (Nov. 1974), pp. 4284–4296.

[135] Fabio Acerbi et al. “Fast Active Quenching Circuit for Reducing Avalanche Charge and Afterpulsing in InGaAs/InP Single-Photon Avalanche Diode”. In: *IEEE Journal of Quantum Electronics* 49.7 (Apr. 2013), pp. 563–569.

[136] O Salehzadeh et al. “Edge breakdown suppression of avalanche photodiodes using Zn diffusion and selective area growth”. In: *IEEE Photonics Technology Letters* 31.10 (2019), pp. 767–770.

[137] Niccolò Calandri et al. “Optical crosstalk in InGaAs/InP SPAD array: Analysis and reduction with FIB-etched trenches”. In: *IEEE Photonics Technology Letters* 28.16 (2016), pp. 1767–1770.

[138] Mark A Itzler et al. “Geiger-mode avalanche photodiode focal plane arrays for three-dimensional imaging LADAR”. In: *Infrared Remote Sensing and Instrumentation XVIII.* Vol. 7808. SPIE. 2010, pp. 75–88.

[139] Tihomir Knežević et al. “Examination of the InP/InGaAs single-photon avalanche diodes by establishing a new TCAD-based simulation environment”. In: *2016 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*. Sept. 2016, pp. 57–60.

[140] Fabio Telesca et al. “Temperature-dependent photon detection efficiency model for InGaAs/InP SPADs”. In: *Opt. Express* 30.3 (2022), pp. 4504–4514.

[141] E Kizilkan et al. “Extended Temperature Modeling of InGaAs/InP SPADs”. In: *ESSDERC 2023-IEEE 53rd European Solid-State Device Research Conference (ESSDERC)*. IEEE. 2023, pp. 140–143.

[142] Wolfgang Fichtner, Donald J Rose, and Randolph E Bank. “Semiconductor device simulation”. In: *SIAM Journal on Scientific and Statistical Computing* 4.3 (1983), pp. 391–415.

[143] Randolph E Bank, Donald J Rose, and Wolfgang Fichtner. “Numerical methods for semiconductor device simulation”. In: *SIAM Journal on Scientific and Statistical Computing* 4.3 (1983), pp. 416–435.

[144] Kazushige Horio and Hisayoshi Yanai. “Numerical modeling of heterojunctions including the thermionic emission mechanism at the heterojunction interface”. In: *IEEE transactions on electron devices* 37.4 (1990), pp. 1093–1098.

[145] Robert F Pierret and Gerold W Neudeck. *Advanced semiconductor fundamentals*. Vol. 6. Addison-Wesley Reading, MA, 1987.

[146] E. D. Palik. *Handbook of Optical Constants of Solids*. Vol. 1. Elsevier, 2012, pp. 503–517.

[147] Sadao Adachi. “Optical properties of $\mathrm{In_{1-x}Ga_xAs_yP_{1-y}}$ alloys”. In: *Phys. Rev. B* 39 (17 June 1989), pp. 12612–12621. DOI: 10.1103/PhysRevB.39.12612.

[148] W. J. Turner et al. “Exciton Absorption and Emission in InP”. In: *Phys. Rev.* 136 (5A Nov. 1964), A1467–A1470.

[149] E. Zielinski et al. “Excitonic transitions and exciton damping processes in InGaAs/InP”. In: *Journal of Applied Physics* 59.6 (1986), pp. 2196–2204.

[150] L. Pancheri, D. Stoppa, and G. Dalla Betta. “Characterization and Modeling of Breakdown Probability in Sub-Micrometer CMOS SPADs”. In: 20.6 (2014), pp. 328–335.

[151] A. Spinelli and A. L. Lacaita. “Physics and numerical simulation of single photon avalanche diodes”. In: 44.11 (Nov. 1997), pp. 1931–1943.

[152] Simon Verghese et al. “Arrays of InP-based Avalanche Photodiodes for Photon Counting”. In: *IEEE Journal of Selected Topics in Quantum Electronics* 13.4 (2007), pp. 870–886.

[153] Y Liu et al. “A planar InP/InGaAs avalanche photodiode with floating guard ring and double diffused junction”. In: *Journal of Lightwave Technology* 10.2 (1992), pp. 182–193.

[154] Fabio Acerbi, Alberto Tosi, and Franco Zappa. “Growths and diffusions for InGaAs/InP single-photon avalanche diodes”. In: *Sensors and Actuators A: Physical* 201 (2013), pp. 207–213.

[155] OJ Pitts et al. “Multiwafer zinc diffusion in an OMVPE reactor”. In: *Journal of crystal growth* 352.1 (2012), pp. 249–252.

[156] CA Hampel et al. “Metalorganic vapor phase diffusion using DMZn Part II: Determination of the interstitial zinc charge state from secondary ion mass spectroscopy measurements using the Boltzmann–Matano technique”. In: *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films* 22.3 (2004), pp. 916–920.

[157] GJ Van Gurp et al. “Zinc diffusion in n-type indium phosphide”. In: *Journal of applied physics* 61.5 (1987), pp. 1846–1855.

[158] D Franke, FW Reier, and N Grote. “Post-growth Zn diffusion into InGaAs/InP in a LP-MOVPE reactor”. In: *Journal of crystal growth* 195.1-4 (1998), pp. 112–116.

[159] Alberto Tosi et al. “Low-noise, low-jitter, high detection efficiency InGaAs/InP single-photon avalanche diode”. In: *IEEE Journal of selected topics in quantum electronics* 20.6 (2014), pp. 192–197.

[160] Jun Zhang et al. “Advances in InGaAs/InP single-photon detector systems for quantum communication”. In: *Light: Science & Applications* 4.5 (2015), e286–e286.

[161] Soo-Hyun Baek et al. “Room temperature quantum key distribution characteristics of low-noise InGaAs/InP single-photon avalanche diode”. In: *Journal of the Korean Physical Society* 78.7 (Apr. 2021), pp. 634–641.

[162] Pascal Rustige et al. “Guard-ring free InGaAs/InP single photon avalanche diodes for C-band quantum communication”. In: *2023 Optical Fiber Communications Conference and Exhibition (OFC)*. IEEE. 2023, pp. 1–3.

[163] Yusei Tamura et al. “Development of InGaAs MPPC for NIR photon counting applications”. In: *Optical Components and Materials XV*. Vol. 10528. SPIE. 2018, pp. 172–183.

[164] Francesco Gramuglia et al. “Engineering breakdown probability profile for PDP and DCR optimization in a SPAD fabricated in a standard 55 nm BCD process”. In: *IEEE Journal of Selected Topics in Quantum Electronics* 28.2: Optical Detectors (2021), pp. 1–10.

[165] Pouyan Keshavarzian et al. “Low-noise high-dynamic-range single-photon avalanche diodes with integrated PQAR circuit in a standard 55nm BCD process”. In: *Advanced Photon Counting Techniques XVI*. Vol. 12089. SPIE. 2022, pp. 73–82.

[166] M. Sanzaro et al. “Single-Photon Avalanche Diodes in a 0.16 μm BCD Technology With Sharp Timing Response and Red-Enhanced Sensitivity”. In: 24.2 (Mar. 2018), pp. 1–9.

[167] Justin A. Richardson et al. “Scaleable single-photon avalanche diode structures in nanometer CMOS technology”. In: 58.7 (July 2011), pp. 2028–2035.

[168] G. Zappalà et al. “Set-up and methods for SiPM Photo-Detection Efficiency measurements”. In: *Journal of Instrumentation* 11.08 (Aug. 2016), P08014–P08014.

[169] S. Lindner et al. “A High-PDE, Backside-Illuminated SPAD in 65/40-nm 3D IC CMOS Pixel With Cascoded Passive Quenching and Active Recharge”. In: 38.11 (Nov. 2017), pp. 1547–1550.

[170] Mingguo Liu et al. “High-Performance InGaAs/InP Single-Photon Avalanche Photodiode”. In: 13.4 (Sept. 2007), pp. 887–894.

[171] Eric A. G. Webster et al. “Per-Pixel Dark Current Spectroscopy Measurement and Analysis in CMOS Image Sensors”. In: 57.9 (Sept. 2010), pp. 2176–2182.

[172] W Maes, K De Meyer, and R Van Overstraeten. “Impact ionization in silicon: A review and update”. In: *Solid-State Electronics* 33.6 (1990), pp. 705–718.

[173] WN Grant. “Electron and hole ionization rates in epitaxial silicon at high electric fields”. In: *Solid-State Electronics* 16.10 (1973), pp. 1189–1203.

[174] Canali Jacoboni et al. “A review of some charge transport properties of silicon”. In: *Solid-State Electronics* 20.2 (1977), pp. 77–89.

[175] A Lacaita, A Spinelli, and S Longhi. “Avalanche transients in shallow p-n junctions biased above breakdown”. In: *Applied physics letters* 67.18 (1995), pp. 2627–2629.

[176] C. Niclass et al. “A Single Photon Avalanche Diode Implemented in 130-nm CMOS Technology”. In: 13.4 (July 2007), pp. 863–869.

[177] M. Gersbach et al. “A single photon detector implemented in a 130nm CMOS imaging process”. In: *Proc. 38th Eur. Solid-State Device Res. Conf.* Nov. 2008, pp. 270–273.

[178] Justin A. Richardson, Lindsay A. Grant, and Robert K. Henderson. “Low Dark Count Single-Photon Avalanche Diode Structure Compatible With Standard Nanometer Scale CMOS Technology”. In: 21.14 (July 2009), pp. 1020–1022.

[179] Eric A. G. Webster et al. “A Single-Photon Avalanche Diode in 90-nm CMOS Imaging Technology With 44% Photon Detection Efficiency at 690 nm”. In: 33.5 (May 2012), pp. 694–696.

[180] T. Leitner et al. “Measurements and Simulations of Low Dark Count Rate Single Photon Avalanche Diode Device in a Low Voltage 180-nm CMOS Image Sensor Technology”. In: 60.6 (June 2013), pp. 1982–1988.

[181] E. Charbon, H. Yoon, and Y. Maruyama. “A Geiger mode APD fabricated in standard 65nm CMOS technology”. In: *Proc. 2013 IEEE Int. Electron Devices Meeting.* 2013, pp. 27.5.1–27.5.4.

[182] Federica Villa et al. “CMOS SPADs with up to 500 μm diameter and 55% detection efficiency at 420 nm”. In: *J. Mod. Opt.* 61.2 (June 2014), pp. 102–115.

[183] C. Veerappan and E. Charbon. “A Substrate Isolated CMOS SPAD Enabling Wide Spectral Response and Low Electrical Crosstalk”. In: 20.6 (Nov. 2014), pp. 299–305.

[184] Myung-Jae Lee, Pengfei Sun, and Edoardo Charbon. “A first single-photon avalanche diode fabricated in standard SOI CMOS technology with a full characterization of the device”. In: *Opt. Express* 23.10 (May 2015), pp. 13200–13209.

[185] C. Veerappan and E. Charbon. “CMOS SPAD Based on Photo-Carrier Diffusion Achieving PDP >40% From 440 to 580 nm at 4 V Excess Bias”. In: 27.23 (Dec. 2015), pp. 2445–2448.

[186] Isamu Takai et al. “Single-Photon Avalanche Diode with Enhanced NIR-Sensitivity for Automotive LIDAR Systems”. In: *Sensors* 16.4 (Mar. 2016).

[187] S. Pellegrini and B. Rae. “Fully industrialised single photon avalanche diodes”. In: *Proc. Adv. Photon Count. Techniq. XI*. Vol. 10212. 2017, pp. 16.5.1–16.5.4.

[188] Hesong Xu et al. “Design and characterization of a p+/n-well SPAD array in 150nm CMOS process”. In: *Opt. Express* 25.11 (May 2017), pp. 12765–12778.

[189] S. Pellegrini et al. “Industrialised SPAD in 40 nm technology”. In: *Proc. 2017 IEEE Int. Electron Devices Meeting*. 2017, pp. 16.5.1–16.5.4.

[190] Manuel Moreno-García et al. “Low-noise single photon avalanche diodes in a 110nm CIS technology”. In: *2018 48th European Solid-State Device Research Conference (ESSDERC)*. IEEE. 2018, pp. 94–97.

[191] Ion Vornicu et al. “Low-noise and high-efficiency near-IR SPADs in 110nm CIS technology”. In: *ESSDERC 2019-49th European Solid-State Device Research Conference (ESSDERC)*. IEEE. 2019, pp. 250–253.

[192] Ion Vornicu et al. “Design of high-efficiency SPADs for LiDAR applications in 110nm CIS technology”. In: *IEEE Sensors Journal* 21.4 (2020), pp. 4776–4785.

[193] Fabio Telesca, Fabio Signorelli, and Alberto Tosi. “Double zinc diffusion optimization for charge persistence reduction in InGaAs/InP SPADs”. In: *IEEE Journal of Selected Topics in Quantum Electronics* (2023).

Ekin **Kızılkan**

Curriculum Vitae

Education

2019 -	**Ph.D. in Photonics (EDPO)**, *AQUA/EPFL*, Lausanne, Switzerland.
2017 - 2019	**MSc. Microengineering**, *EPFL*, Lausanne, Switzerland, *GPA: 5.17/6.0*.
2013 - 2017	**B.S in Physics (Double Major)**, *METU Middle East Technical University*, Ankara, Turkey, *GPA: 3.60/4.0 (High Honors Degree)*.
2011 - 2016	**BS. Electrical and Electronics Engineering (Major)**, *METU Middle East Technical University*, Ankara, Turkey, *GPA: 3.65/4.0 (High Honors Degree)*.
2007 - 2011	**Highschool**, *Antalya Anatolian High School*, Antalya, Turkey.

Research Experience

2019 - 2022	**AQUA lab**, *EPFL, Switzerland*. Development of infrared efficient single-photon-avalanche diodes based on narrow bandgap semiconductors (III-V, Germanium) for Lidar/3-D vision applications.
2017 - 2019	**AQUA lab**, *EPFL, Switzerland*. Design and characterization of III-V compounds semiconductors based single photon avalanche detectors.
2016 - 2017	**Quantum Devices & Nanophotonics Research Laboratory**, *METU, Turkey*. Design, fabrication and characterization of III-V compounds semiconductors based focal plane arrays for short wave infrared imaging.

Publications

- **Kizilkan, E.***, Karaca, U.*, Pesic, V., ... Charbon, E. (2022). Guard-Ring-Free InGaAs/InP Single-Photon Avalanche Diode Based on a Novel Zn-diffusion Technique. IEEE Journal of Selected Topics in Quantum Electronics.
- **Kizilkan, E.***, Karaca, U.*, Pesic, V.*, Lee, M. J., Bruschini, C., SpringThorpe, A. J., ... & Charbon, E. (2023, September). Extended Temperature Modeling of InGaAs/InP SPADs. In ESSDERC 2023-IEEE 53rd European Solid-State Device Research Conference (ESSDERC) (pp. 140-143). IEEE.
- Gramuglia, F.*, Keshavarzian, P.*, **Kizilkan, E.***, Bruschini, C., Tan, S. S., Tng, M., ... & Charbon, E. (2021). Engineering Breakdown Probability Profile for PDP and DCR Optimization in a SPAD Fabricated in a Standard 55 nm BCD Process. IEEE Journal of Selected Topics in Quantum Electronics, 28(2), 1-10.
- Karaca, U.*, **Kizilkan, E.***, Bruschini, C., & Charbon, E. (2023). A new double multiplication region method to design high sensitivity and wide spectrum SPADs in standard CMOS technologies. Scientific Reports. (Submitted)
- Lee, M. J., Karaca, U., **Kizilkan, E.**, Bruschini, C., & Charbon, E. (2023). A 73% Peak PDP Single-Photon Avalanche Diode Implemented in 110 nm CIS Technology with Doping Compensation. IEEE Journal of Selected Topics in Quantum Electronics.
- Ha, W. Y., Park, E., Eom, D., Park, H. S., Gramuglia, F., Keshavarzian, P., **Kizilkan, E.**, ... Lee, M. J. (2023). SPAD Developed in 55 nm Bipolar-CMOS-DMOS Technology Achieving Near 90% Peak

PDP. IEEE Journal of Selected Topics in Quantum Electronics.

- Keshavarzian, P., Gramuglia, F., **Kizilkan, E.**, Bruschini, C., Tan, S. S., Tng, M., ... Charbon, E. (2022, May). Low-noise high-dynamic-range single-photon avalanche diodes with integrated PQAR circuit in a standard 55nm BCD process. In Advanced Photon Counting Techniques XVI (Vol. 12089, pp. 73-82). SPIE.
- Wu, M. L., Ripiccini, E., **Kizilkan, E.**, Gramuglia, F., Keshavarzian, P., Fenoglio, C. A., ... Charbon, E. (2022). Radiation Hardness Study of Single-Photon Avalanche Diode for Space and High Energy Physics Applications. Sensors, 22(8), 2919.
- Wu, M. L., Ripiccini, E., Benserhir, J., Gramuglia, F., **Kizilkan, E.**, Bruschini, C., Charbon, E. (2023, July). Single-Photon Avalanche Diode for Scalable Particle Detection. In IEEE EUROCON 2023-20th International Conference on Smart Technologies (pp. 123-127). IEEE.
- Gramuglia, F., Keshavarzian, P., **Kizilkan, E.**, Wu, M. L., Bruschini, C., Tan, S. S., ... Charbon, E. (2022). A 7.5-ps, 60% PDP low-noise SPAD fabricated in CMOS technology. In ISSW 2022 (International SPAD Sensor Workshop).
- Antolovic, I. M., Ulku, A. C., **Kizilkan, E.**, Lindner, S., Zanella, F., Ferrini, R., ... & Bruschini, C. (2019, February). Optical-stack optimization for improved SPAD photon detection efficiency. In Quantum Sensing and Nano Electronics and Photonics XVI (Vol. 10926, p. 109262T). International Society for Optics and Photonics.
- Uzgur, F., Karaca, U., **Kizilkan, E.**, & Kocaman, S. (2017, May). Al/Sb free InGaAs unipolar barrier infrared detectors. In Infrared Technology and Applications XLIII (Vol. 10177, p. 1017706). International Society for Optics and Photonics.
- Uzgur, F., Karaca, U., **Kizilkan, E.**, & Kocaman, S. (2018). All InGaAs unipolar barrier infrared detectors. IEEE Transactions on Electron Devices, 65(4), 1397-1403.
- Uzgur, F., Karaca, U., **Kizilkan, E.**, & Kocaman, S. (9-13.04.2017). Al/Sb Free InGaAs Unipolar Barrier Infrared Detectors. SPIE. Defense and Commercial Sensing: Infrared Technology and Applications XLIII, Anaheim/Los Angeles.

*First authors with equal contribution

Work Experience

2018 -	**Intern**, *Fastree3D*, Lausanne,Switzerland.
Summary	InGaAs/InP type single-photon-avalanche-diode design.

Achievements & Awards

2021-2022	School of Engineering Teaching Assistant Award Nomination
2017-2018	Swissphotonics scholarship
2014-2017	TUBİTAK (The Scientific and Technological Research Council of Turkey) 2205: Domestic Bachelor Scholarship Programme

Computer skills

Centos, Matlab, Cadence, Sentaurus TCAD, SolidWorks, Python, LabVIEW, Verilog, VHDL

Laboratory Instruments

- Clean room experience: Metal deposition, lift-off, photolithography, wet dry etching.
- Characterization experience: Spectral efficiency measurements, Scanning Electron Microscope, XRD measurements, timing jitter, noise, breakdown voltage, dark current measurements.

Certification

PRINCE2 Foundation certificate

Languages

Turkish	**Native**	
English	**Advanced: C1**	*Fluent in writing and speaking*
French	**A2**	

Interests

- Chess
- SUP
- Board Games
- Ski
- Debate
- Hike