Dissertation

submitted to the

Combined Faculty of Natural Sciences, Engineering and Mathematics

of Heidelberg University, Germany

for the degree of

Doctor of Natural Sciences

Put forward by

Sonam Sharma

born in: Banga, India

Oral examination: 03-07-2025

Design and Development of Single Photon Avalanche Diode in Silicon-On-Insulator Technology

Referees: Prof. Dr. Hans-Christian Schultz-Coulon

Prof. Dr. Peter Fischer

Abstract

This thesis investigates the design, development, and characterization of a novel single photon avalanche diode (SPAD), fabricated in silicon-on-insulator (SOI) wafer technology, for blue and near-UV (NUV) light sensitivity. SPAD sensors are semiconductor photodetectors known for their ability to detect single photons with low noise and high detection efficiency.

Initial TCAD Synopsys simulations demonstrate the design's feasibility and allow for specifying dimensions and process parameters for the sensor. The key parameters, including doping concentrations, junction depths, and electric field distribution, are optimized. The results of these simulations indicated a breakdown voltage of 15.25 V and a photon detection probability of 44% at 405 nm of wavelength.

With the production of SPAD, a proof-of-concept is possible by measuring a working quench mechanism, although the observed breakdown voltage is much higher than expected. Discrepancies between the simulated and experimental results are further investigated with additional tests, such as a light emission test and secondary ion mass spectroscopy, pointing to a mismatch in doping concentrations.

This study proposes a revised design with adjusted doping concentrations, which has been simulated to evaluate its impact on performance, especially concerning the p-n junction. These refinements are expected to align the experimental results more closely with the simulations, enhancing the SPAD design for better photon detection in the UV and NUV wavelength range.

Zusammenfassung

Diese Arbeit untersucht das Design, die Entwicklung und die Charakterisierung einer neuartigen Einzelphotonen-Avalanche-Diode (SPAD), die in Silizium-auf-Isolator (SOI)-Wafer-Technologie hergestellt wird und für blaues und Nah-UV-Licht (NUV) empfindlich ist. SPAD-Sensoren sind Halbleiter-Photodetektoren, die für ihre Fähigkeit bekannt sind, einzelne Photonen mit geringem Rauschen und hoher Effizienz zu erkennen.

Erste TCAD-Synopsys-Simulationen zeigen die Machbarkeit des Entwurfs und ermöglichen die Festlegung der Abmessungen und Prozessparameter für den Sensor. Die Schlüsselparameter, einschließlich der Dotierungskonzentrationen, Übergangstiefen und der Verteilung des elektrischen Feldes, werden optimiert. Die Ergebnisse dieser Simulationen ergaben eine Durchbruchspannung von 15,25 V und eine Photonendetektionswahrscheinlichkeit von 44% bei einer Wellenlänge von 405 nm.

Mit der Herstellung von SPAD ist ein Konzeptnachweis möglich, indem ein funktionierender Quench-Mechanismus gemessen wird, obwohl die beobachtete Durchbruchspannung viel höher als erwartet ist. Die Diskrepanzen zwischen den simulierten und den experimentellen Ergebnissen werden mit zusätzlichen Tests, wie z. B. einem Lichtemissionstest und Sekundärionen-Massenspektroskopie, weiter untersucht, was auf ein Missverhältnis der Dotierungskonzentrationen hindeutet.

In dieser Studie wird ein überarbeitetes Design mit angepassten Dotierungskonzentrationen vorgeschlagen, das simuliert wurde, um seine Auswirkungen auf die Leistung zu bewerten, insbesondere im Hinblick auf den p-n-Übergang. Es wird erwartet, dass diese Verfeinerungen die experimentellen Ergebnisse besser mit den Simulationen in Einklang bringen und das SPAD-Design für eine bessere Photonenerkennung im UV- und NUV-Wellenlängenbereich verbessern.

Contents

Chapter 1

Motivation

Photodetectors, which convert light into electrical signals, have significantly contributed to medical imaging and scientific research advances. The evolution of these devices began with photomultiplier tubes (PMTs) in the mid-20th century, known for their ability to detect individual photons. However, PMTs were large and complex, making them unsuitable for many modern applications that required smaller solid-state devices [1].

The rise of semiconductor technology led to the development of photodiodes and eventually avalanche photodiodes (APDs) in the 1960s [2], which could amplify light signals through an avalanche multiplication process within a reverse-bias p-n junction. The need for even higher sensitivity, particularly in the areas of single-photon detection, led to the invention of Single-Photon Avalanche Diodes (SPADs) later. The ability of SPADs to operate in Geiger mode, where the detection of a single photon triggers a self-sustaining avalanche, marked a breakthrough in photon detection technology [3].

Despite their advancements, conventional SPADs still face challenges, especially in applications that require detection in the blue and near-ultraviolet (NUV) spectral ranges, where conventional SPADs often perform poorly. This thesis aims to develop a new SPAD sensor using Silicon-on-Insulator (SOI) technology and enabling the integration of these detectors through a 3D integration approach.

The research focuses on optimizing the SPAD's design to enhance performance in a specific wavelength range, initially using TCAD simulation tools. Following the simulations, experimental verifications are conducted to test the actual performance of the fabricated SPADs in comparison to the outcomes anticipated from the simulations. This step is crucial for understanding the real behavior of the sensors and for identifying any discrepancies between the expected and measured outcomes. Based on the findings, a new approach is formulated, which involves revision of design parameters and the conduct of further simulations to align with the expected SPAD performance.

1.1 Thesis outline

The thesis is organized as follows:

Chapter 2 - This chapter provides an overview of the fundamental device physics underlying SPADs, including light-matter interaction and photon detection principles. The motivation towards the 3D stacked SPAD integration with CMOS is explained by the doping dependency on the sensor fabrication, crucial for understanding its sensitivity in the blue and NUV range of wavelengths.

Chapter 3 - The third chapter details the use of Technology Computer-Aided Design (TCAD) simulation tools to model the SPAD. It explains the simulation processes used to validate the sensor design before fabrication, highlighting how simulations are calibrated and how they guide the device design before fabrication.

Chapter 4 - This chapter presents the comprehensive results from the TCAD simulation. The SPAD device geometry is elaborated, including a detailed study of doping profiles and the results of optical simulation that predict photon detection probabilities.

Chapter 5 - Experimental results are presented in this chapter, where the fabricated SPADs are tested under various conditions to verify the simulation predictions, including discrepancies between simulated predictions and actual measurements. Techniques such as light emission test and secondary ion mass spectroscopy are used to diagnose the issues.

Chapter 6 - Based on the insights gained from both the simulations and experimental verifications, this chapter discusses the adjustments made to the initial design. It also outlines the future direction for research, focusing on further enhancements of the SPAD's efficiency and reliability in photon detection.

Chapter 7 - This chapter summarizes the research.

Author's contributions

The project was initiated by Dr. Wei Shen, a project leader of the detector development group, who developed the original idea of a 3D-integrated SPAD and designed the first sensor.

The author contributed significantly to the technical development and experimental validation throughout the project.

The TCAD simulations were carried out by the author to understand the sensor design and validate its key parameters, including the detailed simulation of breakdown triggering probability and the modeling of photon detection probability.

The SPAD sensors were fabricated by Fraunhofer EMFT, München. The author designed the PCB readout boards used for sensor testing and assembled a few of the SPAD sensors by bonding them onto the boards. All the laboratory measurements of the fabricated

sensors were performed by the author. An automated code for wafer-level DC characterization was developed by the author to enable large-scale measurement of sensors across the wafers.

Based on the simulation and laboratory findings, the author proposed a revised set of ion implantation parameters and conducted new TCAD simulations for the second production cycle.

Chapter 2

Fundamentals of physics and vision towards 3D integrated sensors

The conceptual design of photon detection requires a deep understanding of the interaction between light and matter. This chapter aims to provide a concise overview of the theoretical background necessary to comprehend the physical processes and phenomena related to this interaction, essential to understand the detection mechanisms in photon detectors such as Single Photon Avalanche Diodes (SPADs). Moreover, this chapter discusses advancements in 3D integrated sensors with different light illumination strategies along with a focus on design considerations for blue light sensitivity.

2.1 Photon detection with semiconductors

In classical theory, light exhibits both a wave-like and a particle-like behavior depending on the context of the interaction. However, with quantum theory, light exhibits particle-like properties in which it behaves as a discrete entity called photons, depending on frequency f and wavelength λ [4]. The energy of a photon can be calculated as

$$E = h \cdot f = \frac{h \cdot c}{\lambda} \tag{2.1}$$

In semiconductor detectors, the detection of photons is based on the interaction between photons and matter, which is crucial for defining the properties of detectors designed for specific applications. These interactions allow for the conversion of photon energy into measurable electronic signals.

Photon interactions with matter are characterized by the nature of the target, such as electrons, atoms, or nuclei, and by the type of event, such as scattering, absorption, or pair production. These interactions result in a partial or full transfer of the photon energy to the material [5].

2.1.1 Light-matter interaction

Depending on the energy of the incident photons and the properties of the medium, the photons interact through three primary mechanisms: the photoelectric effect, compton scattering, and pair production, as shown in Figure 2.1. These mechanisms are fundamental in explaining how photons can transfer their energy to electrons, which then results in a measurable signal in a photon detector. In comparison, the photoelectric effect is significant only at low photon energies (below $\sim$ 0.3 MeV) and for materials with high atomic numbers. Pair production becomes important at higher photon energies and high atomic numbers. For intermediate photon energies, compton scattering remains the most prevalent interaction across all atomic numbers [6].

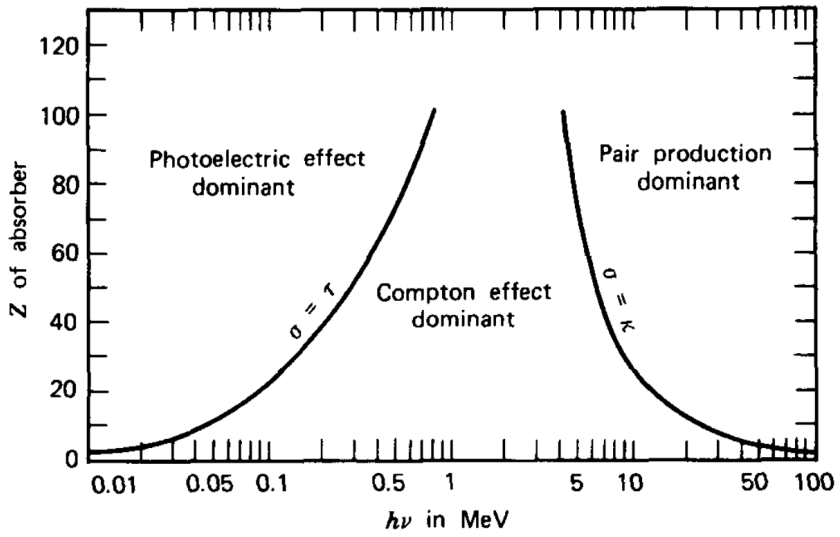


Figure 2.1. Three major types of light-matter interaction as a function of the atomic number and the photon energy. [6]

Photoelectric effect

The photoelectric effect occurs when a photon transfers its entire energy to a bound electron within an atom, causing the electron to be ejected as a photoelectron. The ejected electron typically originates from the atom's K or L shells. The energy of the ejected electron E_e can be expressed as [7]:

$$E_e = h\nu - E_b \tag{2.2}$$

Here, hν represents the energy of the photon and E_b is the binding energy of the ejected electron. For this interaction to happen, the photon's energy must be greater than the electron's binding energy, which makes this an inelastic process.

The phenomenon is divided into two categories based on the electron's emission location: the outer and inner photoelectric effects. The outer photoelectric effect occurs when electrons are ejected from a material's surface into a vacuum, commonly seen in metals. The inner photoelectric effect, on the other hand, involves electrons excited from lower to higher energy levels within the material, such as from the valence band to the conduction band in semiconductors.

The kinetic energy of these photoelectrons, which directly correlates with the incident photons' energy, enables semiconductor detectors to not only detect photons but also to accurately measure their energy. The inner photoelectric effect enables photon detection in devices like SPADs.

Compton scattering

Compton scattering involves a photon colliding with a loosely bound or free electron, transferring some of its energy to the electron, and scattering in a new direction. The energies of the scattered electron (E'_e) and the scattered photon (E'_γ) can be derived using the conservation of energy and momentum law:

$$E'_e = h(\nu - \nu') = h\nu \left[1 - \frac{1}{1 + \frac{h\nu}{m_e c^2}(1 - \cos(\theta))}\right] \tag{2.3}$$

and

$$E'_\gamma = h\nu' = \frac{h\nu}{1 + \frac{h\nu}{m_e c^2}(1 - \cos(\theta))} \tag{2.4}$$

Here, c represents the speed of light in a vacuum, θ is the scattering angle, m_e is the electron mass, and ν is the initial photon frequency.

Compton scattering depends on the number of electrons available as scattering partners and therefore scales linearly with Z [5].

Pair production

Pair production is a phenomenon where a photon is converted into an electron and its antiparticle, a positron. This occurs when the energy of the photon is at least twice the rest energy of an electron, specifically exceeding 1.02 MeV:

$$\gamma = e^- + e^+ \tag{2.5}$$

with

$$E_\gamma > 2m_e c^2 \approx 1.02\,\text{MeV} \tag{2.6}$$

where $m_e c$ denotes the rest mass energy of an electron. Due to the conservation of momentum, this process is only possible in the presence of a nucleus. The probability of

pair production increases with the photon's energy and is approximately proportional to the atomic number (Z^2) [5].

2.1.2 Semiconductor properties

In order to understand the working principle of semiconductor detectors, this section provides a brief overview of the main properties of semiconductors.

Energy band structure

In a semiconductor's crystal lattice, the relationship between energy and momentum is crucial, as both are conserved during the interaction between charge carriers and photons or phonons. This interaction leads to the concept of an energy gap in the material. The electronic band structure of crystalline solids is defined by solving the Schrödinger equation for a periodic potential for which the solution can be expressed as

$$\psi_k(r) = u_k(r)e^{-ikr} \tag{2.7}$$

where e^{-ikr} represents a plane wave and $u_k(r)$ is a function that reflects the periodic nature of the crystal lattice. A one-electron wavefunction of the form of Equation 2.7 is called a Bloch function, which describes how electrons move freely within a perfectly periodic crystal [8].

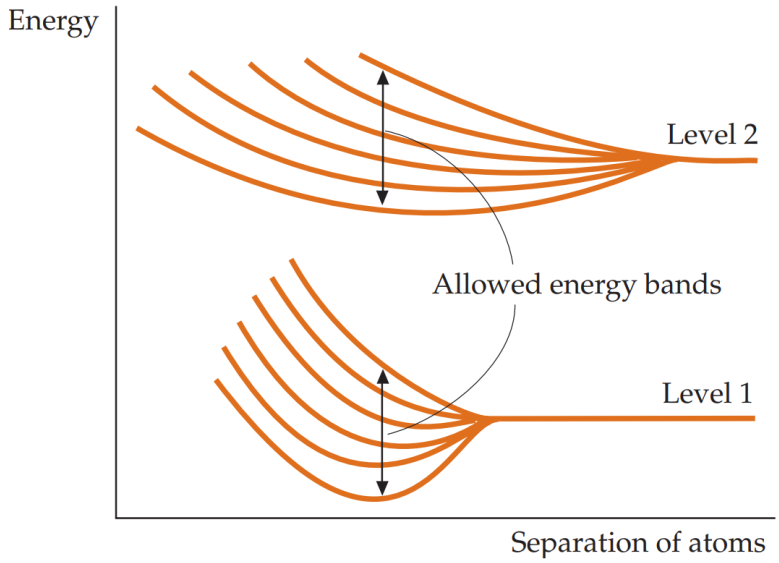
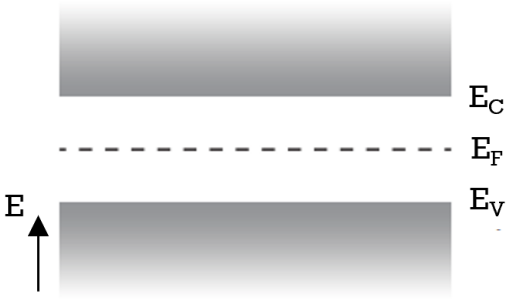


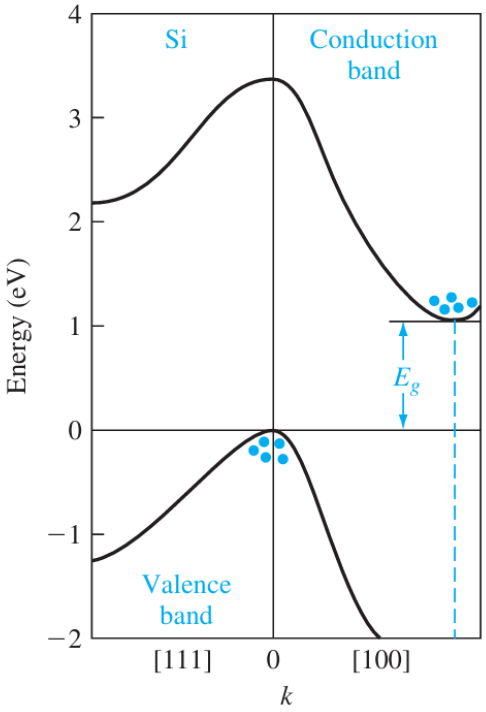
Figure 2.2. Energy splitting of two energy levels for six atoms as a function of the separation of the atoms [9].

To understand what a semiconductor is, the band model of solid state materials has to

be understood. The interaction of the lattice atoms in solid states transforms discrete energy levels of single atoms into energy band structures. Initially, each atom has discrete energy levels, and when atoms are distant, these levels are identical. However, as atoms approach one another, their energy levels start to influence each other, leading to a splitting into slightly different energy levels. This effect becomes more pronounced as more atoms come close, where each original energy level splits into a large number of levels called a band. The levels are spaced almost continuously within the band. There is a separate band of levels for each particular energy level of the isolated atom. The bands may be widely separated in energy, they may be close together, or they may even overlap, depending on the kind of atom and the type of bonding in the solid [10]. Figure 2.2 shows the energy splitting of two energy levels for six atoms as a function of the separation of the atoms.



(a) Energy bands showing valence and conduction levels with the Fermi energy.



(b) Energy band structure of Silicon.

Figure 2.3. Band structure of Silicon [10].

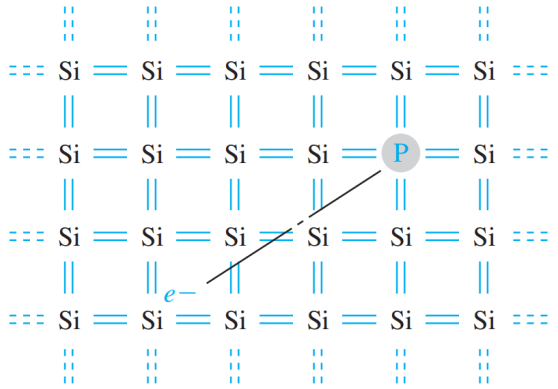
In the band structure depicted, electrons fill the energy levels up to the Fermi energy E_F in accordance with the Pauli Exclusion Principle. The highest band that is completely filled with electrons is known as the valence band, denoted by E_V, at its upper boundary, while the conduction band, which is either partially filled or completely unoccupied, starts at E_C, the energy level at the bottom of this band. A completely filled band does not contribute to carrier transport, since no free states are available. Materials with the Fermi level within an absolute band gap are semiconductors, depending on the size of the energy gap E_g. In

metals, the Fermi level is within the conduction band, which means this band is partially filled. Holes are introduced as imaginary equivalent particles to describe the interactions of the electrons within an almost completely filled band.

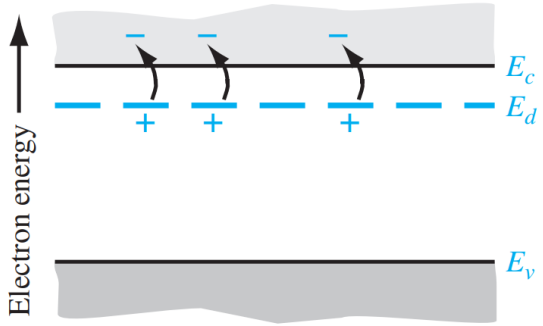
The E-k diagram provided for silicon shown in Figure 2.3 illustrates key aspects of its band structure [10]. In silicon, the highest energy point in the valence band occurs at k = 0, whereas the lowest energy point in the conduction band is not at k = 0, but along the [100] direction. This mismatch in the momentum space (k-space) between the valence band maximum and the conduction band minimum defines the energy gap E_g and classifies silicon as an indirect bandgap semiconductor. In indirect bandgap materials like silicon, electron transitions from the valence band to the conduction band require not only an energy change but also a momentum change. A transition in an indirect bandgap material must necessarily include an interaction with the crystal so that crystal momentum is conserved.

Doping

The ability to modify the electrical properties of semiconductors by introducing impurities into their crystal lattice is a fundamental aspect of semiconductor technology.



(a) Two-dimensional representation of the silicon lattice doped with a phosphorus atom.

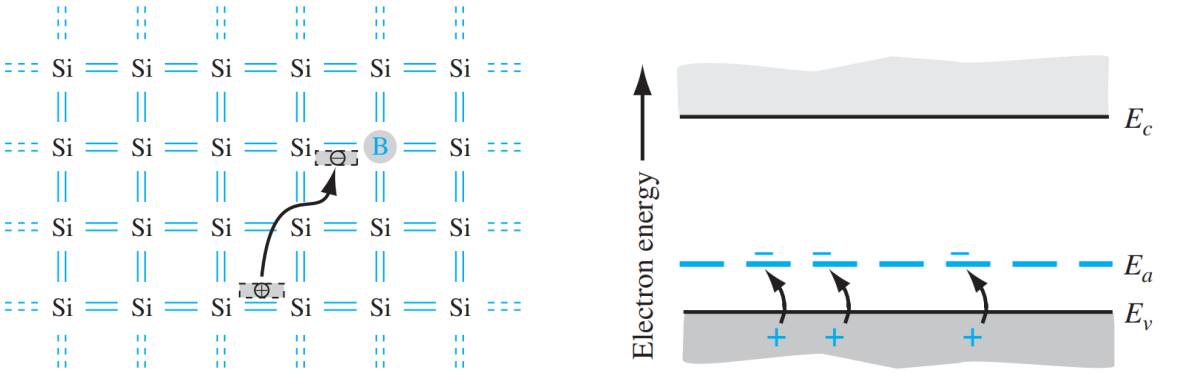


(b) The energy-band diagram showing the effect of a donor state being ionized.

Figure 2.4. Donor impurity level with a band diagram [10].

By doping a semiconductor such as silicon with a group V element like phosphorus, which possesses five valence electrons, four of these electrons contribute to the covalent bonds with silicon atoms. The fifth electron, however, remains loosely bound to the phosphorus atom at low temperatures, called a donor electron, as depicted in Figure 2.4(a). As the temperature increases, even a small amount of thermal energy can excite this donor electron to move into the conduction band, leaving behind a positively charged phosphorus ion, and the energy required to move the donor electron into the conduction band is significantly

less than that for an electron involved in the covalent bonding. The semiconductor thus becomes an n-type semiconductor by donating negatively charged electrons to the conduction band, enhancing its electrical conductivity without the creation of holes in the valence band (Figure 2.4(b)).



(a) Two-dimensional representation of a silicon lattice doped with a boron atom.

(b) The energy-band diagram showing the effect of an acceptor state being ionized.

Figure 2.5. Acceptor impurity level with a band diagram [10].

By replacing the Si atom with a group III element like boron, three valence electrons are taken up by the covalent bonding, making one covalent bonding position empty. The valence electron from the neighbouring Si atom occupies this position, leaving behind a vacancy, and this vacant electron position can be described as holes in the valence band, and impurities with these characteristics are called acceptors (Figure 2.5(a)). The acceptor atom can generate holes in the valence band without generating electrons in the conduction band (Figure 2.5(b)). The resulting material is referred to as a p-type semiconductor.

Figure 2.6 shows the ionization energies for different materials in silicon.

The type of semiconductor determines which type of charge carrier dominates the conductivity. In intrinsic semiconductors, where there are no dopants, both acceptors (N_A) and donors (N_D) are absent, resulting in equal densities of electrons (n) and holes (p), both equal to the intrinsic carrier density (n_i). From the mass-action law, the semiconductor under equilibrium conditions will follow:

$$n \cdot p = n_i^2 \tag{2.8}$$

In extrinsic semiconductors, if the concentration of electrons is larger than the holes, the current is carried predominantly by electrons, called the n-type semiconductor. On the other hand, if the concentration of holes is larger than electrons, the dominant conductivity is by holes, called the p-type semiconductor. The concentration of the corresponding majority carrier will then be given by $n = N_D - N_A$ if $N_D > N_A$ and by $p = N_A - N_D$ if $N_A > N_D$ [11].

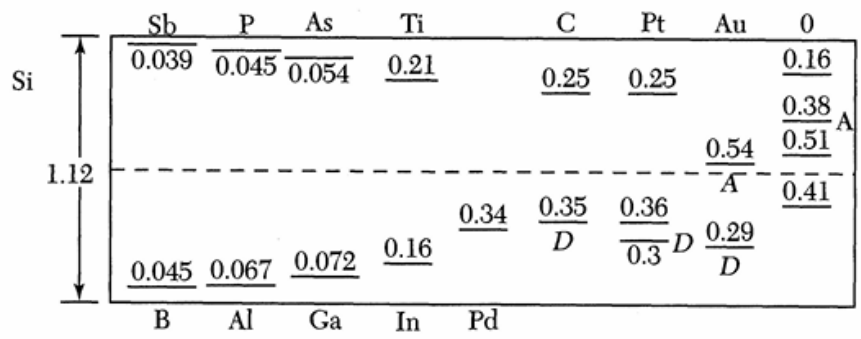


Figure 2.6. Measured ionization energies (in eV) for various impurities in Si. The levels below the gap center are measured from the top of the valence band and are acceptor levels unless indicated by D for the donor level. The levels above the gap center are measured from the bottom of the conduction band and are donor levels unless indicated by A for acceptor level [12].

Optical absorption

The photon absorption process is essential for photogeneration, where electron-hole pairs (EHPs) are created depending on the photon's energy being at least equal to the bandgap energy E_g of the semiconductor to excite an electron from the valence band to the conduction band. The threshold wavelength, or upper cut-off wavelength, λ_g, for this photogenerative absorption is determined by the bandgap energy E_g and can be calculated as

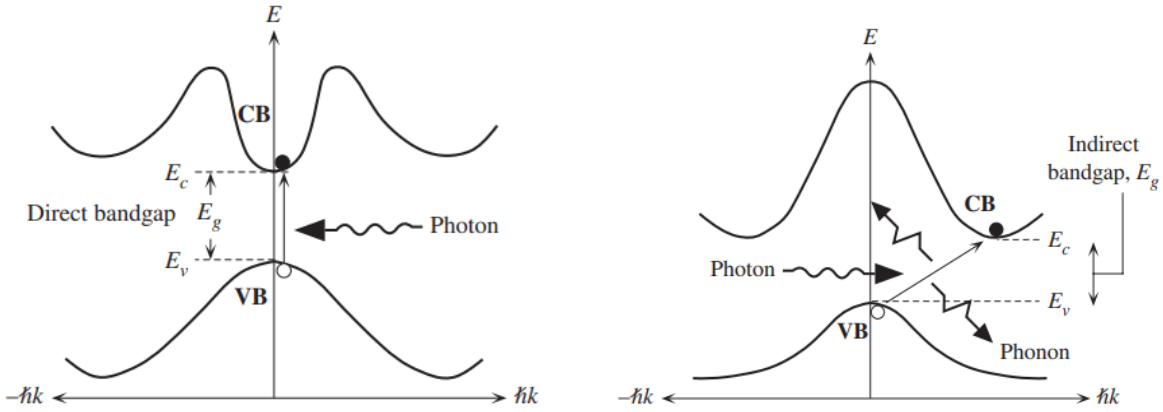
$$\lambda_g(\mu m) = \frac{1.24}{E_g(eV)} \tag{2.9}$$

Absorption in semiconductors needs an understanding of the relationship between electron energy E and electron momentum $\hbar k$ in the crystal, referred to as crystal momentum. E versus $\hbar k$ behavior for electrons in valence and conduction bands of direct and indirect bandgap semiconductors are shown in Figure 2.7.

In direct bandgap semiconductors, such as III-V semiconductors like GaAs and InP, photon absorption is a direct process which requires no assistance from lattice vibrations, where the photon is absorbed and excites an electron directly from the valence band to the conduction band without any change in the electron's wavevector k or crystal momentum $\hbar k$. The change in the electron momentum from the valence to the conduction band is

$$\hbar k_{\text{CB}} - \hbar k_{\text{VB}} = \text{Photon momentum} \approx 0 \tag{2.10}$$

This type of absorption corresponds to a vertical transition on the E versus $\hbar k$ shown in Figure 2.7(a).



(a) GaAs (direct bandgap). **(b)** Si (indirect bandgap).

Figure 2.7. Electron energy versus momentum and photon absorption for direct and indirect bandgap semiconductors [13].

In indirect bandgap semiconductors like silicon (Si) and germanium (Ge), the photon absorption process near the bandgap energy involves both the absorption and emission of lattice vibrations or phonons shown in Figure 2.7(b). If K is the wavevector of a lattice wave, then $\hbar K$ represents the phonon momentum. When an electron in the valence band excites to the conduction band, a change in its momentum in the crystal happens, which cannot be supplied by the momentum of an incident photon which is very small. Thus, the momentum difference must be balanced by a phonon momentum.

$$\hbar k_{\text{CB}} - \hbar k_{\text{VB}} = \text{Phonon momentum} = \hbar K \tag{2.11}$$

The absorption process is said to be indirect, as it depends on the lattice vibrations, which in turn depend on the temperature [13].

2.1.3 PN junction

Transitioning from the understanding of semiconductor properties and photon absorption characteristics, next is to consider how this is utilized within a fundamental semiconductor device, a p-n junction. When a p-type semiconductor comes into the contact of an n-type semiconductor, a p-n junction is formed with the creation of a gradient of two charge carrier concentrations at the boundary that results in the diffusion current, meaning the excess electrons diffuse from the n-region to the p-region and the excess holes diffuse from the p-region to the n-region, so the charge carriers try to reach the equilibrium state of $n = p = n_i$ (see Equation 2.8).

As electrons and holes migrate into the opposite regions, they result in the formation of a depletion zone or space charge region, which extends W_{Dn} on the n-side and W_{Dp}

on the p-side with no free carriers remaining. Due to the opposite space charge in the p- and n- layer, an intrinsic electric field is generated, leading to a drift current in the opposite direction to the diffusion current shown in Figure 2.8, and a dynamic equilibrium is reached, which stops the further flow and causes a constant potential difference between n- and p-type, called the built-in voltage V_{bi}. The electric field reaches its peak value at the junction and can be calculated as

$$E_{\text{peak}} = -\frac{q_0 \cdot N_D \cdot W_{Dn}}{\epsilon_r} = -\frac{q_0 \cdot N_A \cdot W_{Dp}}{\epsilon_r} \tag{2.12}$$

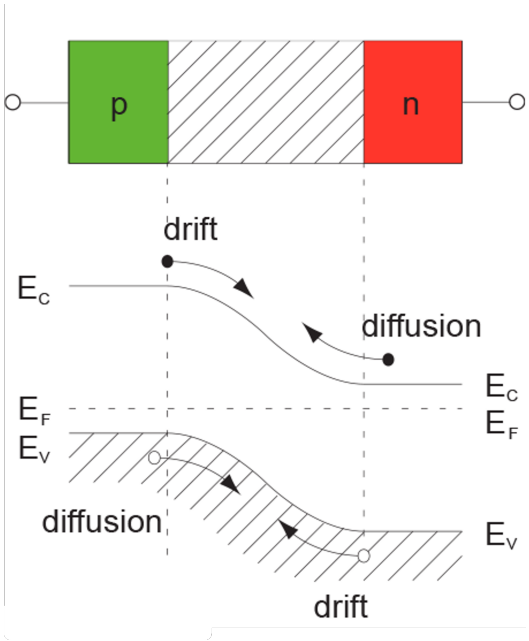


Figure 2.8. Drift and diffusion currents for pn boundary [14].

where q_0 is the elementary charge, ϵ_r is the relative permittivity of the solid, N_D and N_A are the donor and acceptor concentrations and W_{Dn}, W_{Dp} are the depletion widths in the n- and p-regions. The depletion width with a function of the built-in voltage as well as doping concentrations can be calculated as

$$W_D(V_{bi}) = W_{Dn}(V_{bi}) + W_{Dp}(V_{bi}) = \sqrt{\frac{2\epsilon_r}{q_0} \cdot \left(\frac{N_A + N_D}{N_A \cdot N_D}\right) \cdot V_{bi}} \tag{2.13}$$

When an external potential is applied across a p-n junction, the equilibrium condition within the junction is no longer present, and the current starts flowing depending on the bias conditions. Due to the low charge density, the whole voltage V drops over the space charge region:

$$\Delta V = V_{bi} - V \tag{2.14}$$

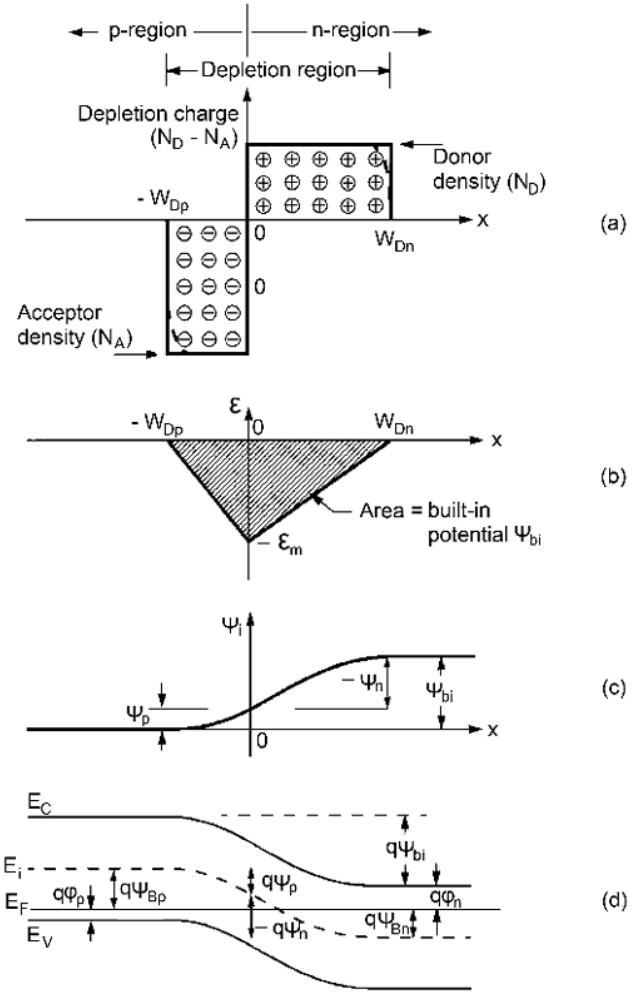


Figure 2.9. Abrupt p-n junction in thermal equilibrium. (a) Space-charge distribution. (b) Electric-field distribution. (c) Potential distribution. (d) Energy-band diagram [12].

The width of the depletion zone is then altered $\propto \sqrt{1 - V/V_{bi}}$, since V_{bi} in Eq. 2.13 is replaced by $V_{bi} - V$.

Figure 2.9 shows an abrupt junction of p- and n-type material. In forward bias, the p-side is at a positive potential with respect to the n-side, and $V > 0$, leading to a decrease in the band bending. It results in a decrease in the depletion width because the charge carriers are pushed towards the junction. The excess electrons from the p-region are pushed to the n-region and the excess holes from the n-region to the p-region, where they recombine. The charge carriers removed due to recombination are replenished by the contacts, resulting in a forward current flow. The contributions of these recombining charge carriers to the overall current can be categorized into diffusion current, occurring in the neutral zones of doping, and recombination current, taking place in the space charge region. With increased forward bias, the depletion zone narrows, thus reducing the recombination current until the diffusion current dominates.

Conversely, in reverse bias $V < 0$, the band deflection increases. The depletion width increases as it pulls free charge carriers away from the junction, resulting in a small current flow. The reverse current in diodes primarily arises from two contributions: generation and diffusion current. The generation current dominates in silicon p-n junctions and results from thermally generated electron-hole pairs that get separated by the applied voltage. Additionally, diffusion current occurs by the movement of electron-hole pairs formed outside the depletion region, moving from areas of higher to lower concentration due to thermal velocity. Electrons from the n-region and holes from the p-region are pushed back from reaching the edge, whereas holes from the n-region and electrons from the p-region are swept toward the opposing side by the electric field of the space-charge region. This directional movement results in diffusion current predominantly carried by minority charge carriers, while the majority carriers are reflected back into their respective zones. In both cases, the charge carriers will then be extracted via the electric contacts, which results in a current called leakage current [10].

Impact ionization and avalanche breakdown

As reverse bias across a p-n junction increases, the electric field within the depletion region becomes strong enough that an electron drifting in this region gains significant kinetic energy, which is enough to ionize a silicon atom or break a Si-Si bond, and this process of ionization is called impact ionization, as depicted in Figure 2.10. For an electron to effectively ionize an atom and create an electron-hole pair, it must gain energy at least equal to the bandgap energy E_g to excite an electron from the valence to the conduction band.

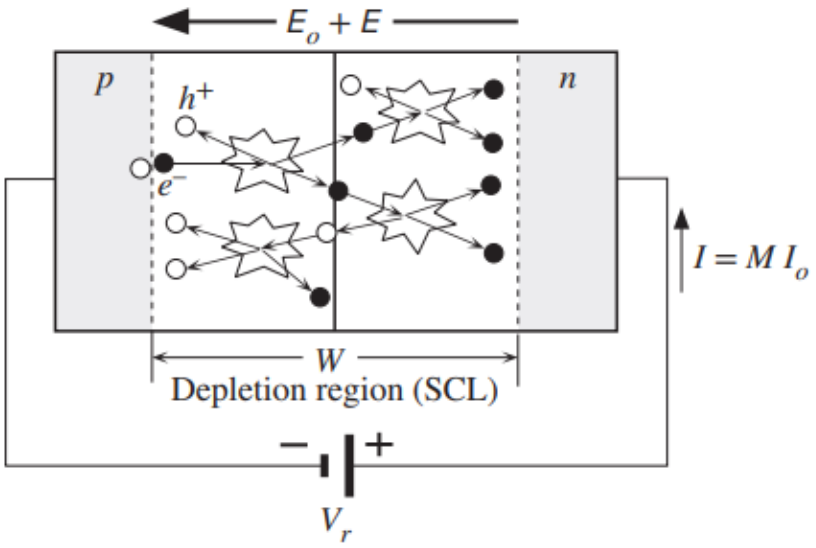


Figure 2.10. Avalanche breakdown by impact ionization [15].

When a thermally generated electron in the p-side inside the depletion region accelerates due to the field, it can collide with a silicon atom, lose energy equivalent to E_g, but accelerate for another ionizing collision, and potentially cause more ionizations along the depletion region. This leads to an avalanche effect where one initial carrier can trigger the generation of many carriers through an avalanche of impact ionization in the depletion region.

The presence of impact ionization significantly amplifies the reverse current in the junction. If the reverse current in the depletion region in the absence of impact ionization is I_0, then due to an avalanche effect, the reverse current becomes MI_0, where M is the multiplication factor, representing the net number of carriers generated by the avalanche effect per carrier in the depletion region. Impact ionization strongly depends on the electric field and a small rise in current can sharply increase the multiplication process. The relationship can be expressed as

$$M = \frac{1}{1 - \left(\frac{V_r}{V_{bd}}\right)^n} \tag{2.15}$$

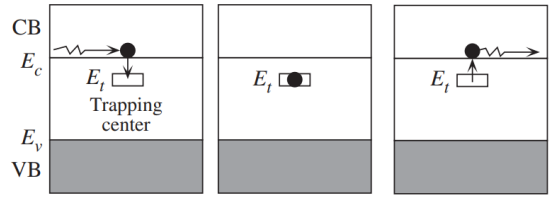
Here, V_r is the reverse bias voltage, V_{bd} is the breakdown voltage, and n typically ranges between 3 and 6, depending on the semiconductor and the type of substrate. As V_r approaches the breakdown voltage V_{bd}, avalanche breakdown occurs, which corresponds to $M \rightarrow \infty$. Once this condition is fulfilled, there is no need for external carriers to support the avalanche, resulting in a self-sustained avalanche and a consequent electric current flow.

This free carrier can be the result of a photogeneration process directly in the depletion zone, once a photon is absorbed in that zone or even a diffusion process from the neutral zones when the carrier is generated there. During avalanche breakdown, photon emission can occur due to hot carrier luminescence [16].

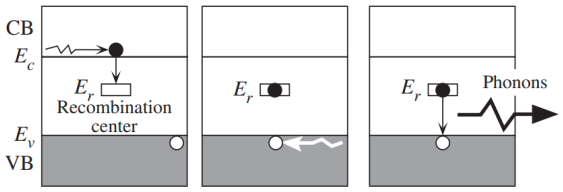
2.1.4 Generation/Recombination phenomenon

In an ideal semiconductor, the bandgap between the valence and conduction bands contains no allowed energy states. However, in reality, during doping or due to chemical impurities, defects such as vacancies or interstitials can introduce deep energy states within the bandgap. These defect states act as traps where electrons or holes can be temporarily captured. Unlike free carriers in the conduction or valence bands, carriers in these states are localized around the defect site. The electrons or holes can be temporarily or permanently removed from the conduction band or valence band and captured in these deep energy states. In the first case, the electrons or holes are temporarily removed from the conduction or valence band, falls into the trapping center E_t and later with the incident energetic lattice vibrations, are excited back into the conduction or valence band, shown in Figure 2.11(a). In the second case, the electrons from the conduction band can be captured by the recombination center E_r, remain

localized until a hole arrives and recombines with it by losing some energy in the form of lattice vibrations called phonons, as shown in Figure 2.11(b). Also, with the absorption of these lattice vibrations, these defects can allow the generation of electrons and holes. The defect states are near the mid-gap of bands. This non-radiatve recombination is the dominant recombination process in indirect bandgap materials such as Si and Ge, and is known as the Shockley-Read-Hall (SRH) process.



(a) Trapping and de-trapping of electrons by trapping centers. A trapping center has a localized energy level in the bandgap.



(b) Recombination in Si via a recombination center that has a localized energy level at E_r in the bandgap, usually near the middle.

Figure 2.11. Recombination and trapping process. [15]

The net transition rate for the mid-gap trap states for SRH is

$$U_{SRH} = \frac{n_c p_v - n_i^2}{\tau_p(n_c + n_i) + \tau_n(p_v + n_i)} \tag{2.16}$$

where τ_n and τ_p are the electron and hole lifetimes, respectively and n_i is the intrinsic charge carrier density. The minority carrier lifetimes are written as

$$\tau_p = \frac{1}{\sigma_p v_{th} N_t} \tag{2.17}$$

$$\tau_n = \frac{1}{\sigma_n v_{th} N_t} \tag{2.18}$$

with the trap density N_t, the thermal velocity $v_{th} = \sqrt{2k_B T/\pi m}$, and σ_n and σ_p are the electron and hole capture cross sections, respectively.

The Shockley-Read-Hall (SRH) mechanism plays an important role in carrier generation within the space charge region of avalanche photodetectors, such as SPADs, because they represent a potential source of noise. In the depletion region, $pn \ll n_i^2$, generation of carriers dominates. The rate of thermal generation is directly related to the density of traps N_t from the above equations that contribute to the dark noise in the photodetectors.

2.2 Single photon avalanche diode

A single photon avalanche diode (SPAD) is a p-n junction that works in reverse bias above the breakdown voltage in the Geiger mode region [17]. When an incident photon is absorbed in the depletion region, it creates electron-hole pairs by the photoelectric effect, and are separated by an internal electric field. The reverse biasing is strong enough to accelerate the charge carriers to create new e-h pairs by the physical mechanism of impact ionization. Once a photon hits the active area, a self-sustained avalanche is triggered in picoseconds.

Figure 2.12 shows the current-voltage characterization of a SPAD. When there is no carrier, the SPAD is ready for detection and stays in region 1 until the arrival of an incident photon or a dark carrier that initiates an avalanche. With the absorption of a photon, an impact ionization process starts, an avalanche is triggered, and the current rapidly increases, thus enters region 2. However, SPAD is not able to stay in this region for a long time due to the heating effect of the huge self-sustained avalanche current, which can burn the device. Therefore, SPAD must be quenched properly with an external series resistor that reduces the reverse voltage below the breakdown voltage to suspend the avalanche where it enters region 3. Once the avalanche current is extinguished, the voltage is increased again through the recharge circuit, making the SPAD ready for the next detection.

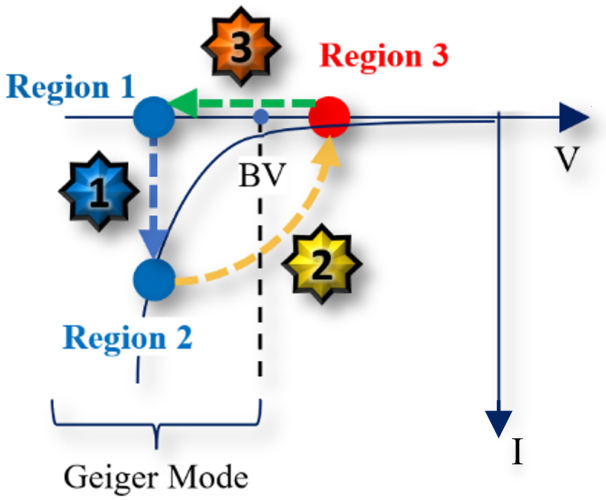


Figure 2.12. I-V characteristics of SPAD [18].

Unlike in linear mode, where the magnitude of the electric field is only enough to provide ionization caused by electrons, the photodiode operating in Geiger mode has an electric field strong enough to cause significant ionization from both electrons and holes and the avalanche will become self-sustaining. In this approach, the multiplication of impact

ionization can be characterized by the ionization rates for electrons α_e and holes α_h. They are defined as probabilities of ionization per unit length and are assumed to be functions of the electric field F. The ionization rates for electrons and holes can be written as

$$\alpha_e(F) = A \exp(-b_e/F) \tag{2.19}$$

$$\alpha_h(F) = B \exp(-b_h/F), \tag{2.20}$$

with A and B being the constants [19].

Quantitatively, a diode is in Geiger mode when it meets the condition [20]:

$$1 \leq \int_0^W \alpha_e \exp\left(\int_x^W (\alpha_h - \alpha_e)\, dx'\right) dx \tag{2.21}$$

where W is the width of the space charge region, α_e and α_h are the ionization coefficients for electrons and holes, respectively. The right term is called the ionization integral and depends on the ionization rate values along the electric field lines [21]. The bias at which Equation 2.21 is an equality is called the breakdown voltage. Since SPADs operate above the breakdown voltage, the difference between the bias voltage and the breakdown voltage is called overvoltage V_{ov}. The breakdown voltage depends on the doping profile of the p-n junction and the temperature.

An important performance parameter of SPAD is the dark count rate (DCR), which corresponds to the Geiger discharge of a pixel without any incoming or detected photons but is caused by the thermal generation of carriers, tunneling of carriers through the junction, and the generation of carriers due to the presence of defects in the material. The detected signals are similar to a photon detection event and cannot be distinguished from regular photon detection. It is generally expressed in counts per second (cps).

Another crucial performance parameter that quantifies its ability to detect low levels of light is sensitivity. It is expressed in terms of photo detection probability, which is the probability that a photon incident on the pixel generates an avalanche.

2.2.1 3D SPAD architecture

Recently, SPAD sensors have profited from the availability of increasingly advanced CMOS technology such as 3D-stacking, which enables new possibilities in terms of sensor architectures and process nodes. In a 3D-stacked architecture, the SPAD sensor resides on one silicon die and is vertically integrated (stacked) on top of a second die containing the supporting electronics using micro-bumps. This approach allows the sensor layer to be optimized independently of the readout layer. As a result, 3D-stacked SPAD sensors can achieve higher photon detection efficiency than traditional implementations [22] [23].

This two-tier method helps avoid compromises between pixel density, sensitivity, and the size of the silicon area. It scales the SPAD-based imaging systems to higher resolutions under Moore's Law, further expanding their utility in cutting-edge applications [24]. 3D stacking allows for higher functional density within a smaller chip footprint.

Traditionally, SPAD sensors are fabricated with frontside illumination (FSI), where light enters from the top of the sensor, passing through any overlying passivation and metal gaps to reach the p–n junction. This approach is simpler to fabricate, where no wafer thinning is required and is compatible with standard CMOS processes. With an ultra-shallow junction in FSI, the UV detection efficiency can be improved with a proper depletion region width. In particular, an optimum ultra-shallow junction has been developed with the PureB implementation for the UV wavelength [25]. However, frontside illumination suffers from several drawbacks. The fill-factor is inherently limited because a significant fraction of the pixel's front area may be occupied by metal lines or other structures that cause light loss due to reflections from mismatched layers above the SPAD or blocks, reducing the sensor's effectiveness.

In contrast, backside-illuminated (BSI) SPADs offer a better alternative by allowing light to enter from the back of the surface of the sensor. In a BSI SPAD, the wafer is thinned from the backside so that photons can reach the SPAD active volume from behind without any losses seen in FSI, like optical losses associated with metal coverage and mismatched refractive indices in the optical stack. The result is a "metal-free" light path, as the entire pixel area on the backside can be photosensitive, drastically increasing the fill-factor, improving photo detection probability and reducing dark count rates [26] [27]. Additionally, the design consideration of the sensor needs to be carefully chosen to minimize the diffusion region related to the distance from the backside of the surface to the depletion region edge towards frontside, as under BSI, it encounters challenges for UV detection. It requires a deeper depletion region with a thin silicon body to extend the drift region close to the surface on the backside for improving the UV sensitivity [27]. However, the SPAD structures on bulk Si also suffers from defects during the backside thinning process for BSI. For this improvement at the UV wavelength range, the use of Silicon-On-Insulator (SOI) technology for BSI SPAD on a thin silicon body, stacked on a buried oxide layer (BOX), has indicated a higher NUV and blue sensitivity with a good DCR [28] [23].

Switching from FSI to BSI in SPAD sensors with a 3D integration approach represents a significant step forward in photodetector design, addressing key challenges and paving the way for advanced, compact photosensors.

2.2.2 Design considerations for NUV light-sensitive SPADs

Extending the sensitivity of SPADs into the NUV regions ($\leq$400 nm) has design challenges. Short-wavelength photons in this range are absorbed very close to the surface of silicon [25]. Therefore, to optimize the optical conversion efficiency, the photosensitive depletion region should be very close to the surface of the light-entrance window, which means the device structure and electric field must be carefully implemented to ensure that those photogenerated carriers trigger an avalanche. It requires careful consideration of several key factors: quantum efficiency, avalanche triggering probability, and geometrical efficiency for efficient detection [21]. Figure 2.13 illustrates a n+/p/π/p+ structure, which is built on a 4 μm thick low-doped epitaxial layer grown on p+ substrate. A junction is created between n+ and p, with n+ doped on the top of the structure and a p region is defined as smaller than n+ to obtain a virtual guard ring structure.

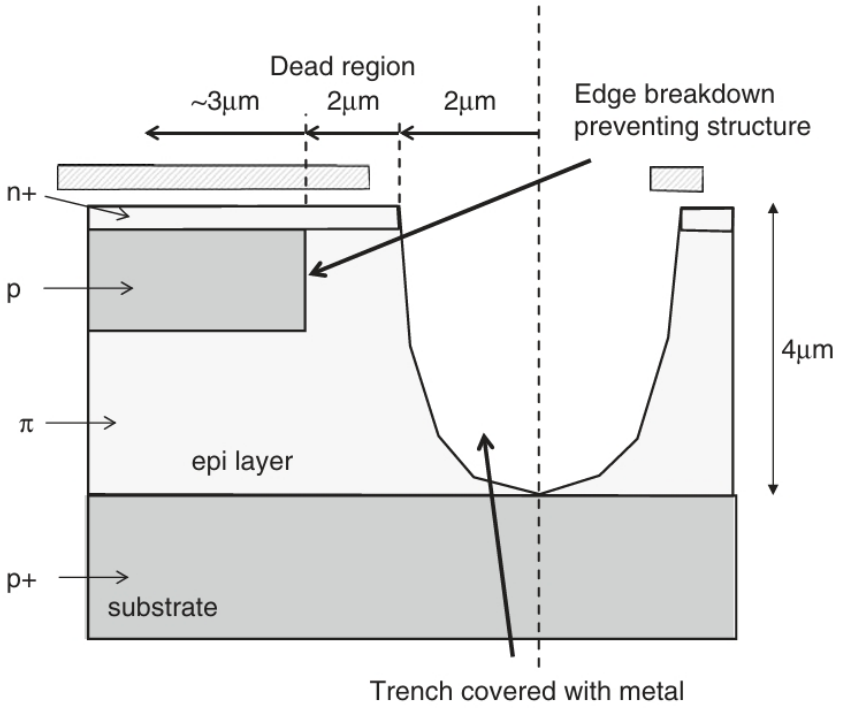


Figure 2.13. Sketch of a conventional structure [21].

The quantum efficiency (QE) of a SPAD at a given wavelength determines the probability that an incident photon is absorbed in the active area and generates an electron-hole pair. However, not every absorbed photon will produce a detectable avalanche; the photon detection probability (PDP) is the product of QE and the probability that the generated carrier initiates a self-sustaining avalanche, called the avalanche triggering probability. For blue light, most photons are absorbed within the first 500 *nm* of silicon, making it essential to minimize recombination losses in this region. This can be achieved through optimized

anti-reflective coatings and shallow n+ layers that limit recombination at the silicon surface. In conventional structures, recombination in undepleted regions and at the silicon/oxide interface can significatly reduce QE, if not addressed. Therefore, precise doping profiles are necessary to enhance performance, particularly the n+, which is close to the surface, and should be as shallow as possible.

The avalanche-triggering probability plays a vital role in determining whether carriers generated by absorbed photons initiate an avalanche. This depends on the alignment of the high-field region with the photogeneration zone and the strength of the applied bias voltage. The high-field region should be as thin as possible to photogenerate more. As avalanche triggering probability depends on the electric field, the strong field maximizes the triggering probability. Figure 2.14 illustrates the importance of electric field design, showing the relationship between the electric field distribution and photon absorption for different wavelengths. This highlights the need for precise alignment between the high-field region and the primary photon absorption depth. For example, at 420 *nm*, only 90% of the photons are absorbed before the maximum triggering probability region, even after such an optimized structure.

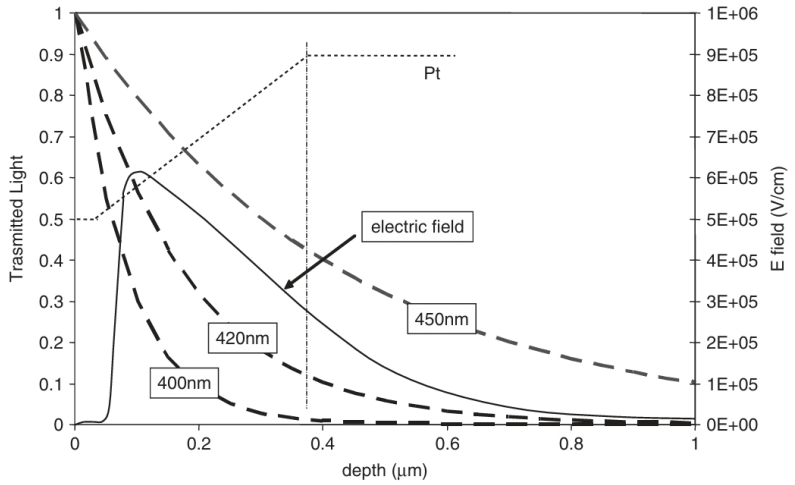


Figure 2.14. Electric field distribution for an n+/p diode with optimized doping profiles for short-wavelength detection. On the same graph, the absorption curves for three wavelengths and a representation of the triggering probability are shown [21].

Geometrical efficiency, defined as the ratio of the active to total device area, is another critical factor. Guard rings, which are necessary to prevent edge breakdown, cause dead regions and reduce this ratio, particularly in conventional designs. Techniques like rounded SPAD geometries help avoid field hot-spots at corners, allowing the guard region to be narrower, thus enlarging the active area for a given pixel.

Fabrication techniques play a crucial role in enhancing SiPM performance. Advanced

ion implantation processes enable precise doping profiles, ensuring uniform electric field distributions and consistent breakdown voltages across the device.

By addressing these considerations, SPADs can achieve superior efficiency for NUV and blue light detection, making them indispensable for applications such as scintillator detectors and low-light imaging systems. The combination of optimized quantum efficiency, triggering probability, and geometrical efficiency, supported by advanced fabrication processes, positions SPADs as a leading technology for detecting low-intensity light in challenging spectral regions.

2.3 Proposed 3D integrated BSI SPAD

Figure 2.15 illustrates a proposed design for the development of a 3D-integrated SPAD. The conceptual design aims to enhance the functionality of SPADs by employing a stacked layer approach. This assembly starts with two distinct tiers, where Tier 1 is dedicated to the readout circuit, which manages all data processing and output functionalities. This circuit is placed on the bottom-tier chip and fabricated using advanced CMOS technology to ensure efficient signal processing, high-speed operation, and low power consumption.

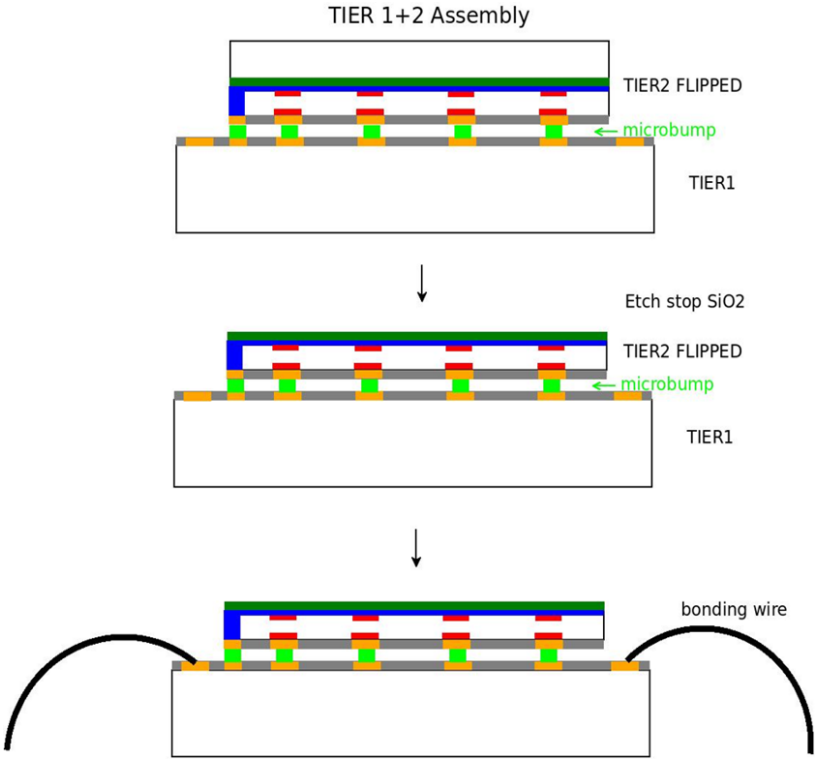


Figure 2.15. Proposed 3D back-side illuminated (BSI) sensor architecture.

Tier 2 comprises an SPAD sensor fabricated on an SOI wafer, positioned above Tier 1. The 3D-stacked architecture provides the freedom to optimize both processes individually,

and therefore DCR and PDP can be improved simultaneously by using a better technology for SPADs and optimizing doping levels and profiles.

The Tier 2 structure begins with an SOI wafer topped with a buried oxide layer on a p-type substrate, as shown in Figure 2.16. The design incorporates the BOX layer on the SOI wafer, serving a dual purpose. It not only supports the structural integrity of the SPAD but also acts as an etching stop during the wafer backside thinning process. Subsequent steps include the deposition of a p-type impurity in the p-type substrate with a p-epitaxially growing layer. After the base, n-type and p-type dopants are implanted into the epitaxial layer to establish n+ and p+ regions essential for forming the pn-junctions. These junctions are vital for the avalanche breakdown process necessary for photon detection. Metal contacts are added to facilitate electrical connectivity and signal transmission on the proposed thin BSI SPAD.

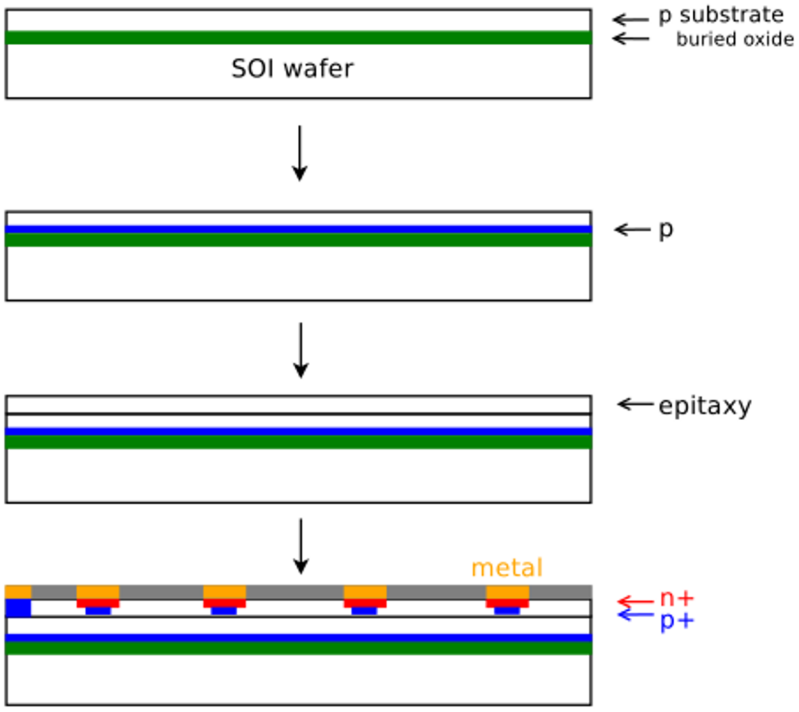


Figure 2.16. Tier 2 profile for 3D integrated BSI SPAD.

In the proposed structure, Tier 2 is flipped and aligned precisely onto Tier 1 using micro-bumps, providing reliable electrical pathways that eliminate the need for bulky traditional wire bonds. These micro-bumps ensure effective communication across the stacked tiers, facilitating the seamless transfer of electronic signals without loss or interference. The wafer is then thinned to make it BSI, where silicon dioxide (SiO_2) acts as an etch stop, protecting the underlying structures. The assembly is completed with wire bonding to establish robust external electrical connections, and providing the electrical inputs and outputs needed for device operation.

The architecture significantly enhances photon detection efficiency by maximizing the fill factor through vertical integration. It cleverly avoids the typical trade-offs between photon detection and signal processing by vertically integrating the pixel processing unit [29]. The design also incorporates vertical control circuitry to manage noisy pixels, which enhances overall image quality. By using adapted fabrication technology tailored specifically for SPAD and CMOS circuits, the design optimizes each component's performance, making it a sophisticated solution for advanced photodetection applications.

This thesis focuses on the development of the SPAD sensor architecture by initially using the simulation tools to validate the design, followed by characterization to ensure it meets the simulated expectations. The research concentrates firstly characterizing the SPAD in an FSI to confirm the working mechanism of the proposed structure.

2.4 Potential applications of SPAD

Due to the high sensitivity, low fabrication cost, excellent timing resolution, and compatibility with CMOS technology, SPADs are used in different applications such as medical fields and dark matter physics experiments.

SPADs are pivotal in the field of dark matter physics, where they are used to detect scintillation photons in rare-event detection experiments like those involving liquified noble gases due to their low background noise, and high photon detection efficiency. Experiments such as DarkSide, nEXO, and DARWIN utilize liquified noble gases, like argon or xenon, as detection mediums [30] [31] [32]. These media emit scintillation photons in the vacuum ultraviolet (VUV) range when interacting with particles such as neutrinos. Traditionally, photomultiplier tubes (PMTs) coated with wavelength shifters have been used to detect these VUV photons, as the shifters convert VUV to blue light, which is easier to detect. However, SiPMs offer a robust alternative. They operate effectively at low voltages, are not susceptible to magnetic fields, and their dark count rates are significantly reduced when cooled to cryogenic temperatures.

Focusing on the DARWIN, the DARk matter WImp search with liquid xenoN, is a next-generation experiment aimed at the direct detection of dark matter. Its main objective is to construct a highly sensitive detector using a multi-ton liquid xenon (LXe) target inside a time projection chamber (TPC). The next-generation dual-phase TPC is filled with liquid xenon and surrounded by light reflectors designed to maximize the collection of VUV light. Two arrays of photosensors are placed on the top and bottom part of the TPC to detect light signals. The bottom sensor array is immersed in the liquid xenon, while the top array is located in the gaseous xenon, as shown in Figure 2.17.

When a particle interacts with the liquid xenon, it deposits energy, producing both

prompt scintillation light and ionization electrons. The scintillation signals are immediately detected by the photosensors, the S1 signal. Meanwhile, the free electrons are guided upwards by a drift field generated between the cathode at the bottom and the gate electrode at the top, separated by approximately 3 meters. Upon reaching the liquid-gas interface, these drifting ionization electrons are extracted into the xenon gas phase. Under a strong extraction field, typically around 5 kV/cm, these electrons generate a secondary scintillation light through electroluminescence, known as the S2 signal, which is then detected by the same arrays of photosensors. The time delay between S1 and S2, in addition to the localization of the S2 light pattern on the top photosensor array, allows precise reconstruction of the three-dimensional interaction vertex. To replace the PMTs, the photosensors should meet the requirement of low dark count rate, high photon detection efficiency, operates at cryogenic temperatures, low power dissipation and highest radio-purity [33].

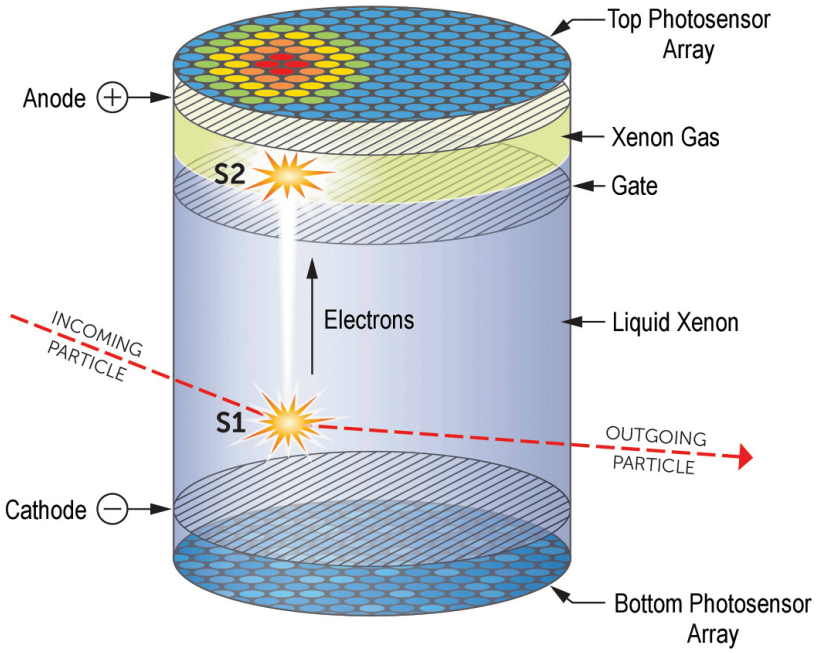


Figure 2.17. Principle of a dual-phase liquid xenon TPC. Energy from a particle interaction within the active liquid xenon volume produces prompt scintillation light (S1) and a delayed signal (S2) from electroluminescence (proportional scintillation) in the gaseous xenon layer. The localization of the S2 signal and the time difference between S1 and S2 allow for the determination of the original vertex location [33].

The initial plan of the project was to design SPADs optimized for the DARWIN experiment. However, before moving toward application-specific optimization, it is necessary to develop a functional SPAD that meets the general requirements set by DARWIN. Therefore, the current focus is on the development of SPAD with the necessary characteristics to eventually integrate it into the DARWIN experiment.

Other applications of SPAD include radiation detectors, particularly in PET, widely utilized in medical fields such as oncology, cardiology, and neuroimaging to visualize physiological activities and metabolic processes within the human body [34], SPAD arrays can be used by combining them with scintillators that are able to absorb high-energy radiation and convert it into a pulse of visible photons. A positron emitter is inserted in the patient. Once emitted, these positrons travel a short distance before annihilating upon collision with electrons in the tissue. This annihilation generates pairs of gamma photons (each with 511 keV of energy) traveling in opposite directions. PET imaging involves capturing these gamma rays through a two-step detection process.

As the first step, the incoming gamma rays are absorbed by a scintillation material that emits visible photons in response to the absorption of higher-energy radiation. The produced visible photons are then detected with silicon photomultipliers (SiPMs), an array of SPADs. By making statistical fluctuations negligible through a sufficiently high number of acquisitions, it is possible to locate the source of these photons along the axis of emission within the patient's body. The spatial resolution of PET can be further enhanced by measuring the Time-of-Flight (TOF) of the detected gamma rays. Including TOF information allows precise localization of photon interaction along the line, effectively increasing spatial resolution and potentially reducing radiation dose or imaging time [35] [36].

Chapter 3

TCAD simulation tools

Technology Computer-Aided Design (TCAD) is an essential tool to simulate the semiconductor fabrication process and analyze device performance. It allows engineers and researchers to model and optimize devices from the initial design stage to the final implementation. Synopsys Sentaurus TCAD is one of the most widely used tools for this purpose, providing a structured approach to semiconductor development while reducing time and cost.

The development of a new semiconductor device involves multiple iterations of simulation, fabrication, and measurement. By simulating the process flow, potential fabrication challenges can be identified early, allowing adjustments before physical prototypes are manufactured. Once a device is fabricated, experimental measurements validate the simulation results, enabling refinements to the model parameters. This feedback loop enhances the accuracy of TCAD predictions and helps optimize the final device design.

TCAD tools solve fundamental semiconductor equations, including Poisson's equation, continuity equations for electrons and holes, and drift-diffusion equations. These equations describe the transport of charge carriers within semiconductor materials and predict device behavior under different operating conditions. The simulations address both static and dynamic characteristics, which are critical for understanding overall device behaviour.

TCAD is broadly categorized into two main branches: process simulation and device simulation. Process simulation models the fabrication steps such as doping, oxidation, implantation, etching, and deposition, allowing users to refine manufacturing techniques and predict how variations in the process impact device performance. Device simulation analyzes electrical performance, including charge transport, optical generation, current-voltage characteristics, and breakdown behaviour under different conditions. Together, these simulations provide insights that guide the design and optimization of semiconductor devices before fabrication.

A modern TCAD software package includes tools like process, device, and visualiza-

tion simulators particularly useful for developing new detectors, such as Single Photon Avalanche Diodes (SPADs). The results obtained from TCAD simulations can be integrated into further studies, such as the investigation of breakdown probability, quantum efficiency, and noise analysis.

The following chapter gives a detailed description of the simulations and measurement methods used in this thesis.

3.1 Process simulation

Sentaurus SProcess is a highly flexible multidimensional process modeling environment. With its modern software architecture, it serves as a solid base for semiconductor process simulation. It enables early evaluation of a device's feasibility, identifies potential challenges, and optimizes performance through virtual pre-characterization. Various design options can be explored and understood quickly and efficiently [37].

In process simulation, fabrication steps such as etching, deposition, ion implantation, thermal annealing, and oxidation are modeled based on the fundamental physical equations governing these processes by offering precise control over doping profiles and topological structures. Users can also select between different implantation and diffusion modeling methods, including analytical and Monte Carlo approaches. The silicon wafer structure is discretized into a finite-element mesh, providing a computational framework for solving complex diffusion equations during thermal annealing or simulating oxidation growth while considering oxygen diffusion and mechanical stress effects.

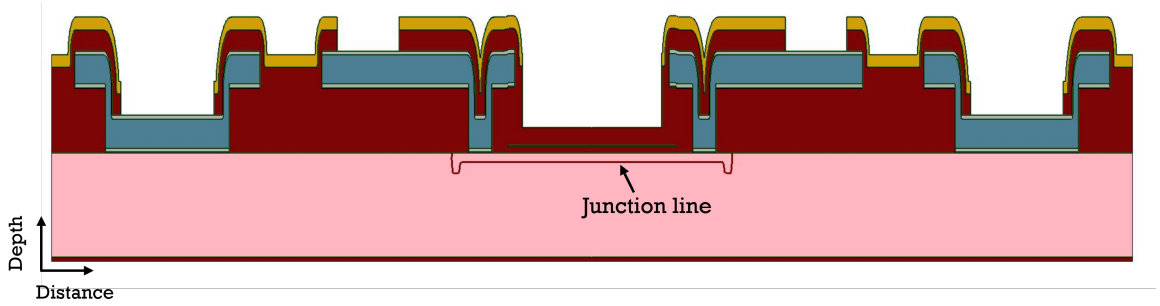


Figure 3.1. Simulated cross-sectional structure of a Single Photon Avalanche Diode (SPAD), indicating the junction line. The pink region represents the active detection area, with deposited metal contacts and anti-reflective coating layers on top, and a thin buried oxide layer below.

Additionally, process simulations generate the input data required for subsequent device simulation, ensuring smooth integration between the process and device modeling stages.

The critical regions of the device where doping profiles change significantly, such as photoresist edges or topological transitions are simulated in detail. This targeted simulation allows for finer mesh resolution and precise results with less computing time. Later, these detailed simulation results are combined to form the complete sensor structure for comprehensive device-level simulation.

Figure 3.1 shows the simulated cross-section of a SPAD. In this representation, various layers resulting from different fabrication processes are clearly visible. The central highlighted junction line indicates the position of the pn-junction formed by doping processes. The structure also includes multiple layers, such as the metal, contacts and anti-reflective coatings above the active region and a thin buried oxide layer underneath, necessary for electrical isolation and optimal sensor performance.

3.2 Device simulation

Sentaurus device is a numeric device simulator designed to analyze the electrical, thermal and optical characteristics of a semiconductor device. The basic working principle of device simulations is to numerically solve the equations that describe the transport of charge carriers and electrostatic interactions within semiconductor materials [38].

To accurately simulate semiconductor devices, Sentaurus Device solves a set of fundamental equations, including Poisson's equation for electrostatic potential and continuity equations for electrons and holes. Selecting appropriate physical models is crucial to ensure realistic and accurate simulation results. Specifically, for SPADs, critical phenomena like carrier transport, photon-induced carrier generation, impact ionization, trapping mechanisms, and band-to-band tunneling must be modeled precisely.

The structure of a typical device simulation setup in Sentaurus Device follows a structured approach. Firstly, the file section handles input files for device geometry and doping profiles taken from SProcess module and defines the output files. The electrode section defines the electrical device contacts with applied initial boundary conditions. In the physics section, the relevant semiconductor physical models such as carrier mobility, carrier transport and generation-recombination, are selected. The plot section specifies which electrical or optical variables are to be monitored and recorded during the simulation. Numerical methods and solvers, essential for efficiently handling complex differential equations, are defined in the math section. Lastly, the solver section consists of a series of commands that are executed sequentially with the steady-state approximation to perform analysis over a range of operating conditions and ensure the convergence for the initial Poisson equation with the specific applied bias conditions.

The simulation process treats the device as a finite-element mesh structure. Every node

in this mesh is assigned physical properties, including doping concentration and material type. Device simulations compute several variables at each node, such as carrier concentrations, current densities, electric fields, and recombination rates. Electrodes, represented as specific boundary areas in the mesh, have electrical conditions like applied bias voltages. Solving the governing equations yields measurable outcomes like electrical currents at device contacts, enabling an in-depth understanding and optimization of device performance.

3.2.1 Carrier transport

In semiconductor device simulations, the classical drift-diffusion model is widely utilized to characterize the transport of charge carriers, electrons and holes, and is the default carrier transport model in Sentaurus Device. Within this model, the current flow arises due to two fundamental phenomena: drift, driven by the presence of an electric field, and diffusion, which results from gradients in carrier concentration [12]. The drift-diffusion equations mathematically represent these phenomena, expressing electron and hole current densities. The electron current density is described as

$$\vec{J_n} = e\mu_n n\vec{E} + eD_n\vec{\nabla}n \tag{3.1}$$

Similarly, the hole current density is defined as

$$\vec{J_p} = e\mu_p p\vec{E} - eD_p\vec{\nabla}p \tag{3.2}$$

In these equations, *e* represents the elementary charge (1.6×10^{-19} C), μ_n and μ_p denote the mobilities of electrons and holes, respectively; *n* and *p* indicate their respective densities, and D_n and D_p correspond to diffusion constants, and $\vec{\nabla}n$ and $\vec{\nabla}p$ are concentration gradients.

Now, from the Einstein relation which relates the diffusion coefficient with temperature and mobility

$$D = \frac{kT}{e}\mu \tag{3.3}$$

and the electric field relates to the electrostatic potential by

$$\vec{E} = -\frac{dV}{dx}, \tag{3.4}$$

The current densities can be rewritten to include quasi-Fermi potentials, ϕ_n and ϕ_p:

$$\vec{J_n} = -ne\mu_n\vec{\nabla}\Phi_n \tag{3.5}$$

$$\vec{J_p} = -pe\mu_p\vec{\nabla}\Phi_p \tag{3.6}$$

Poisson's equation is employed to determine the electrostatic potential distribution within the semiconductor. This equation links the electric field to the charge density within the device as

$$\nabla \cdot \vec{E} = \frac{\rho}{\varepsilon} \tag{3.7}$$

where ρ represents the total charge density, including ionized dopants as well as free electrons and holes, and ϵ is the material permittivity.

The electrostatic potential calculated by Poisson's equation provides the electric field distribution inside the device, essential for modeling the drift currents and thus the overall device behavior.

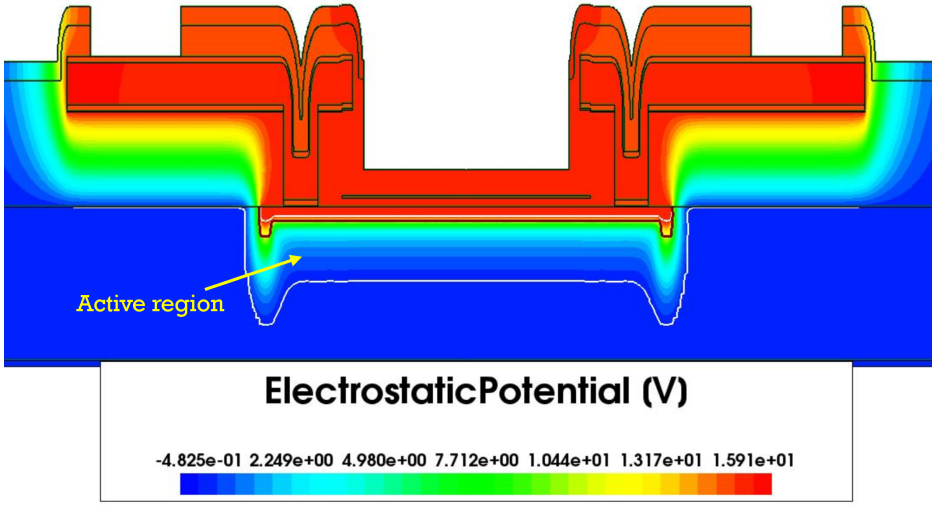


Figure 3.2. Simulated electrostatic potential distribution across the SPAD. The color gradient indicates the variation in potential (V), highlighting regions of high potential near junction interfaces and electrodes, and lower potential within the substrate region.

Figure 3.2 shows the simulated electrostatic potential distribution across SPAD calculated by solving Poisson's equation. The color map indicates potential values in volts (V), ranging from negative values (blue regions) to higher positive potentials (red regions). The regions with dark red color indicate areas of higher positive potential, corresponding to heavily doped regions or metal contacts. The blue region shows the lower potential region, the substrate or the deeper active region. The variation in potential clearly shows the formation of high electric fields at junction interfaces, which can lead to critical phenomena such as avalanche breakdown or impact ionization.

3.2.2 Carrier generation in space charge region

In semiconductor devices, especially within the space-charge region (SCR), electrons and holes can be created or recombined through different physical processes. These mechanisms involve carriers transitioning between the conduction and valence bands. For SPADs, it is particularly important to accurately simulate these generation processes, because electron-hole pairs created under high electric fields can trigger avalanche multiplication, directly affecting device performance. One of the critical processes modeled here is Shockley-Read-Hall (SRH) recombination and generation [39]. SRH generation and recombination occur due to defect levels (traps) located within the semiconductor's bandgap [40]. In TCAD simulations, the net SRH recombination rate ($R_{\text{net}}^{\text{SRH}}$) is calculated by subtracting the generation rate from the recombination rate, as described by the following equation:

$$R_{\text{net}}^{\text{SRH}} = \frac{np - n_{i,\text{eff}}^2}{\tau_p(n + n_1) + \tau_n(p + p_1)} \tag{3.8}$$

Here, *n* and *p* represent electron and hole concentrations, respectively, and $n_{i,eff}$ denotes the effective intrinsic carrier concentration, which takes bandgap narrowing effects into account. Additionally, n_1 and p_1 depend on the trap energy level within the bandgap, calculated by

$$n_1 = n_{i,\text{eff}} \exp\left(\frac{E_{\text{trap}}}{kT}\right) \tag{3.9}$$

and

$$p_1 = n_{i,\text{eff}} \exp\left(-\frac{E_{\text{trap}}}{kT}\right) \tag{3.10}$$

where E_{trap} is the trap energy level relative to the intrinsic level (mid-gap), typically assumed to be zero by default (mid-gap traps), and kT is the thermal energy.

The SRH model acts as a recombination mechanism when $np > n_i^2$ ($R_{\text{net}}^{\text{SRH}} > 0$). In contrast, when $np < n_i^2$, the model acts as a generation mechanism, typically in a dark regime.

Another important carrier generation mechanism, especially significant in heavily doped p-n junctions, is band-to-band tunneling (B2B). Band-to-band tunneling occurs when the electric field at the junction exceeds a certain threshold [41]. Under these conditions, electrons in the valence band of the p-type region directly tunnel to the conduction band of the n-type region, creating electron-hole pairs. In this thesis, the simplest band-to-band tunneling model is employed, which can be mathematically described as:

$$G^{B2B} = AE^p \exp\left(-\frac{B}{F}\right) \tag{3.11}$$

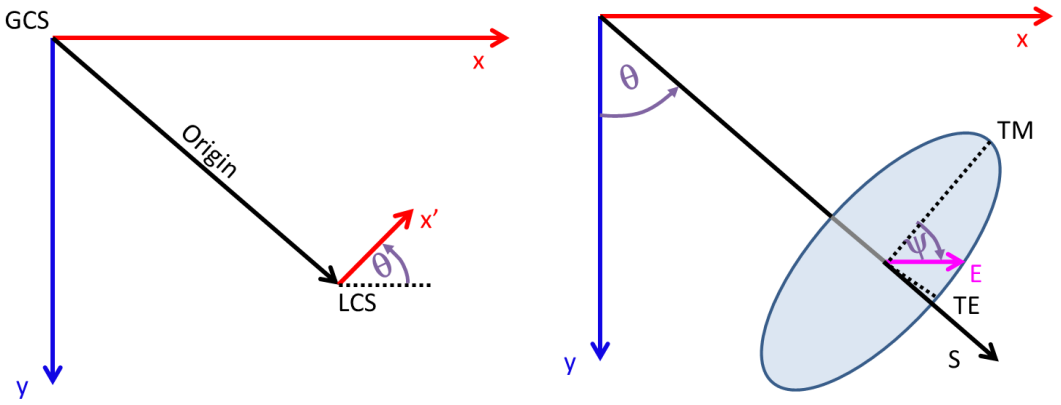
where G_{B2B} represents the tunneling generation rate, F is the electric field, and A and B are material-dependent constants used as fitting parameters. p takes the value 1, 1.5, or 2,

depending on the model.

3.2.3 Optical device simulation

To calculate the optical generation rate within a semiconductor device, various optical solvers are available in the Sentaurus Device. The optical solver employed in this thesis is the Transfer Matrix Method (TMM), which specifically computes the optical generation from a monochromatic optical source. When an optical wave penetrates the device structure, it is partially absorbed, generating electron-hole pairs. The rate of this optical generation is derived from the product of the absorbed photon density, calculated by the optical solver, and the quantum yield.

A crucial aspect of optical simulations for photodetectors, is the concept of an illumination window. This window confines incident light onto a particular area of the device structure. The Sentaurus Device offers a flexible user interface that allows defining one or multiple illumination windows and provides the capability to dynamically move these windows during the simulation by adjusting relevant parameters. Each illumination window is defined within a local coordinate system (LCS) that can be shifted with respect to the global coordinate system (GCS). In two dimensions, the LCS is a line (x'), shown in Figure 3.3(a).



(a) Definition of the LCS with respect to the GCS in two dimensions.

(b) Angle definition for excitation direction and polarization in two dimensions.

Figure 3.3. Coordinate system definition in two dimensions [38].

The excitation section describes the properties of the excitation of plane wave and its direction is defined with respect to the GCS given by the grid. It includes parameters such as Intensity, Wavelength, Theta θ, and PolarizationAngle ψ or Polarization. Specifically, θ defines the angle between the propagation direction and the positive y-axis with counterclockwise orientation in two-dimensional simulations, shown in Figure 3.3(b). The illumination window section, combined with the excitation section, specifies illumination parameters

essential for precise optical simulation using the TMM solver.

The TMM solver requires extracting a one-dimensional (1D) layer stack from the actual multidimensional device structure, simplifying the analysis to a 1D optical problem, as shown in Figure 3.4. The layer stack extraction algorithm works by extracting all grid elements along a line normal to the corresponding illumination window starting from a user-specified position within the window. By default, all elements belonging to the same region will form a single layer, which is part of the entire stack. This assumes that the material properties do not vary within the region. The surrounding media at the top and bottom of the extracted layer stack are assumed to have the material properties of vacuum by default.

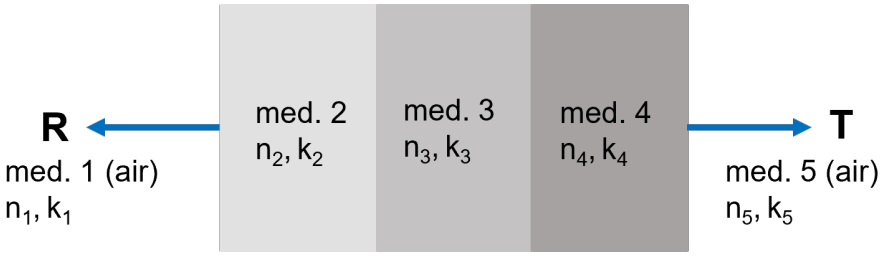


Figure 3.4. Schematic of the Transfer Matrix Method (TMM) for optical generation calculation in Sentaurus Device.

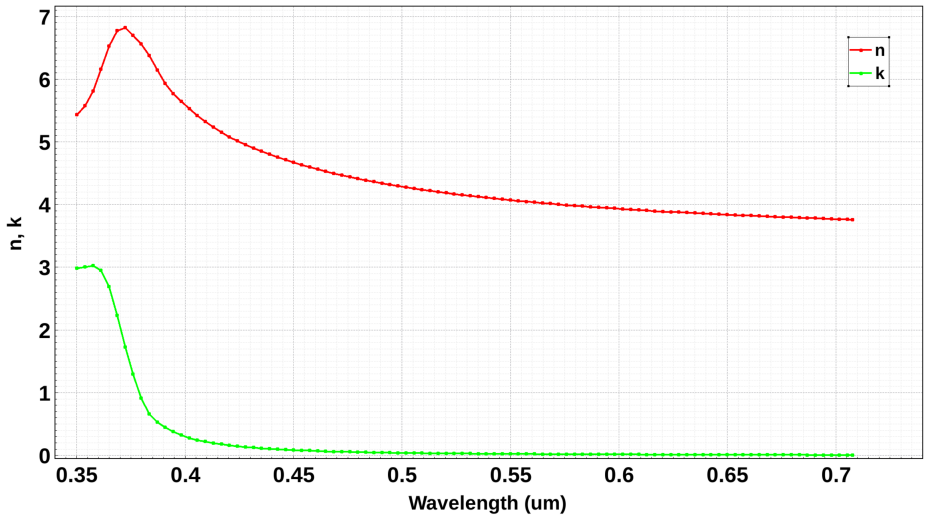


Figure 3.5. Simulated wavelength-dependent refractive index (n) and extinction coefficient (k).

The TMM approach models monochromatic plane waves penetrating parallel planar layers at various angles of incidence and polarization states. In the simulation setup, each region of the device is characterized by a complex refractive index, expressed by the refractive index, n and the extinction coefficient, k. Figure 3.5 illustrates an example of a wavelength-dependent complex refractive index used for the simulation results shown in

Chapter-4. The values of *n* and *k* can be specified using various models such as constant values, wavelength-dependent models, temperature-dependent models, carrier-dependent models, or custom-defined models. The optical solver, based on these complex refractive index values, calculates the absorbed photon density at each vertex in the simulation mesh, quantifying the number of photons absorbed per unit volume and unit time.

The simplest quantum yield model defines how the absorbed photon density is converted to the optical generation rate, and it is assumed that each absorbed photon generates exactly one electron-hole pair in this thesis.

Figure 3.6 shows the simulated absorption coefficient as a function of wavelength that calculates how quickly the intensity of light decreases as it penetrates into the material. This behavior is described by Beer-Lambert's Law [10], which states that the intensity of light (I) passing through a semiconductor decreases exponentially with depth (x), given by

$$I(x) = I_0 e^{-\alpha x} \tag{3.12}$$

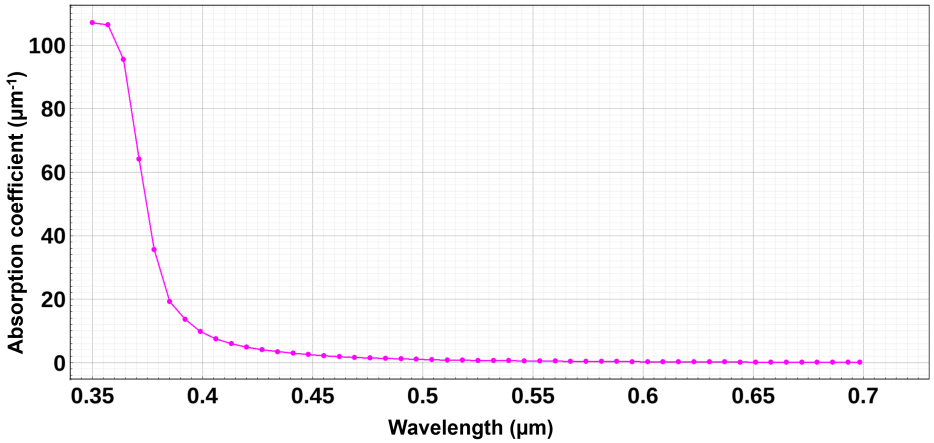


Figure 3.6. Simulated absorption coefficient as a function of wavelength.

where I_0 is the incident light intensity, and α is the absorption coefficient, which is simulated from *k* and wavelength λ according to

$$\alpha = \frac{4\pi}{\lambda} \cdot k \tag{3.13}$$

3.3 SVisual

Sentaurus Visual is an advanced visualization tool to visualize and analyze results obtained from TCAD simulations [42]. It provides interactive and flexible visualization capabilities through a graphical user interface (GUI), making it easier to interpret complex simulation

data. Additionally, Sentaurus Visual supports the automation of repetitive tasks and analysis through scripting, using the Tcl scripting language.

Users can visualize simulation data clearly, facilitating an intuitive understanding of complex semiconductor device behavior. It can handle various simulation outputs, such as doping profiles, current-voltage (IV) curves, carrier concentrations, electric fields, electrostatic potentials, and optical properties. Furthermore, tasks like data extraction, plotting, and visualization customization can be automated with built-in Tcl scripting support, making the analysis process efficient and repeatable.

Chapter 4

SPAD layout and TCAD simulation results

This chapter presents the TCAD simulation of the SPAD, along with the post-processing and analysis of the results. Initially, the results from the process simulation environment are discussed, focusing on essential fabrication steps such as epitaxial growth, doping profiles, diffusion, and implantation processes. Subsequently, electrical characterization results, including current-voltage (IV) characteristics, electric field distribution, and ionization coefficients, are analyzed. The simulated results are further processed to estimate the photon detection probability. Various models used in the simulation are described and justified according to their specific roles and suitability.

4.1 TCAD simulation results

This section presents the simulation results of the SPAD device using Synopsys TCAD tools. It includes both process simulation and device-level results, focusing on geometry setup, doping profiles, junction formation, and physical parameter analysis such as electric field, IV characteristics, and impact ionization.

4.1.1 Geometry definition and doping processing steps

The desired active structure of SPAD for blue and NUV sensitivity with the corresponding doping profile and electric field can be seen in Figure 4.1, where the maximum electric field occurs at the n+/p+ junction. The geometry and doping profiles are set up in the Sentaurus SProcess module, as previously described in Chapter 3. The introduction of doping impurities can be carried out by different techniques: diffusion, ion implantation, and growth of a doped silicon layer on an existing substrate (epitaxy). Impurity diffusion is carried out

at high temperatures, where the impurity atoms can diffuse throughout the crystal lattice through interactions with point defects such as interstitials and vacancies. Epitaxy is a processing technique in which a single-crystal layer of silicon is grown on silicon at a high temperature. In the ion implantation process, implanted ions are accelerated and come to rest at a specific penetration depth called the projected range after giving up some energy, which depends on the implantation energy used. This loss of ions occurs due to collisions and interactions with the atoms in the substrate. This behavior is crucial for controlling where the ions settle within the material, which directly affects the electrical and optical properties of the sensors. Generally, the higher the energy of the implanted ions, the deeper they penetrate into the material. The profile of the implanted ions can be described with reasonable accuracy using a Gaussian distribution. The concentration of implanted impurities is described by the following relationship:

$$C(x) = C_p \exp\left[-\frac{1}{2}\left(\frac{x - R_p}{\Delta R_p}\right)^2\right] \tag{4.1}$$

where C_p is the concentration at the peak of the Gaussian distribution, R_p is the projected range and ΔR_p is the standard deviation called the straggle, which measures the width of the distribution [43]. Photolithography is used to perform ion implantation or etch a material in selected areas on the silicon surface by transferring a required pattern.

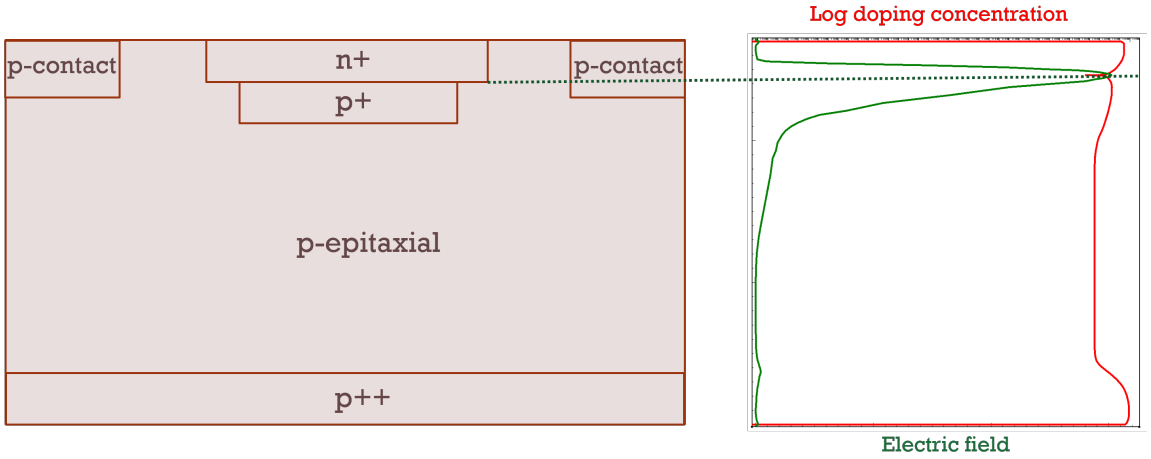


Figure 4.1. Desired active structure of SPAD with doping concentration and electric field distribution.

The following steps illustrate the fabrication processes used for simulating the SPAD structure, which include oxidation, epitaxial growth, diffusion, implantation, and annealing.

1. **Buried oxide layer deposition and initial boron doping**: The first step involves the deposition of a thin buried oxide layer (BOX) on the SOI wafer, which acts as an etch stop during the wafer backside etching process [44]. The initial layout parameters, including positioning and mesh resolution, are defined. An oxide layer of 0.1 μm thickness is deposited at the bottom. On top of this BOX layer, a highly doped boron layer with a concentration of 1.0×10^{-19} cm^{-3} and a thickness of 0.2 μm is deposited to form the initial p-type region, illustrated in Figure 4.2. The standard wafer orientation is specified as (100), offers a flat and smoother surface.

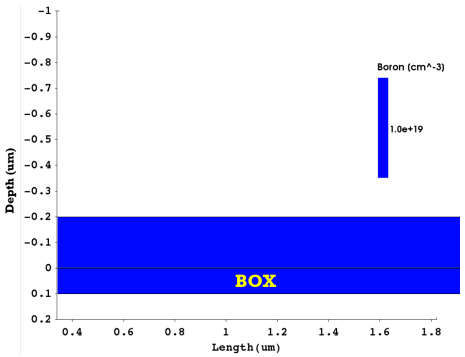


Figure 4.2. Initial structure showing the buried oxide (BOX) deposition with a p-type (boron) highly doped layer on top.

2. **Epitaxial layer growth and thermal diffusion**: A lightly doped epitaxial layer is grown on top of the highly doped p-type layer and acts as an active silicon region of SPAD, shown in Figure 4.3. The thickness of this epitaxial layer is set at 2.5 μm, close to the value used in the standard configuration for short wavelength light detection [21]. After deposition, the layer undergoes thermal diffusion at high temperatures to achieve a uniform doping profile and proper junction formation.

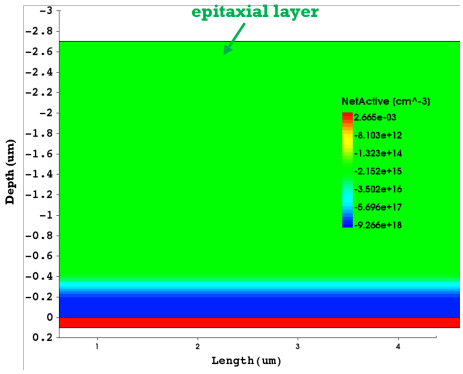


Figure 4.3. Simulated doping profile after epitaxial layer growth and thermal diffusion.

3. **Screening oxide growth**: After the epitaxial layer formation, a thin oxide layer, commonly known as a screening oxide, is grown on the top surface, as shown in Figure 4.4. This oxide layer typically serves to protect and isolate the underlying epitaxial layer from damage during further processing steps such as implantation and diffusion.

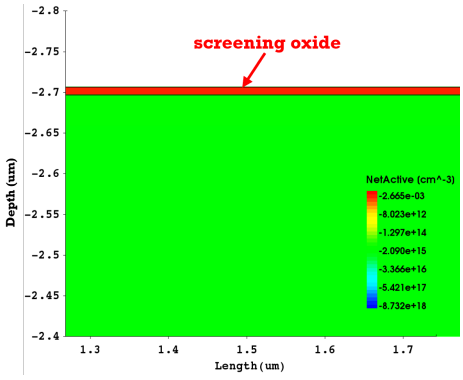


Figure 4.4. Screening oxide layer grown on top of the epitaxial layer (highlighted in red).

4. **p-type implantation**: A shallow p-type doping region is created on the top of the epitaxial layer through low-energy implantation at 10 *keV*. This step is confined to the top tens of nanometers of silicon and designed to avoid the channeling effect that can lead to unintended dopant penetration depth and should be incorporated with a screening oxide. Figure 4.5 presents the simulation result, clearly showing the shallow p-type doping profile as a lighter blue region near the top surface of the epitaxial layer.

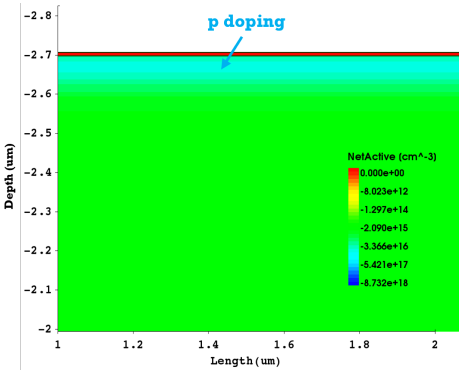


Figure 4.5. Boron doping (p-type implantation) showing a shallow, lightly doped region close to the epitaxial layer surface.

5. **High-energy p-type implantation**: Following the shallow implantation, another p-type implantation is performed at a higher energy of 95 *keV* by defining the photoresist masks to limit the area where p-type doping will occur. This higher energy allows boron atoms to penetrate more deeply into the epitaxial layer, creating a moderately doped region deeper inside the substrate. Figure 4.6 displays the simulation result, highlighting the deeper, moderately doped p-region clearly visible as the lighter blue area extending deeper into the substrate.

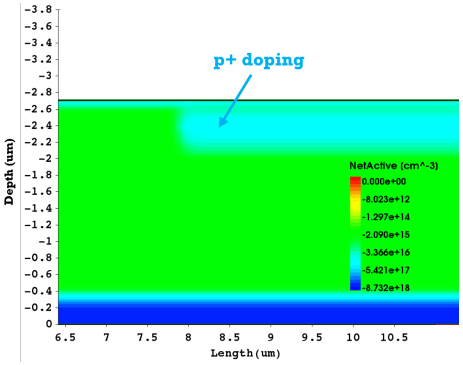


Figure 4.6. High-energy boron implantation, resulting in a deeper, moderately doped region within the epitaxial layer.

Total boron profile: To evaluate the overall doping distribution within the device, the total boron concentration profile is visualized in Figure 4.7. This cumulative view reflects the result of multiple boron implantation steps, both low and high energy, forming a gradient doping structure throughout the vertical depth. The gradient helps in shaping the electric field across the device, ensuring better control over breakdown behavior.

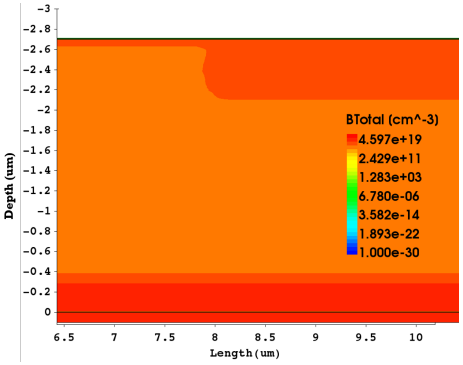


Figure 4.7. Simulated total boron concentration showing a gradient doping profile formed by sequential p-type implantations.

6. **n-type implantation**: To form the required p-n junction for the SPAD operation, 70 *keV* energy of n-type doping (phosphorus) is implanted near the surface of the epitaxial layer by defining the photoresist masks, shown in Figure 4.8. This results in a shallow n+ region on the previously implanted p-type regions, forming a junction.

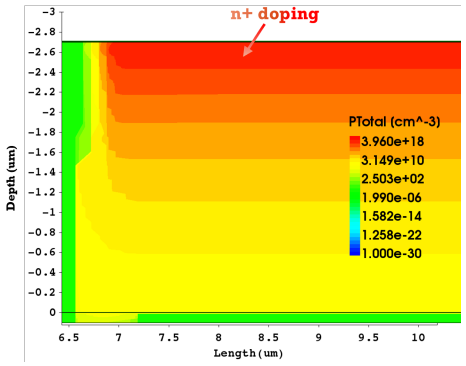


Figure 4.8. Simulated phosphorus (n-type) doping profile implanted at the top of the high-energy p-type implantation.

7. **Ohmic contact formation**: An additional shallow n-type (arsenic) doping is implanted using photoresist masks with 30 *keV* of energy near the surface to enable a reliable ohmic contact with the metal contact layer. Figure 4.9 illustrates the arsenic doping distribution in red color.

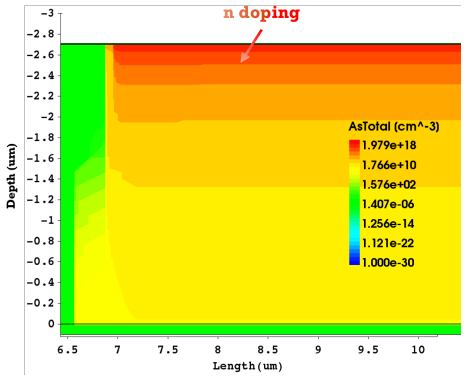


Figure 4.9. Arsenic (n-type) doping profile implanted for ohmic contact formation.

8. **p contact formation**: To enable proper biasing of the SPAD device under reverse voltage conditions, a p-type doping is implanted for p-contact formation using photoresist masks. This doping ensures a well-defined ohmic contact on the p-side of the junction, which is critical for stable electrical operation during negative biasing. Figure 4.10 shows the boron concentration profile in the region where the strong and shallow p+ region is formed for p-contact.

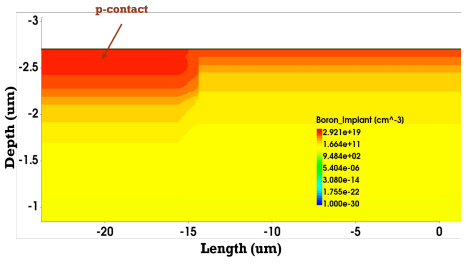


Figure 4.10. Boron doping profile for p-contact formation showing a highly doped p+ region near the surface for negative bias connection.

After the doping processes are completed, the annealing step is carried out together with the anti-reflective coating deposition. An oxide layer is first deposited on the top of

the surface using the LPCVD TEOS method. To densify the interface between thermal oxide and TEOS, annealing is performed at 900°C. Following this, a silicon nitride layer is deposited to serve as the ARC. Additional layers are then added to complete the device structure. These include metal layers and contacts, which are introduced through processing techniques such as diffusion with patterned masks, etching, and oxide deposition.

4.1.2 SPAD schematic representation

In order to have a reliable model, the doping parameters need to be precisely selected and simulated since the intrinsic characteristics of SPAD are strongly dependent on them. The final cross-sectional schematic of the SPAD device is shown in Figure 4.11. This is a half-device view, meaning the other half (to the right) is a mirror image of the shown structure. It is a square-shaped n+/p+/p-epi/p++ structure with the p-n junction formed between the n+ and p+ regions. The active depth of the structure is around 2.7 μm. The figure also highlights the important structural components such as the BOX layer, virtual guard ring created due to the gradient profile of n+ and p+ type implantations, p-epitaxial region, and p-contact, along with the different deposited layers like Ni, Ti, TiN, Al, and oxide used for electrical isolation and connections.

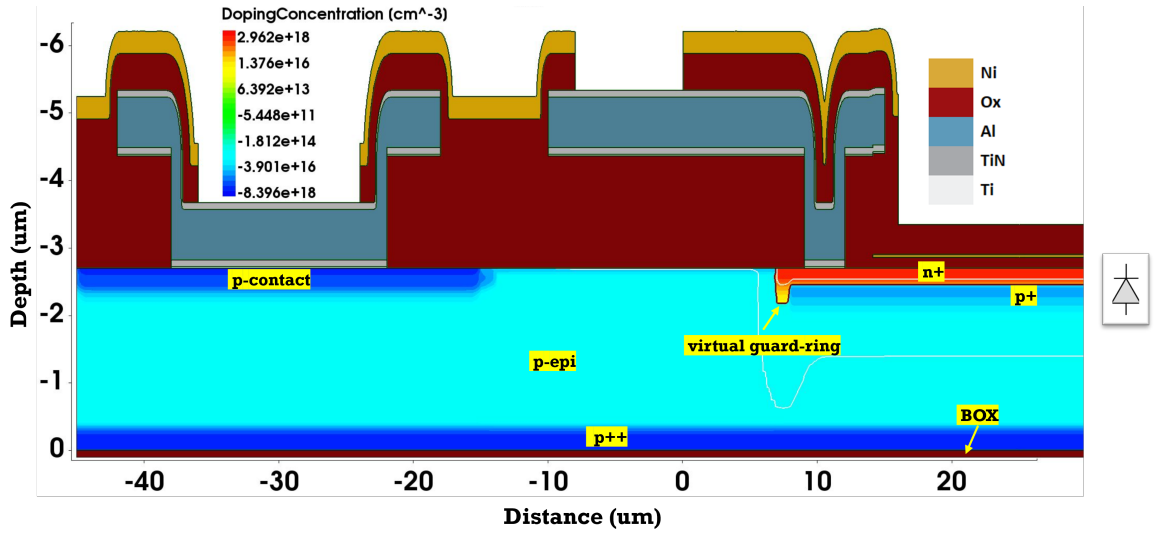


Figure 4.11. Final half cross-sectional view of SPAD showing all regions and material layers.

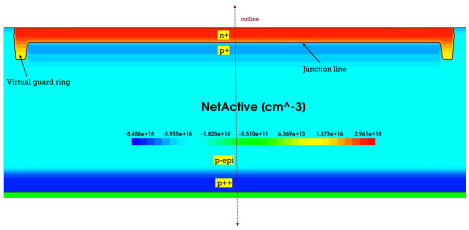
4.1.3 Doping profile extraction

The doping profile of the SPAD device is extracted using the SProcess module by making a cutline through the central region of the structure. This analysis provides insight into how different dopant species, primarily boron (B), phosphorus (P), and arsenic (As), are dis-

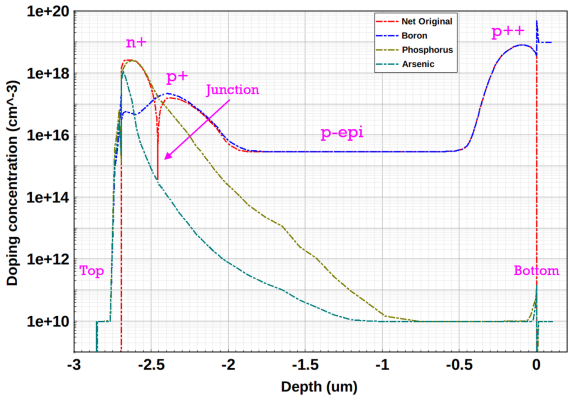
tributed throughout the device. These dopants form critical junctions and contacts necessary for the SPAD operation.

Figure 4.12(a) shows the 2D net doping concentration distribution across the SPAD. The p-type and n-type regions are clearly distinguishable by color gradients, with red regions indicating a high concentration of donors (n-type) and blue indicating acceptors (p-type) with a cutline in the middle.

Figure 4.12(b) shows the extracted doping profile across depth, taken along the vertical cutline. It plots the doping concentrations of boron, phosphorus, and arsenic, along with the net doping, on a logarithmic scale. The sharp dip in the plot marks the n-p junction, which is the active multiplication region of the SPAD. The arsenic doping is located nearer the top surface, corresponding to contact formation.



(a) Net doping concentration distribution with a vertical cutline across the SPAD structure.



(b) 1D simulated doping profile across the vertical cutline on a logarithmic scale: boron, phosphorus, arsenic, and net concentration.

Figure 4.12. Doping profile distribution from the simulated SPAD structure.

4.1.4 Electrical simulation

After simulating the device geometry and doping steps in the process simulation environment, the resulting structure is imported into the SDevice module for electrical simulations. In this step, electrical contacts are defined, and biasing conditions are applied. Several physical models are activated within the simulator, such as carrier transport, carrier generation under illumination, and impact ionization, as discussed earlier in Chapter 3. These models must be carefully selected to ensure that the behavior of the simulated device matches the characteristics of an actual fabricated SPAD.

Figure 4.13 shows the simulated SPAD structure, highlighting the depletion region formed under reverse bias. The depletion depth grows with the increase in applied reverse voltage.

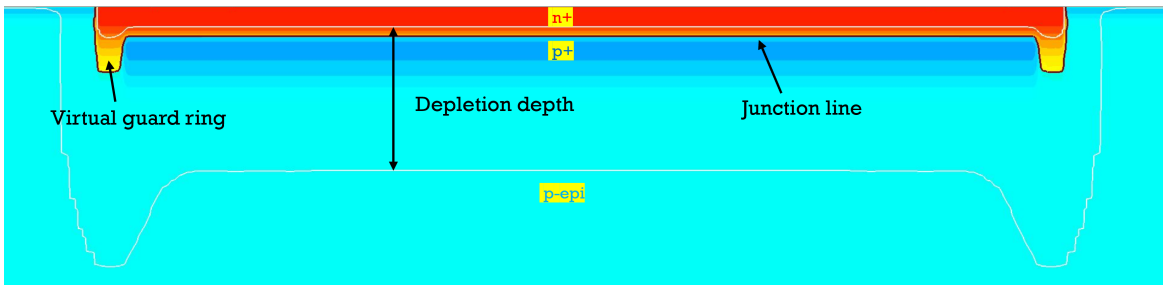
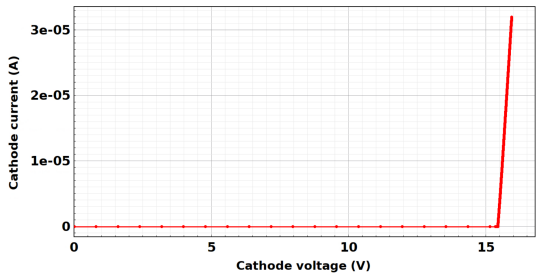


Figure 4.13. Simulated SPAD cross-section showing the depletion depth under reverse bias, along with labeled regions. The junction line and depletion boundary are clearly visible.

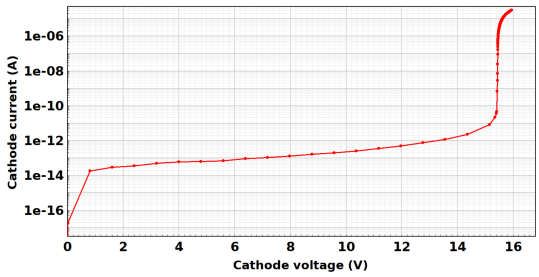
IV characteristics

When a reverse bias is applied across the SPAD device, the electric field within the depletion region increases significantly. Once the applied voltage exceeds the breakdown voltage, avalanche multiplication occurs due to the impact ionization, leads to a rapid rise in current.

The simulated result of the reverse I-V characteristics of the SPAD is shown in Figure 4.14 in a linear and logarithmic scale at room temperature. The logarithmic scale is particularly useful in visualizing the leakage and dark currents at low voltages and the significant rise in current as breakdown initiates. Since SPADs operate above breakdown voltage, it is extracted and comes out to be 15.25 V.



(a) Linear IV curve.



(b) Logarithmic IV curve.

Figure 4.14. Simulated reverse I-V characteristics of the SPAD at room temperature. Breakdown occurs at 15.25 V, as seen from the sharp increase in cathode current.

The reverse IV characteristics are simulated at different temperatures and plotted in Figure 4.15, where the temperature coefficient comes out to be 21.2 mV/°C. With the increase in temperature, the breakdown voltage rises. The energy loss of a carrier to the lattice during its travel through the depletion region increases with the rise in temperature, reducing in turn the ionization rate. To compensate for this loss, the field strength, and that means the applied voltage, must be increased to get a specific value of multiplication factor and thus generate a self-sustained avalanche process [45].

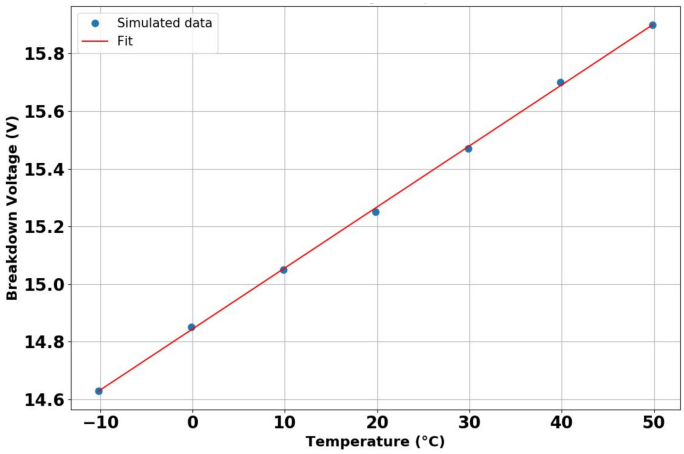


Figure 4.15. Simulated breakdown voltage at different temperatures.

Ionization coefficient and avalanche generation

As the SPAD device is subjected to increasing reverse bias, the electric field across the depletion region becomes strong enough to accelerate free carriers, electrons and holes, within the junction. These high-energy carriers can initiate impact ionization by colliding with silicon atoms and generating additional electron-hole pairs, thereby initiating a carrier multiplication process that leads to avalanche multiplication.

The rate of electron-hole pair generation due to impact ionization can be expressed as [41]

$$G^{ii} = \alpha_n n v_n + \alpha_p p v_p \tag{4.2}$$

Here, α_n and α_p are the ionization coefficients for electrons and holes, respectively, while v_n and v_p are their carrier velocities. n and p are the concentrations of electrons and holes, respectively. The ionization coefficients, explained in Chapter 2, represent the number of new electron-hole pairs generated per unit distance traveled and are strongly dependent on the magnitude of the local electric field, usually expressed in cm^{-1}. These coefficients are critical parameters in accurately modeling the avalanche process. In SPADs, they directly influence the reverse I-V characteristics, breakdown voltage, and avalanche triggering probability. Notably, in silicon, the ionization coefficient for electrons is higher than that for holes, $\alpha_n > \alpha_p$.

Avalanche breakdown, also known as Geiger discharge, occurs when the rate of impact ionization becomes sufficiently high to cause a self-sustaining increase in carrier generation. This phenomenon leads to a rapid and irreversible increase in current at the breakdown voltage. Mathematically, this condition can be approximated when either of the following ionization integrals equals one:

$$1 = \int_0^W \alpha_n \exp\left(-\int_0^x (\alpha_n - \alpha_p)dx\right) dx = \phi_n \tag{4.3}$$

$$1 = \int_0^W \alpha_p \exp\left(-\int_0^x (\alpha_p - \alpha_n)dx\right) dx = \phi_p \tag{4.4}$$

Here, W denotes the width of the depletion region, and ϕ_n and ϕ_p are the ionization integrals for electrons and holes, respectively. These integrals are evaluated using the approximate breakdown analysis (ABA) model from SDevice module along the electric field lines through the depletion region and provide a fast yet reasonably accurate estimate of the breakdown voltage. Though not as precise as full I-V simulation, they correlate well with breakdown voltage observed in IV characterization in the above section.

Figure 4.16 illustrates the simulated ionization integrals for both electrons and holes as a function of applied reverse voltage. The breakdown occurs when either integral reaches unity, indicating the onset of avalanche multiplication. The hole integral reaches unity slightly later than the electron curve, as expected with the lower ionization coefficient for holes in silicon.

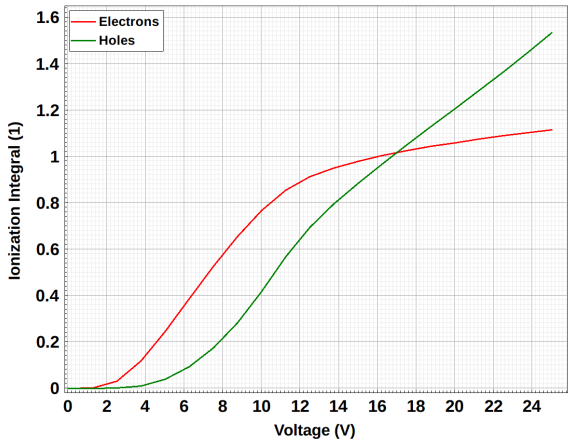


Figure 4.16. Ionization integrals of electrons and holes as a function of applied reverse voltage.

In TCAD simulations, four models are available to describe the impact ionization process in silicon. These models are largely based on the original Chynoweth formulation, where the ionization coefficient is modeled as a function of the electric field [19].

$$\alpha(E) = a \exp\left(-\frac{b}{E}\right) \tag{4.5}$$

where a and b are the coefficients used for calibration.

The four models studied here are Van Overstraeten–de Man model, the Okuto–Crowell model, the Lackner model, and the University of Bologna model. Each of these has been

implemented and compared in this thesis to identify the most suitable option for SPAD simulation.

van Overstraeten - de Man Model

This model is based on Chynoweth's equation and introduces parameter sets optimized for both low and high electric field regions [46]. It is valid for electric field values ranging between 1.75×10^5 V/cm and 6×10^5 V/cm. In addition to the exponential form, the model incorporates a scaling factor γ that accounts for optical phonon energy and introduces temperature dependence into the formulation. The ionization coefficient is expressed as

$$\alpha(E) = \gamma a \exp\left(\frac{-\gamma b}{E}\right) \tag{4.6}$$

Here, a and b are experimentally determined constants, and E represents the electric field strength.

Okuto-Crowell Model

This model is empirical and specifically calibrated for silicon at a reference temperature of 300 K. It is suitable for electric fields in the range of 1×10^5 V/cm and 1×10^6 V/cm [47]. The expression accounts for both field and temperature dependencies.

$$\alpha(F_{\text{ava}}) = a \cdot (1 + c(T - T_0))\, E^{\gamma} \exp\left[-(\frac{b\,(1 + d(T - T_0))}{E})^{\delta}\right] \tag{4.7}$$

By default, the values for this model's coefficients are fixed for silicon. However, it can be tuned to match specific experimental data by modifying its parameters.

Lackner Model

The Lackner model modifies the Van Overstraeten–de Man formulation by introducing a temperature dependence on the electric field in the exponential coefficient. It is designed to better describe ionization behavior over a wider range of high electric fields 10^5 V/cm and 10^6 V/cm [48]. This model is defined as:

$$\alpha(E) = \frac{\gamma a}{Z} \exp\left(\frac{-\gamma b}{E}\right) \tag{4.8}$$

where the additional Z parameter allows for finer control of the field-dependent response.

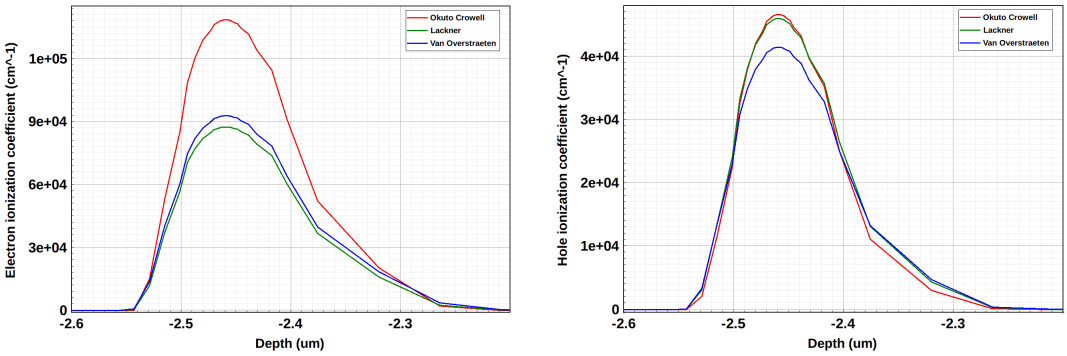
University of Bologna Model

This model is designed for a broad range of electric fields from 5×10^4 V/cm to 6×10^5 V/cm and temperatures up to 773 *K* [49] [50]. It is based on solutions to the Boltzmann transport equation and experimental validation, making it well-suited for devices operating under low to moderate electric fields such as SPADs with wide depletion regions or where temperature effects are important [51] [52]. The model is given as

$$\alpha(E,T) = \frac{E}{a(T) + b(T)\exp\left(\frac{d(T)}{E+c(T)}\right)} \tag{4.9}$$

This model offers precise predictions across a broad temperature and field range, making it ideal for devices affected by heating or large currents. So, this model will be skipped as it does not fit the necessities.

Figure 4.17 presents a comparison of the above mentioned impact ionization models, without the University of Bologna model, by plotting the electron and hole ionization coefficients as a function of depth at an overvoltage of 1 V. The results show that the Okuto-Crowell model, an empirical model, predicts higher ionization coefficients compared to the other two, as expected due to the high electric-field range. The Okuto–Crowell model is more suitable for simulating strong avalanche effects in SPAD devices.



(a) Electron ionization coefficient.

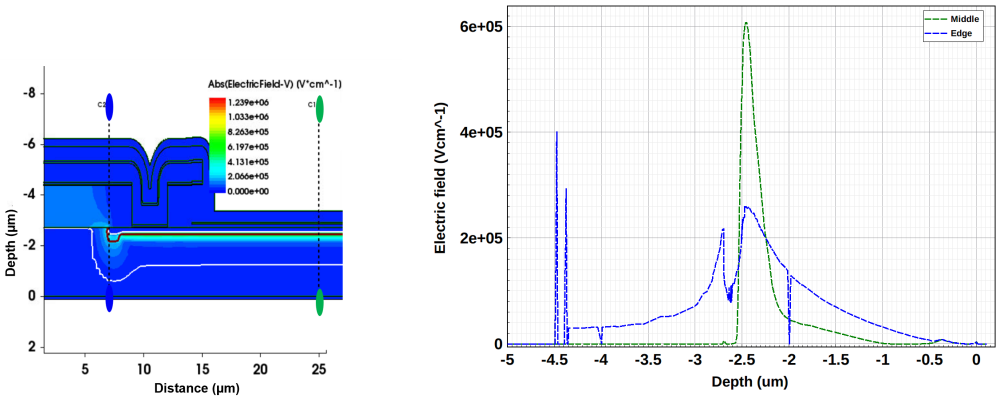
(b) Hole ionization coefficient.

Figure 4.17. Comparison of ionization coefficients across three different models at an overvoltage of 1 V: Van Overstraeten–de Man, Okuto–Crowell, and Lackner. The Okuto–Crowell model exhibits the highest peak value, indicating stronger avalanche behavior under high fields.

The electric field distribution using the Okuto-Crowell model along the two vertical cutlines, one through the center of the pixel and the other at the edge, is shown in Figure 4.18. This comparison helps to identify where the maximum electric field occurs. In this SPAD structure, virtual guard rings are implemented to reduce the electric field at the edges. This is important to prevent premature edge breakdown and ensure that avalanche multiplication starts only in the intended active region. As shown in the plot, the electric field is significantly higher in the middle of the pixel, reaching values in the critical electric field range. At the edges, the electric field is noticeably lower, confirming the effectiveness of the guard ring design.

The distribution of impact ionization across the SPAD structure is illustrated in Figure 4.19. The profile indicates that the ionization is primarily confined to the junction region and

does not extend deeply into the depletion zone. This behavior is expected because impact ionization is highly dependent on the local electric field, which is strong near the junction.



(a) Simulated electric field profile showing the position of cuts.

(b) Simulated electric field distribution in the vertical direction taken from the middle and the edge of the SPAD pixel.

Figure 4.18. Simulated electric field distribution of the SPAD pixel at an overvoltage of 1 *V*.

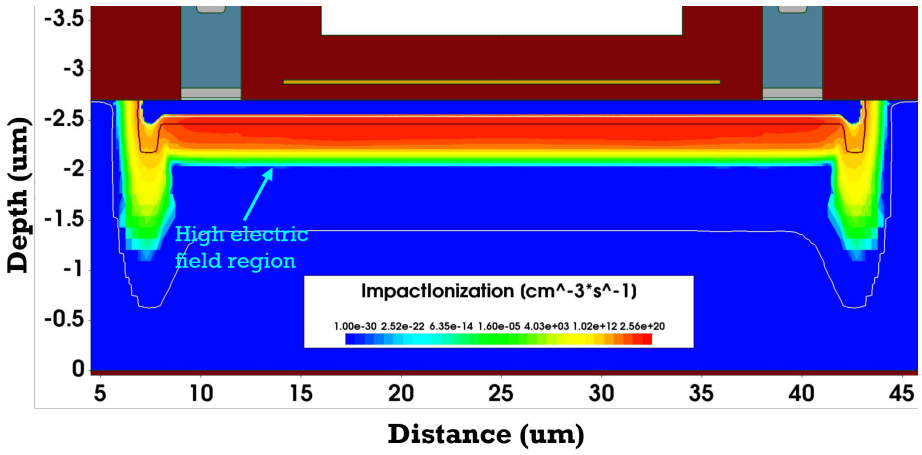


Figure 4.19. Distribution of total impact ionization rate across the SPAD cross-section at an overvoltage of 1 *V*.

Figure 4.20 presents the electron and hole ionization coefficient profiles extracted from a vertical cutline through the center of the active pixel area at an overvoltage of 1 *V*. The result shows that the electron ionization coefficient is consistently higher than that of holes. This is due to the lower effective mass of electrons in silicon. As a result, electrons undergo less scattering and can accelerate more efficiently under the influence of an electric field, making them more likely to gain enough energy to initiate impact ionization, leading to a higher probability of impact ionization compared to holes.

It can be clearly seen from the distribution that the impact ionization coefficients are high in the region where the electric field is strong. In areas where the electric field weakens, the ionization coefficients drop significantly, resulting in a negligible contribution to avalanche multiplication in those regions. The electron and hole ionization coefficients are used to calculate the breakdown triggering probability in the next section.

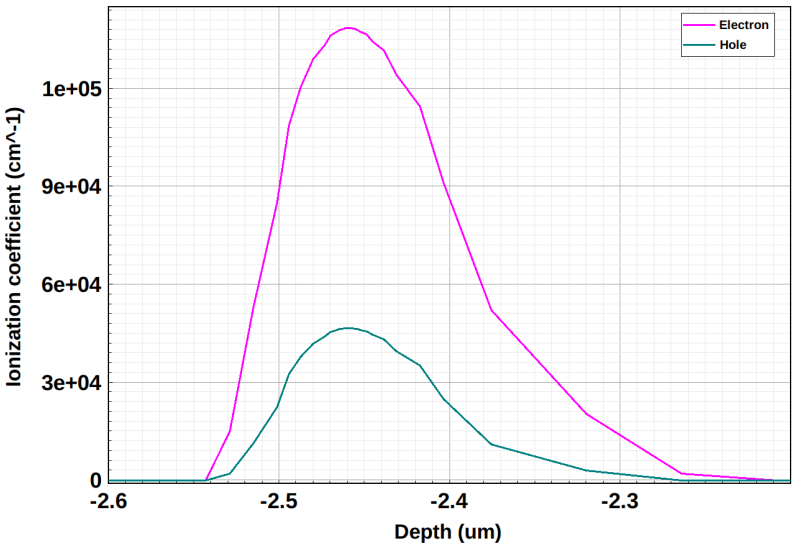


Figure 4.20. Distribution of ionization coefficients for electrons and holes as a function of depth taken from the center of the pixel at an overvoltage of 1 V.

4.2 Photon detection probability

The probability that a photon impinging on the active area triggers an avalanche is known as the photon detection probability (PDP). PDP is a key parameter that describes the detection of the probability of SPADs. Due to the statistical nature of impact ionization and photons' penetration into silicon, this probability is always less than one. It can be calculated as [53]:

$$\mathrm{PDP}(\lambda) = (1 - R(\lambda)) \int_{Z_p}^{Z_n} \alpha(\lambda) e^{-\alpha(\lambda) z} P_t(z) dz + \eta_p^e(\lambda, z_p) P_e(z_p) + \eta_p^h(\lambda, z_n) P_h(z_n) \tag{4.10}$$

In the above equation, there are three terms corresponding to three contributing sources of PDP. The first term considers the electron-hole pairs generated in the depletion region, where Z_n and Z_p are the depths of the respective depletion edges on the n- and p-sides. It calculates how deeply the photons penetrate into the SPAD and their chances of triggering an avalanche at different depths, calculated with the total breakdown probability P_t, triggered by electron-hole pairs photo-generated at z position. In light-matter interactions, the photons enter into the sensor, resulting in an exponential dependence of the absorption depth z and are

described by the Beer-Lambert law $\alpha \exp(-\alpha z)$ [15], where α is an absorption coefficient in the device material, a function of wavelength. *R* corresponds to the reflectivity, which comes from the reflection that happens due to the different refractive indices of the material that a light goes through and is calculated by using Fresnel equations. The details on reducing the reflection and having high absorbance that can contribute to a high photo generation rate are explained later in the next sections.

The second and third terms describe the contribution of those charge carriers that are generated in the neutral regions and, by diffusion, move to reach the depletion region, which has a very limited probability. The second term corresponds to the electrons that are generated in the p-type neutral region, where $P_e(z_p)$ is the electron breakdown probability triggered by electrons at z_p position and $\eta_{ph}^{e}(\lambda, z_p)$ is the electron quantum efficiency that tells the number of electrons generated in the p-type neutral region by one photon with wavelength λ and then diffuses to the p-side depletion edge, z_p, to trigger a breakdown event from there with $P_e(z_p)$.

Figure 4.21 shows the schematic of referencing the depletion edges with corresponding n- and p-type regions.

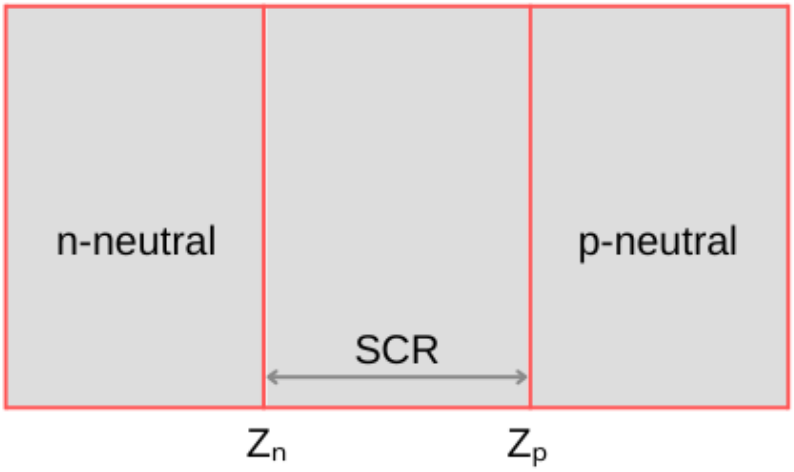


Figure 4.21. Schematic showing the depletion edges of corresponding n- and p-type regions.

Similarly, in the third term, η_{ph}^{h} is the hole quantum efficiency that calculates the number of holes that are generated in the n-type neutral region by one photon with wavelength, λ and then diffuse to the n-side depletion edge, z_n, to trigger a breakdown event from there with $P_h(z_n)$. The TCAD tools and their extracted parameters used to calculate PDP are shown in Figure 4.22, where the basic fabrication steps with geometry layout are done in the SProcess module and the electrical simulations in the SDevice module, as explained earlier in this chapter and Chapter 3. The extracted parameters used to calculate PDP are electron and hole ionization coefficients, α_e and α_h, which are imported into MATLAB to calculate the avalanche triggering probabilities, as there is no model in TCAD that could directly calcu-

late these probabilities. Other parameters, such as absorption coefficient, reflectivity, and electron or hole current densities used to calculate the electron or hole quantum efficiency, are extracted. All these extracted parameters are then imported into the SVisual module to plot the PDP as a function of wavelength.

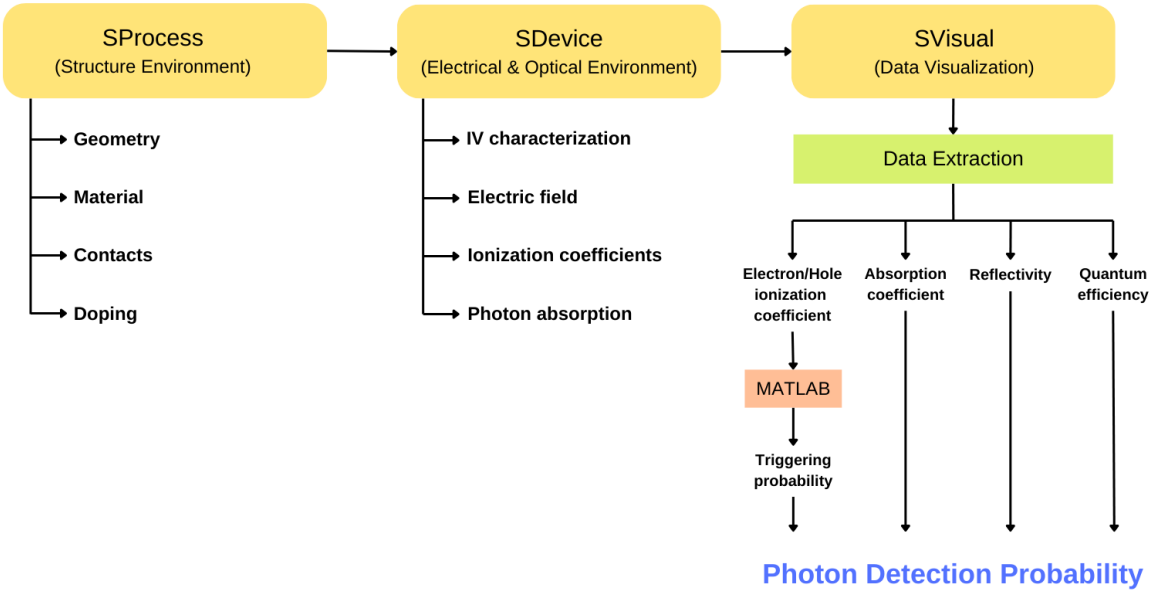


Figure 4.22. Flowchart to calculate photon detection probability (PDP) of SPAD.

4.2.1 Breakdown triggering probability

Avalanche breakdown triggering probability is defined as the probability under which the carriers (electrons and holes) trigger a breakdown process and initiate a self-sustaining avalanche. A photon can be absorbed either in the depletion layer or in the quasi-neutral regions above and below the depletion layer. When it is absorbed in the depletion region, both electrons and holes are separated and accelerate in opposite directions, which may initiate an avalanche due to the impact ionization process, but it does not mean that every interaction results in a self-sustaining avalanche. The probability that an electron or a hole generated at the depth z triggers a self-sustaining avalanche can be obtained by solving the following differential equations [54]:

$$\frac{\mathrm{d}P_e}{\mathrm{d}z} = -(1-P_e)\alpha_e[P_e + P_h - P_eP_h] \tag{4.11}$$

$$\frac{\mathrm{d}P_h}{\mathrm{d}z} = (1-P_h)\alpha_h[P_e + P_h - P_eP_h] \tag{4.12}$$

The terms $P_e(z)$ and $P_h(z)$ represent the electron and hole breakdown triggering probabilities, respectively. α_e and α_h are the impact ionization coefficients of electrons and holes dependent on the electric field, as explained in Section 4.1.4. This breakdown trigger probability depends on the location of photo-generation and the electric field distribution in the

device. The total triggering probability, $P_t(z)$, that either the electron or hole initiates an avalanche can be simply calculated as [54]:

$$P_t = P_e + P_h - P_e P_h. \tag{4.13}$$

To calculate these equations, MATLAB is used with the BVP4C solver to get the solution of the differential equations. The used boundary conditions are $P_e(z_n) = 0$ and $P_h(z_p) = 0$.

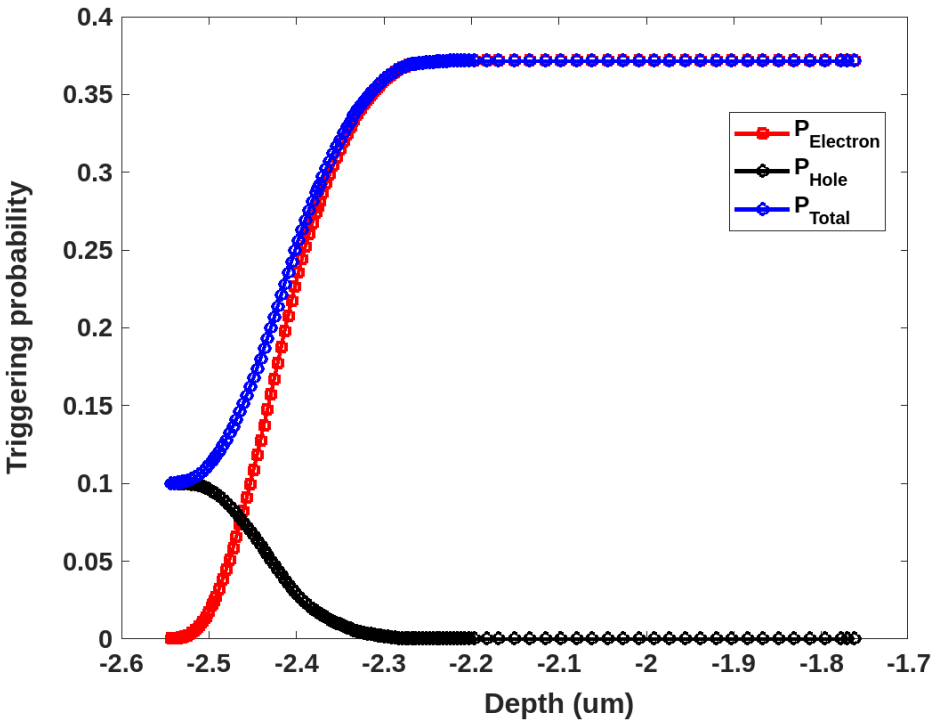


Figure 4.23. Breakdown triggering probability at an overvoltage of 1 V.

The electron, hole, and total breakdown triggering probability are calculated only for the overvoltage of 1 V. The electrical equations used to simulate the device does not converge at higher overvoltages. The result is shown in Figure 4.23, where it is clear that P_h and P_e satisfy the boundary conditions as expected. Notably, the maximum value of P_h is smaller than the maximum of P_e due to electrons having a higher impact ionization coefficient than holes (Section 4.1.4). The total trigger probability P_t is 0.1 at the n-side depletion edge and rises to about 0.38 at the p-side depletion edge. For electrons, P_e increases with z from zero to a maximum value generated inside the multiplication region. The maximum probability of electrons occurs at the edge of the p-side depletion region.

4.2.2 Simulated optical absorption and generation

In this section, the focus is on how incident photons interact with the various material layers of the device due to each material's unique optical properties, such as refractive indices.

These properties significantly affect how light is handled, whether it is absorbed, reflected, or transmitted through the device. Figure 4.24 shows the order of five different materials from the top of the simulated structure: silicon dioxide (SiO_2) for metal isolation, silicon nitride ($Si_3\ N_4$) as an anti-reflective coating (ARC) layer, silicon dioxide (SiO_2) as a screening oxide, silicon as a substrate and buried oxide at the bottom having different refractive indices.

The optical behavior of these layers is quantitatively analyzed using the Transfer Matrix Method (TMM) available in the TCAD SDevice module [38], explained in Chapter 3. It provides a systematic approach by using matrices to calculate how light is reflected, transmitted and absorbed at the interfaces and within the layers of a multilayer structure. The reflection, transmission and absorption coefficients are related as

$$R + T + A = 1 \tag{4.14}$$

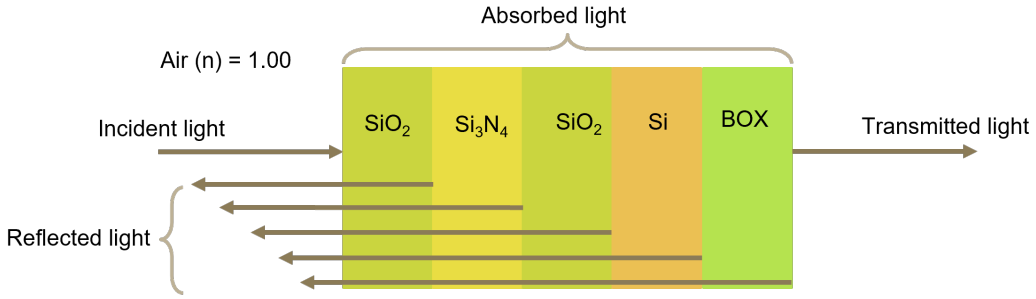


Figure 4.24. Multi-layered structure in the design.

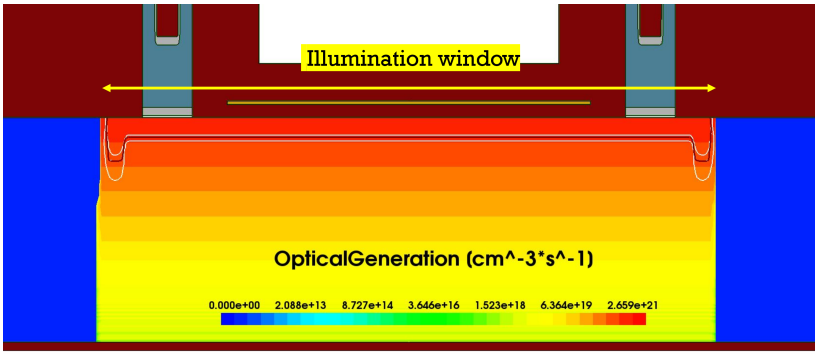


Figure 4.25. An illumination window with the color map of optical generation.

An illumination window is defined as the area where light falls onto the device, as depicted in Figure 4.25. The layer stack extraction method extracts all grid elements along a line normal to the corresponding illumination window, starting from a specified position within the window.

The absorption of photons follows an exponentially decaying distribution, which is simulated and plotted in Figure 4.26 and used for the calculation of PDP in the wavelength range of 0.4 μm to 0.72 μm.

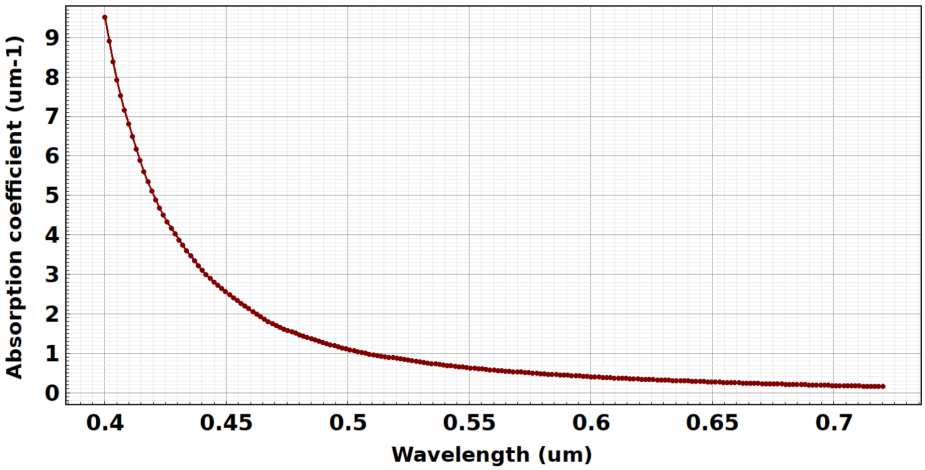


Figure 4.26. Simulated absorption coefficient as a function of wavelength.

An ARC layer $Si_3\ N_4$ is used, which helps to minimize reflections, thereby increasing the amount of light that enters the SPAD. When an incident wave first hits the ARC layer, some of it reflects back. This reflected wave, referred to as wave A, undergoes a 180° phase change, an external reflection. The light that enters and travels in the ARC then becomes reflected at the ARC-semiconductor surface is wave B, shown as an example in Figure 4.27, also suffers a 180° phase change since $n_3 > n_2$. d is the thickness of the ARC. When wave B reaches wave A, it has suffered a total delay corresponding to traversing the thickness d of the coating twice. The phase difference is $k_c(2d)$, where k_c is a propagation constant equals to $2\pi n_2/\lambda$. The phase difference between A and B then becomes $(2\pi n_2/\lambda)$ (2d). With the destructive interference of A and B, the reflected light can be reduced, and the phase difference then becomes odd multiples of π [55].

$$\left(\frac{2\pi n_2}{\lambda}\right) 2d = m\pi \quad \text{or} \quad d = m\left(\frac{\lambda}{4n_2}\right) \tag{4.15}$$

Thus, the thickness of the ARC must be the odd multiples of the quarter wavelength in the coating and depends on the wavelength. To calculate the higher absorbance and less reflectivity, the thickness of the ARC layer can be tuned. The thickness used in the simulated structure is ~ 40 *nm* for the aiming wavelength of around 0.41 μm. The simulated reflectivity is then plotted in Figure 4.28, where the lowest reflection can be seen at 0.435 μm, meaning more absorption is happening at 0.435 μm. Also, few ripples shows the light reflection between different layers, so this model is very close to the results that can be observed in the measurements.

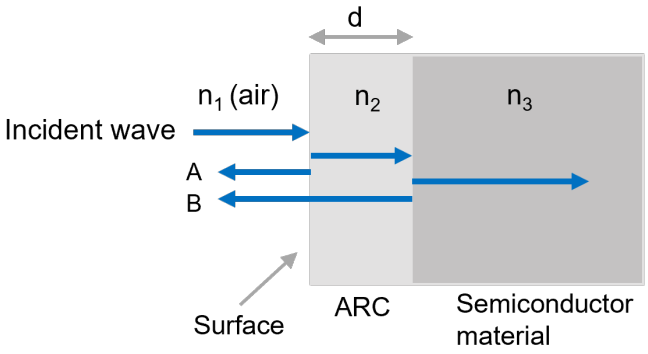


Figure 4.27. An example showing the thickness of ARC and reflections happening from different layers of a structure.

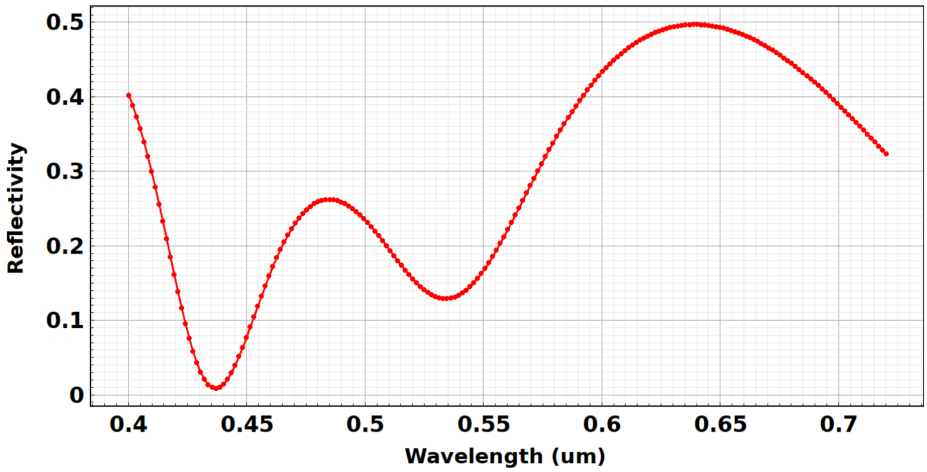


Figure 4.28. Simulated reflectivity as a function of wavelength with the lowest reflection at the wavelength of 0.4 μm.

Quantum efficiency of an electron or hole can be calculated as [41]:

$$\eta_{ph}^{e,h} = \frac{hc}{q\lambda}\frac{J_{e,h}}{P_0} \tag{4.16}$$

Here, $J_{e,h}$ is the photon current density of electrons or holes and P_0 is a power density.

For each ramping wavelength, the photons are illuminated with a constant power density P_0 of 0.05 $W/\ cm^2$ on the device. The distribution of photogenerated current densities of electrons and holes are simulated with a steady-state solution to calculate the quantum efficiencies of electrons and holes. As voltage dependence of PDP only comes from the breakdown triggering probabilities, therefore, no change is observed in the distributions at different bias voltages. The simulation of these distributions are done at 12 V of bias voltage. Figure 4.29 and 4.30 show the 2D electron and hole current density distribution at a wavelength of 0.45 μm. The distributions show a lower current density at the center at the

depth away from the depletion region which is due to the electrical contacts located on the sides. However, in the calculation of quantum efficiency, the densities are extracted from the depletion edges of electrons and holes.

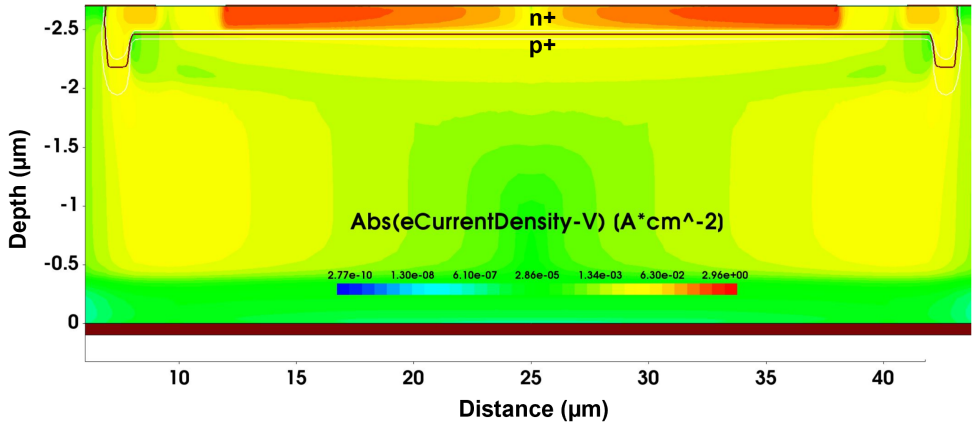


Figure 4.29. Simulated 2D electron current density distribution.

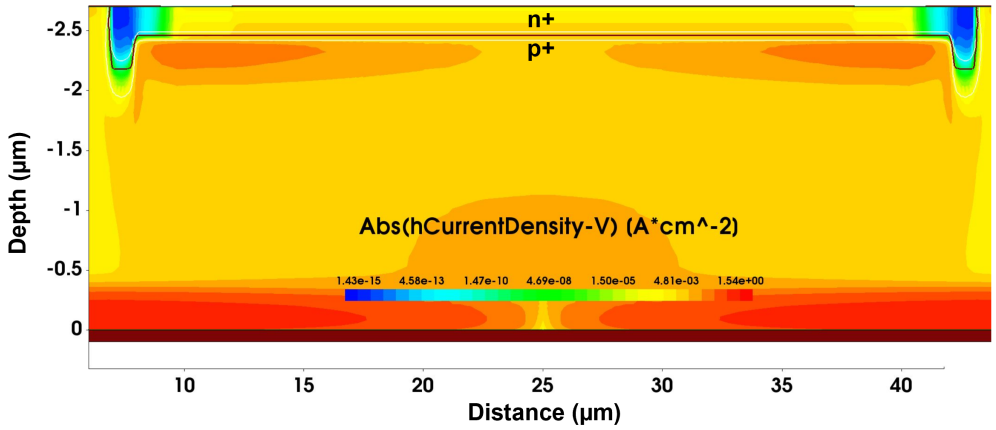


Figure 4.30. Simulated 2D hole current density distribution.

Figure 4.31 shows the 1D cross-sectional profiles in depth at the device center of total, electron, and hole distributions, highlighted with the depletion region. The electron current density diffuses from the lower p-type neutral region is the value at the p-side depletion edge (z_p) around 2.41 μm, marked by a green circle. Conversely, the hole current density diffuses from the upper n-type neutral region is the value at the n-side depletion edge (z_n) around 2.48 μm, indicated by a blue circle.

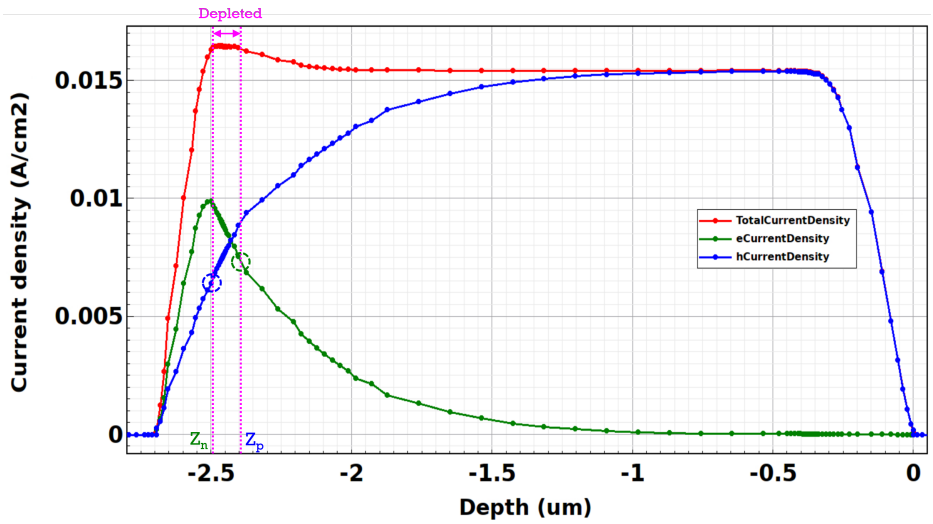
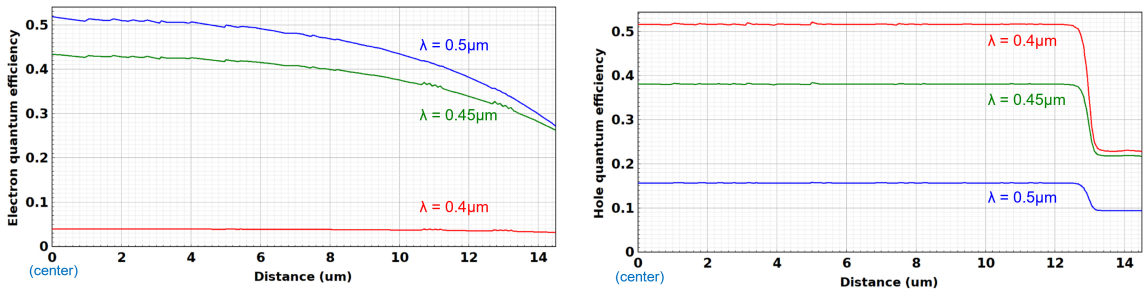


Figure 4.31. 1D profile of total, electron and hole current densities as a function of depth at the center of the device at the wavelength of 0.45 μm.

From these extractions at the depletion edges and solving Equation 4.16, the electron and hole quantum efficiencies, η_{ph}^{e} and η_{ph}^{h}, are calculated and presented in Figure 4.32 across distances taken from the center of the device for wavelengths of 0.4, 0.45, and 0.5μm. It is observed that at 0.4 μm, the hole quantum efficiency η_{ph}^{h} reaches 0.51 due to the high absorption coefficient of *Si* at this wavelength, $\sim 9.52 \times 10^{4}\ cm^{-1}$, that gives an absorption depth of 0.1 μm. Consequently, most photons at 0.4 μm are absorbed within the n-type region extending up to 2.48 μm of depth. In contrast, the electron quantum efficiency η_{ph}^{e} for 0.4 μm wavelength is significantly lower as 0.03, reflecting that only very few photons penetrate to the deeper p-type neutral region beyond 2.41 μm of depth.



(a) Electron quantum efficiency. **(b)** Hole quantum efficiency.

Figure 4.32. Quantum efficiency of electrons and holes as a function of the distance taken from the center of the structure under 0.4, 0.45 and 0.5 μm illumination.

At the wavelength of 0.45 μm with an absorption coefficient of $\sim 2.55 \times 10^{4}\ cm^{-1}$ with a corresponding absorption depth of 0.3 μm, the electron and hole quantum efficiencies are close to 0.4. Moving to the higher wavelength simulated for this observation at 0.5 μm

having an absorption coefficient of $\sim 1.11 \times 10^4\ cm^{-1}$ with an absorption depth of 0.9 μm, the electron quantum efficiency is 0.52, whereas the hole quantum efficiency comes out to be 0.16, showing that photons penetrate deeply with the increase in wavelength.

The quantum efficiencies are plotted from the middle of the device. For both electrons and holes, the quantum efficiencies remain constant until the edge effect starts, and they slowly start to decline when reaching the edge of the device. It is important to consider that the TCAD simulations might overestimate the collection efficiency in the neutral regions due to defect densities generated in the process simulation module than those found in fabrication processes.

4.2.3 PDP dependence on wavelength

After extracting and calculating the parameters according to Equation 4.10, such as electron and hole quantum efficiencies η^e_{ph} and η^h_{ph} at all wavelengths, the reflectivity spectra, and the breakdown trigger probability functions, the PDP spectra at a fixed overvoltage can be calculated in the SVisual module. Figure 4.33 presents the simulated PDP as a function of wavelength, and the maximum PDP reaches to peak value at 0.403 μm of wavelength, about 44 % at an overvoltage of 1 V.

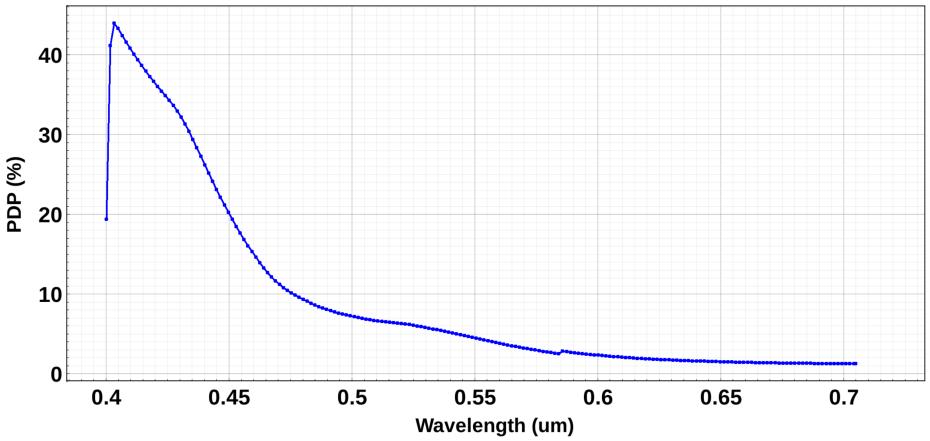


Figure 4.33. Simulated photo detection probability as a function of wavelength at a fixed overvoltage of 1V.

For the longer wavelengths (>500 *nm*) in the visible or near-infrared spectral range, the sensitivity is degraded as the absorption depth of this range is higher than the SPAD active area.

Chapter 5

SPAD measurement studies and analysis

This chapter provides a detailed layout of the first production of SPAD matrices produced using Silicon-On-Insulator (SOI) technology by Fraunhofer EMFT. Both static and dynamic measurement studies, including forward and reverse current-voltage (IV) characteristics with the extraction of breakdown voltage and waveform analysis at different temperatures, are covered. The findings from the analyses of these measurement studies suggest the necessity for additional tests, such as Light Emission Test (LET) and Secondary Ion Mass Spectroscopy (SIMS), to gain a more comprehensive understanding of the performance parameters. The chapter concludes with a discussion of the results by comparing them with the simulations and analysing the mismatch parameter profiles observed from the comparison.

5.1 Sensor layout

The first SPAD sensor is fabricated on 160 *nm* SOI wafer, with a total of 336 reticles on 12-inch wafer. The total size of the reticle is 9.6×9.6 mm^2. The design features an 8×7 die arrangement per reticle, resulting in 56 individual dies per reticle. The GDS mask layout of an individual reticle, illustrated in Figure 5.1, reveals strategic division into three blocks featuring wire-bonded, flip-chips (non-wire bonded) and test structures. The wire-bonded sensors incorporate a variety of configurations, including both single and multi-pixel arrangements. Each pixel configuration is fabricated in four different guard ring variations. In addition to the SPAD sensors, dedicated test structures crucial to understand and validating the step-by-step doping process are implemented as well. The test structures include the variations in the guard ring and the distances between n+ doping, doped on top of p+ to make an n-p junction and p-contact, used for negative biasing (see SPAD design layout from Chapter 4).

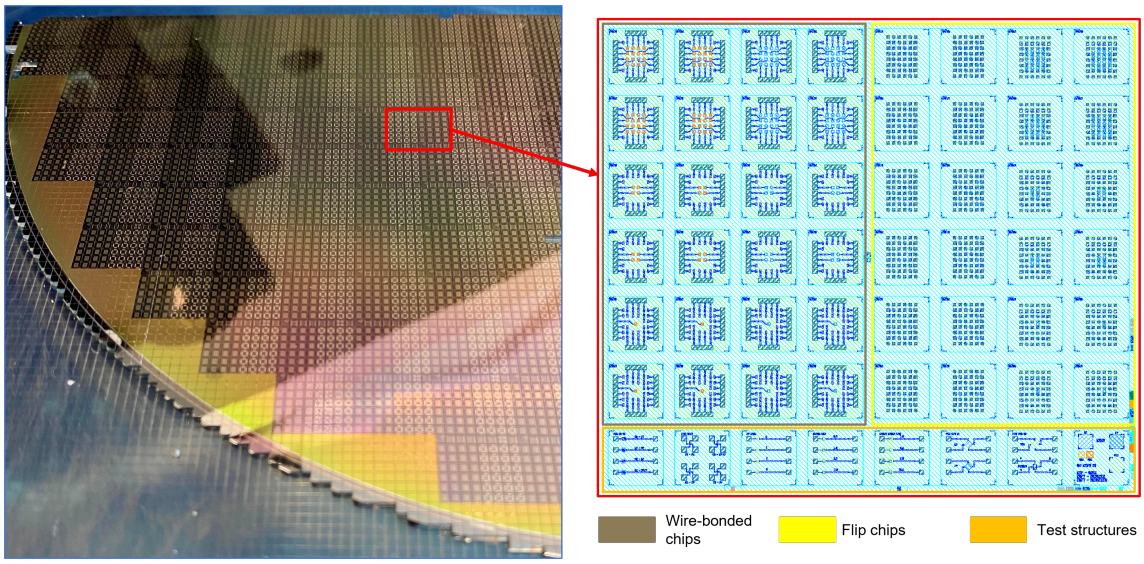


Figure 5.1. Left shows the processed sensors glued on blue tape, highlighting one reticle with a red color, and right shows the GDS mask layout of the reticle, where the highlighted top left, top right, and bottom blocks show the wire-bonded chips, flip chips, and some test structures, respectively.

The reticles are designed to accommodate various needs through pixel configurations of 1×1, 2×2 and 4×4 pixel arrays, with and without an entrance window with four different guard ring widths of 0.5 μm, 1 μm, 1.5 μm, and 2 μm. Figure 5.2 shows the microscopic layout of the square-shaped sensors, highlighting the arrangement of 1×1 and 4×4 pixel designs in the middle. Each pixel has dimensions of 50×50 μm^2. The outer square structures are the bonding pads, used to bond the sensor with wires onto the PCB.

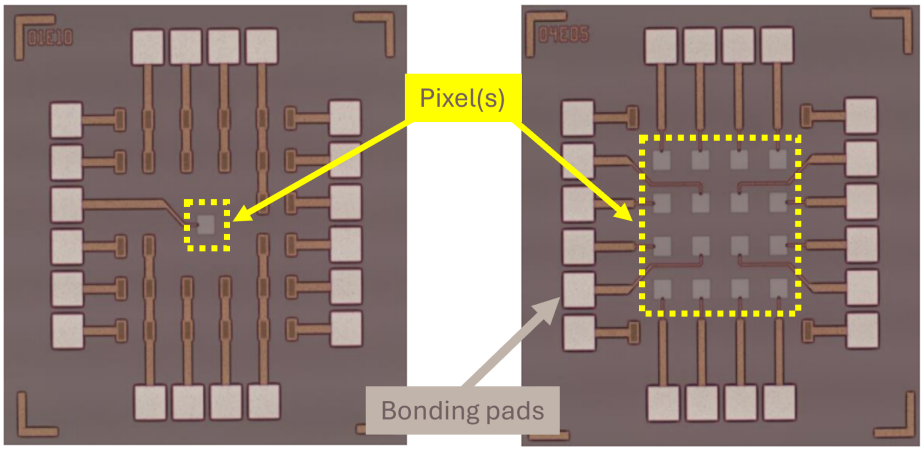
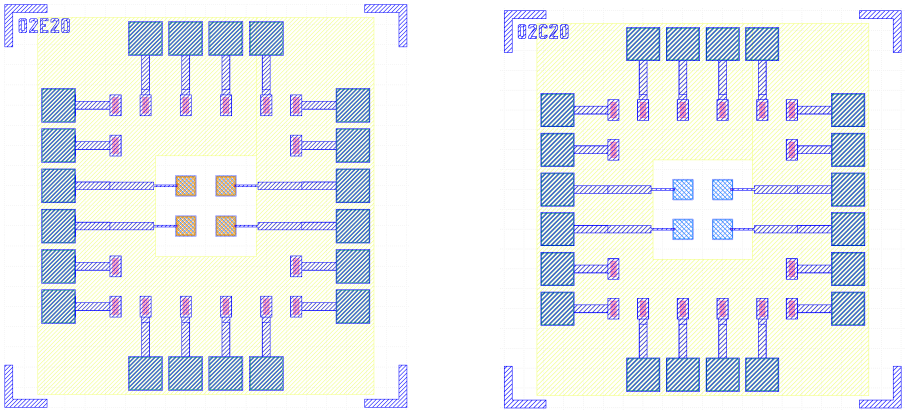


Figure 5.2. Microscopic layout of the sensors. Left shows 1×1 and right shows 4×4 pixel variation with pixel(s) positioned in the middle and surrounded by bonding pads.

Each sensor on the wafer is labeled with an identifier that provides details of its configuration. For instance, the identifier ”01E20” denotes a 1×1 pixel configuration (01), the

presence of an entrance window (E), and a guard ring value of 2.0 μm (20). This labeling facilitates easy identification and alignment of sensors during characterization.

The design variation between pixel configurations, such as the 2×2 pixels with and without an entrance window, is shown in Figure 5.3. The presence of an entrance window, highlighted by an orange region in the middle, is optimized for light capture and enhanced sensitivity, thanks to the deposition of an ARC layer. Conversely, the configuration without an entrance window prevents light from entering due to the absence of an ARC layer.



(a) 2×2 pixels with an entrance window.

(b) 2×2 pixels without an entrance window.

Figure 5.3. GDS mask structures of 2×2 pixels with variation in an entrance window.

5.2 Lab measurements

This section focuses on the static and dynamic characterization of different SPADs in the laboratory. The results are analyzed and discussed.

5.2.1 Static characterization

From the simulation results discussed in Chapter 4, breakdown voltage, an important parameter for the characterization, comes out to be around 15 V. Now, the laboratory measurements for DC characterization of the fabricated sensors involve measuring forward and reverse current-voltage (IV) characteristics at room temperature to observe the behaviour of breakdown voltage. The measurements are done using a Keithley 6487 pico-ammeter, which provides the bias voltage and current measurements for these tests. The setup, illustrated in Figure 5.4, includes a dark aluminum box to shield the sensor from ambient light. A PCB is designed to accommodate both single- and multi-pixel configurations, which allows for easy switching between different sensor types by plugging them in using an edge card connector. The sensor is covered with a cap for optical insulation.

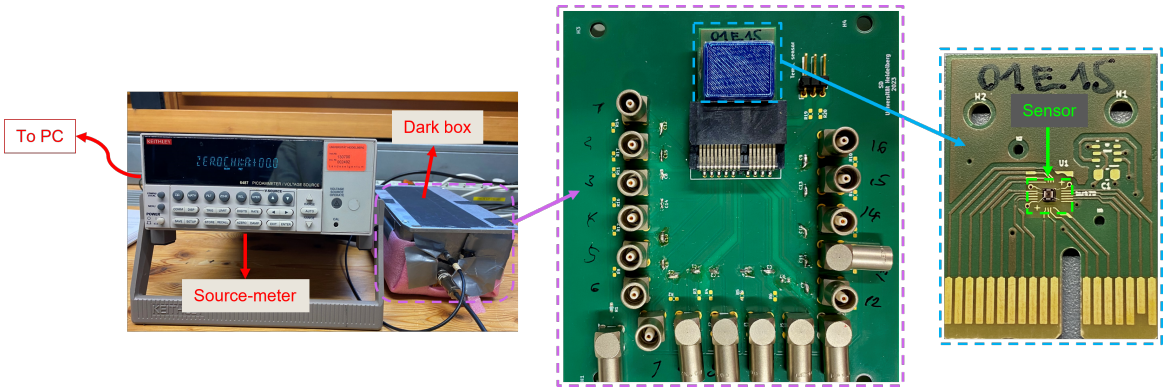


Figure 5.4. IV-curve characterization setup: a sourcemeter and a dark aluminum box in which the PCB with wire-bonded sensor is placed.

The entire assembly of the PCB is then put in a box, ensuring that no light from outside generates any extra current. The box is connected to the ”shield pin” of the Keithley 6487 to minimize electronic noise, expecting a small measurement uncertainty for ultra-low current levels. The setup is controlled via a PC, which automates the data acquisition process, allowing for efficient and reproducible testing.

Forward IV characterization

The forward-biased IV curve of the sensor is shown in Figure 5.5. The characteristic curve is consistently observed across four variations of the sensor, labeled as *01E05*, *01E20*, *02E05*, and *02E20* at room temperature. The voltage is gradually increased in steps of -0.1 V up to a maximum of -1 V to avoid potential damage from higher voltages. The negative voltage is applied to the n-side of the p-n junction corresponding to the forward bias configuration with a zero series resistor. Initially, as the bias voltage V_{bi} increases and is below the knee voltage of -0.6 V, the rise in current is minimal. As bias voltage V_{bi} surpasses -0.6 V, the current starts to increase rapidly.

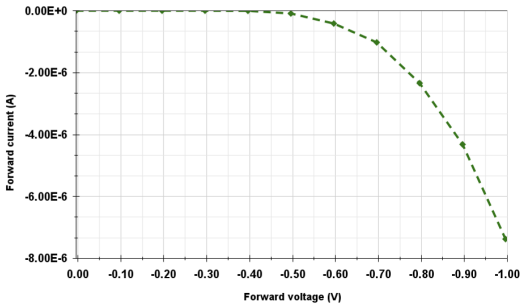
The relationship between the current I and the bias voltage V_{bi} in the forward-biased condition of a p-n junction is effectively modeled by the Shockley diode equation:

$$I \approx I_s \exp\left(\frac{qV_{bi}}{\eta kT}\right), \tag{5.1}$$

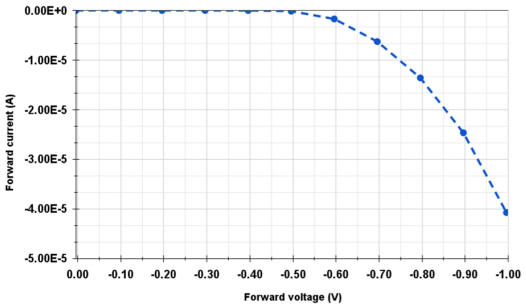
where I_s is the saturation current, q is the elementary charge, k is the Boltzmann’s constant, and η is an ideality factor of the diode.

The ideality factor, η, derived from the slope of the IV curve, informs about the internal conduction mechanisms in the p-n junction. It varies between 1 and 2, according to the main conduction phenomenon in the diode. η equals 1 when the current is due to the minority carrier diffusion in the neutral region, and 2 when it is due to the recombination in the space

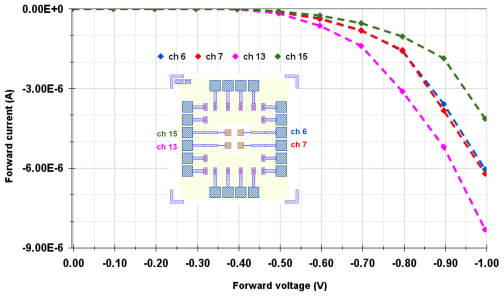
charge region. From the slopes extracted from Figure 5.5, η comes out to be between 1 and 2, indicating both processes play an important role.



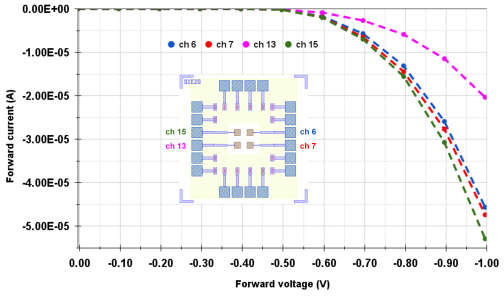
(a) Type 01E05.



(b) Type 01E20.



(c) Type 02E05 with a schematic detailing channel layout.



(d) Type 02E20 with a schematic detailing channel layout.

Figure 5.5. Measured forward characteristics of SPAD: (a) 1×1 pixel with 0.5 μm of guard ring (b) 1×1 pixel with 2.0 μm of guard ring (c) $2\times$ 2 pixels with 0.5 μm of guard ring (d) 2×2 pixel with 2.0 μm of guard ring.

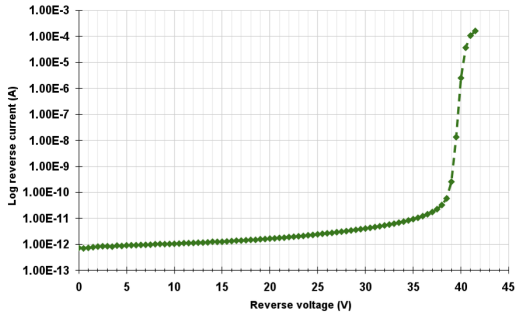
Reverse IV characterization

In the reverse IV characteristic curve of SPAD sensors, the current that flows when the sensor is not exposed to light depends mainly on the free carriers available. The dominant mechanism in the semiconductor material is the Shockley-Read-Hall (SRH) effect, which significantly contributes to the dark current [56]. It describes the generation of electron-hole pairs due to the trapping effect of impurities in the lattice as well as band-to-band tunneling effects.

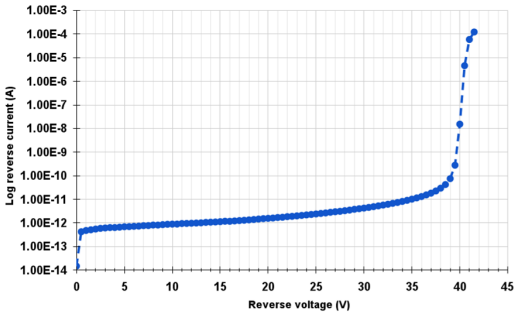
For measuring the IV curve, depicted in Figure 5.6, the bias voltage is incrementally applied from 0 to 41.5V in a step of 0.5V, as exceeding this voltage can damage the sensor. Current measurements are displayed on a logarithmic scale to clearly show the changes over a wide range of values. In the graph, two distinct zones are identified, corresponding to different operational regimes, termed pre-breakdown and post-breakdown zones.

1. Pre-breakdown phase (Linear regime): In this phase, the bias voltage is below the breakdown voltage. The current increases slightly with an increase in bias voltage. This ”dark current” is mainly due to the surface currents and the bulk dark current due to the movement of free carriers.

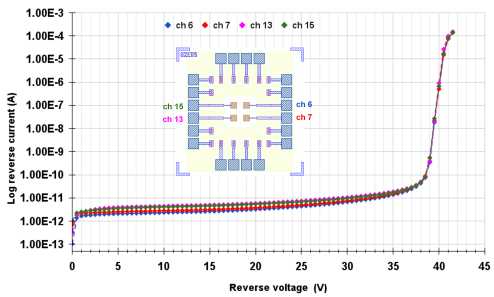
2. Post-breakdown phase (Geiger regime): Once the bias voltage exceeds the breakdown voltage, the current rises dramatically. This trend is due to the avalanche created by the free carriers generated by the process of impact ionization. These free carriers, which usually trigger an avalanche, are created due to thermal SRH generation enhanced by tunneling.



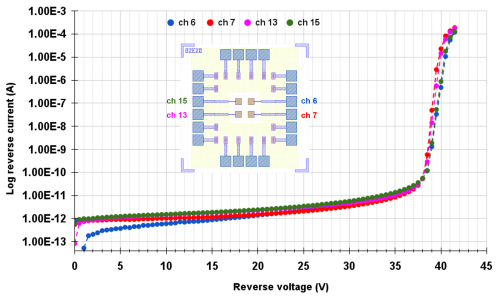
(a) Type 01E05.



(b) Type 01E20.



(c) Type 02E05 with a schematic detailing channel layout.



(d) Type 02E20 with a schematic detailing channel layout.

Figure 5.6. Measured reverse characteristics of SPAD: (a) 1×1 pixel with 0.5 μm of guard ring (b) 1×1 pixel with 2.0 μm of guard ring (c) $2\times$ 2 pixels with 0.5 μm of guard ring (d) 2×2 pixel with 2.0 μm of guard ring.

The breakdown voltage comes out to be around 40 V for all sensor variations tested, regardless of differences in the values of guard rings. This voltage represents the point where the electric field within the sensor's depletion region is strong enough to trigger an avalanche from any free carrier present, whether originating from absorbed photons or the thermal generation process. Interestingly, despite variations in the thickness of guard rings, which are expected to have a small influence on the breakdown voltage, all sensors show

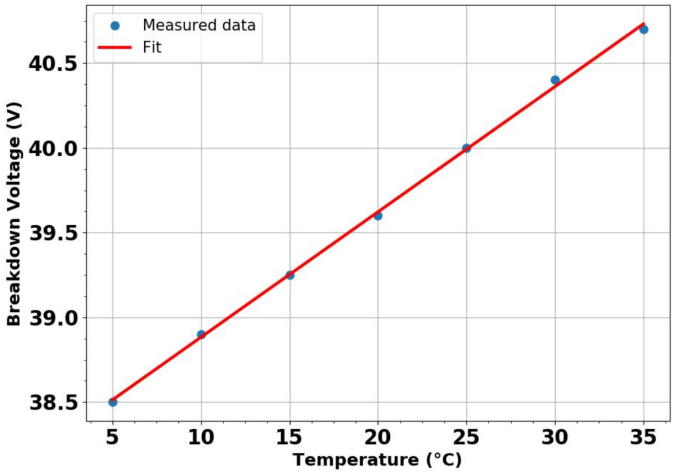


Figure 5.7. Measured breakdown voltage at different temperatures.

similar breakdown characteristics. Even in multi-pixel sensors, each channel displays nearly identical IV behavior. The fabricated sensors reveals a huge difference from the simulated breakdown voltage, a sign of the inconsistencies in the fabricated sensors.

The reverse IV characteristics are measured at different temperatures and is plotted in Figure 5.7. The sensor is placed in a temperature-controlled cabinet and the measurements are taken from 5°C to 35°C. The measured data is fitted on a linear scale, and the temperature coefficient, $\frac{dV_{br}}{dT}$, comes out to be 73.9 mV/°C, a positive coefficient, as the free path of charge carriers decreases with the rising temperature because of the excitation of phonons.

Breakdown voltage analysis

The breakdown voltage is a critical parameter that indicates the transition from the pre-breakdown to the post-breakdown regime in SPAD. Understanding this voltage is essential as it determines when the sensor starts to undergo avalanche breakdown.

Breakdown voltage is traditionally identified using the logarithmic plot of the IV curve, where a sharp rise in current indicates the onset of avalanche breakdown. However, a more precise method involves analyzing the DC analysis of the third derivative of the current with respect to voltage. This method focuses on subtle changes around the breakdown point on the IV curve, allowing for the identification of both turn-on and turn-off voltages, which are critical for understanding the dynamics of avalanche initiation and quenching [57].

Above the breakdown voltage, the current becomes proportional to both the charge produced and the avalanche triggering probability. Since the charge is proportional to bias voltage V, and the triggering probability is also a linear function of V, the current above breakdown ($V > V_{bd}$) follows a quadratic relation.

$$I \propto q(V)\, P_{\mathrm{t}}(V) \propto (V - V_{\mathrm{bd}})(V - V_{\mathrm{bd}}) \tag{5.2}$$

Here, a single breakdown voltage is assumed. However, in practice, the avalanche in SPADs exhibits hysteresis with two distinct voltages: a turn-on voltage (V_{01}) when the avalanche is initiated and a turn-off voltage (V_{10}) when the avalanche fades away. This hysteresis behavior originates from the difference in how the avalanche process starts and stops during a voltage sweep. During a forward voltage sweep (from low voltage to high voltage), the avalanche initiates when the applied voltage reaches the turn-on voltage V_{01} and the current starts rising rapidly. During a reverse sweep (from high voltage to low voltage), the process becomes slow, and there is a difference in the breakdown point, known as the turn-off voltage V_{10} [58].

The hysteresis can be taken into account as a voltage shift between the q(V) and P_t (V) functions. The charge becomes zero below V_{10} and the triggering probability emerges from zero at V_{01}, so the current of SPAD will be a shifted quadratic function of V for $V > V_{01}$.

$$I \propto q(V)\, P_{\mathrm{t}}(V) \propto (V - V_{10})(V - V_{01}) \tag{5.3}$$

and the avalanche voltage hysteresis is given by:

$$h = V_{01} - V_{10}. \tag{5.4}$$

To extract these voltages from experimental data, the third derivative of the reverse IV curve is analyzed. The breakdown phenomenon by distinct avalanche turn-on V_{01} and turn-off V_{10} voltages, predicting that the third derivative of the reverse current curve follows [57]:

$$\frac{\mathrm{d}^3 I}{\mathrm{d}V^3} = A_{\mathrm{3rd}} \cdot \left[2 - \frac{h}{\sigma^2}(V - V_{01})\right] \cdot e^{-\frac{(V-V_{01})^2}{2\sigma^2}} \tag{5.5}$$

where A_{3rd} is the amplitude of the third derivative, V_{01} is the turn-on voltage, *h* represents the hysteresis between turn-on and turn-off voltage (h = V_{01} - V_{10}), and σ is the standard deviation.

Figure 5.8 displays the third derivative of the measured reverse IV curve alongside its model fitting from 33 V to 41.5 V with Equation 5.5. The curve exhibits a bipolar shape, which is characteristic of the transition between the turn-on and turn-off phases. The analysis indicates a turn-on voltage (V_{01}) of 40.76 V. With a hysteresis (h) of 1.69 V, the turn-off voltage (V_{10}) is calculated to be 39.07 V with a σ of 0.5. It is noted that further improvement in measurement resolution, such as using finer voltage steps, can enhance the accuracy of the extracted breakdown parameters.

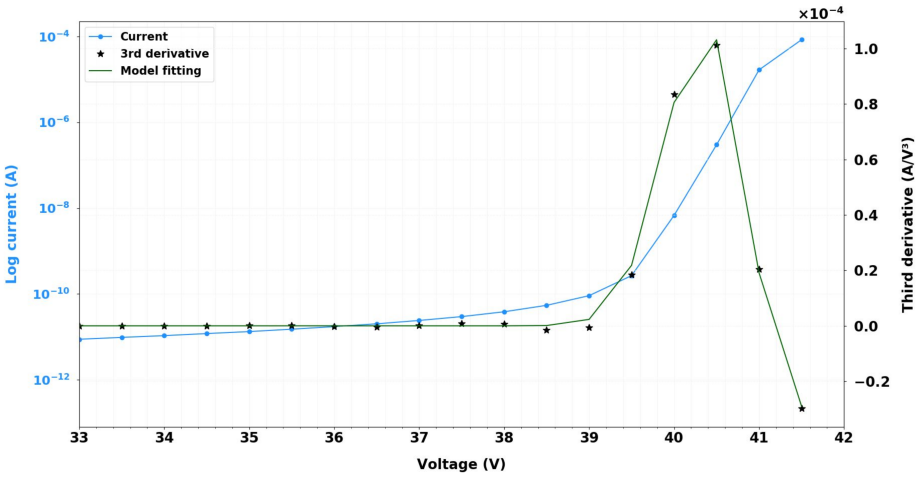


Figure 5.8. Third derivative of the reverse IV curve modeled from 33 V to 41.5 V and its fitting result for type 01E20.

5.2.2 Dynamic characterization

Passive quenching circuit for readout

When SPAD is set above its breakdown voltage, it becomes sensitive enough to detect single photons. Each photon that hits the device can start an avalanche, and to prevent damage to the device and to make sure it can detect another photon, this avalanche must be stopped and the device must be reset to its initial state. This is done through two processes: quenching and recharging. A quenching mechanism is required to limit the avalanche current and to stop the breakdown process, using a passive quenching circuit. In the passive quenching circuits, the discharge is stopped by a voltage drop at a high-value quench resistor, placed in series with the diode, caused by the avalanche current. As the diode voltage is reduced to approximately breakdown voltage, and since the avalanche is a statistical process, there is a certain probability that none of the carriers crossing the junction may impact ionize and the avalanche is quenched. After quenching, the device goes into the recharging phase. During this time, the voltage across the diode is gradually increased to its original level above the breakdown voltage. This period is known as the recovery or dead time. These steps are crucial to ensure that SPAD devices can operate, detecting events repeatedly without getting damaged due to the high currents of avalanches.

Figure 5.9 depicts the passive equivalent circuit of a SPAD, which is reverse-biased using a quenching resistor R_q = 2 kΩ. This includes a junction capacitance C_d typically around 1 pF, and stray capacitance C_s, which is the capacitance to the ground of the diode terminal connected to R_q and is usually a few picofarads. Additionally, R_s, a low-value series resistor on the ground lead is set to 1kΩ to facilitate current measurement.

The resistance R_d representing the diode's resistance is given by the series of space-

charge resistance of the avalanche junction and the ohmic resistance of the neutral semiconductor through which the current passes. When a photon is absorbed by the SPAD, it closes the switch in the circuit, initiating an avalanche. This causes the current to swiftly rise to a macroscopic steady level. The avalanche current discharges the capacitances which effectively quenches the avalanche, and corresponds to opening the switch in the diode, allowing the capacitances to slowly recharge. As a result, the voltage across the diode exponentially recovers toward the bias voltage, preparing the device for the next detection event. This setup helps the SPAD quickly reset after detecting an event, allowing it to work efficiently and continuously [59].

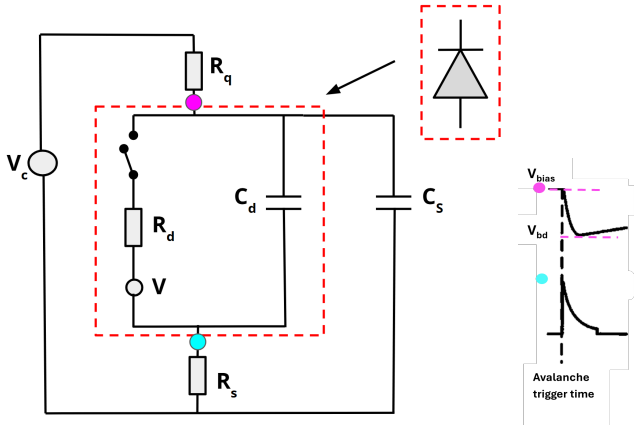


Figure 5.9. Passive quenching circuit for the current-mode output configuration of the SPAD with expected waveform output.

The output is captured from two specific points, reflecting the voltage output from node R_q and current output from node R_s waveform, shown in Figure 5.9. When the quenching happens, the bias voltage, V_{bi}, drops to breakdown voltage V_{bd}, and the time it takes is referred to as the quenching time (T_q), which can be calculated by

$$T_q = (C_d + Cs)\frac{R_d R_q}{R_d + R_q} \cong (C_d + C_s)R_d \tag{5.6}$$

Here, R_d and R_q act in parallel. Following the quenching, the circuit requires a recovery period to recharge the capacitances, preparing for the next detection cycle. This recovery time (T_r) is given by

$$T_r = R_q(C_d + C_s) \tag{5.7}$$

The current mode output taken through R_s on the ground lead shows the waveform of the diode current directly. It is crucial that the stray capacitance C_s must be comparable to or greater than the diode capacitance C_d for a significant voltage pulse to appear across R_s. Otherwise, only a small fraction of the avalanche current will pass through R_s; only the

current that discharges C_s flows in the loop including R_s, whereas the current discharging C_d flows in the internal loop within the diode.

Calculation of electrical parameters

To quantify the performance and operational parameters of the SPAD, the voltage output waveform from the node R_q is analysed.

1. Diode and stray capacitances (C_d and C_s): The rise time is derived from the waveform at node R_q and is calculated using the formula from Equation 5.7.

   Given R_q is 2 kΩ, and from the waveform data at node R_q, the rise time T_r is observed, leading to the computation of the mean of combined capacitance (C_d + C_s) as 0.107 nF.

2. Diode resistance (R_d): Using the previously derived mean value of (C_d + C_s), the resistance R_d from quenching time T_q can be calculated from Equation 5.6.

   Solving this with the known values gives a mean R_d of 1.1 kΩ.

3. Diode current (I_d): The diode current is calculated when the overvoltage (V_{ov}), which is the difference of bias voltage and breakdown voltage, applied across the diode is known:

$$I_d = \frac{V_{ov}}{R_d} \tag{5.8}$$

   With an overvoltage of 500 mV, the calculated diode current I_d is 0.4 mA.

Transient response

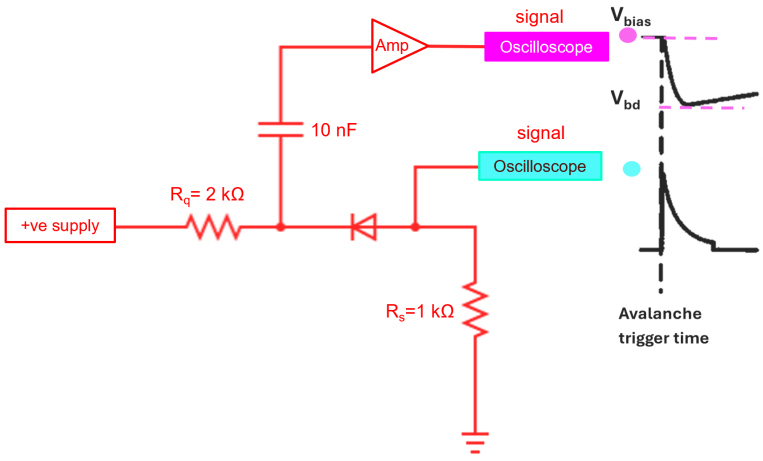
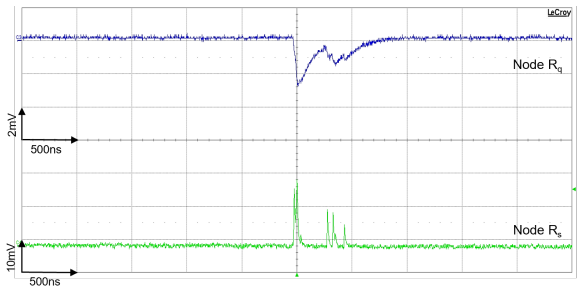
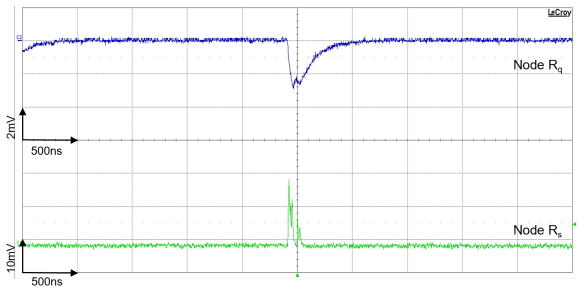


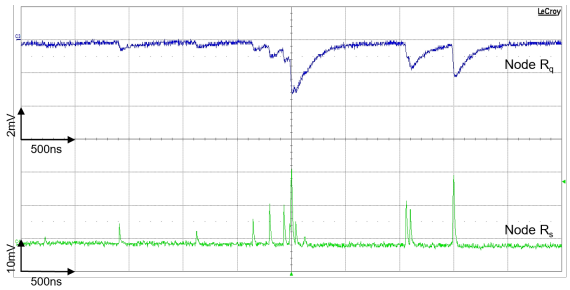
Figure 5.10. Detailed circuit layout for transient response of SPAD.



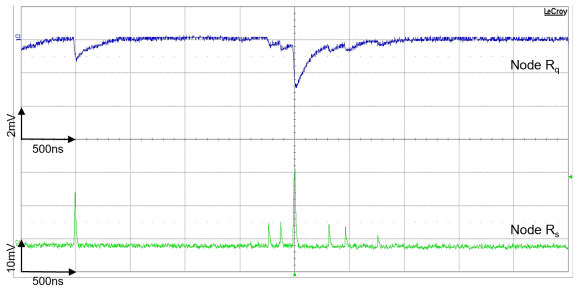
(a) At 50mV



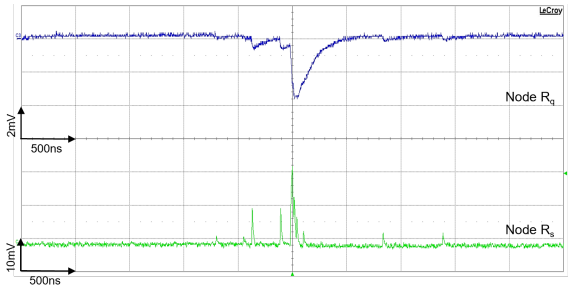
(b) At 100mV



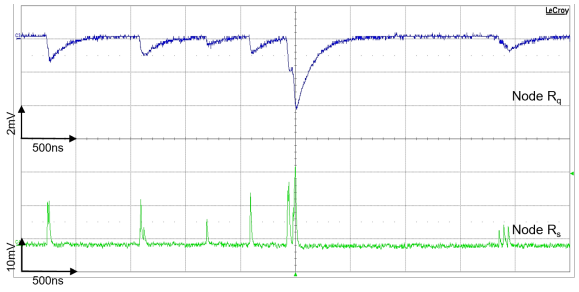
(c) At 200mV



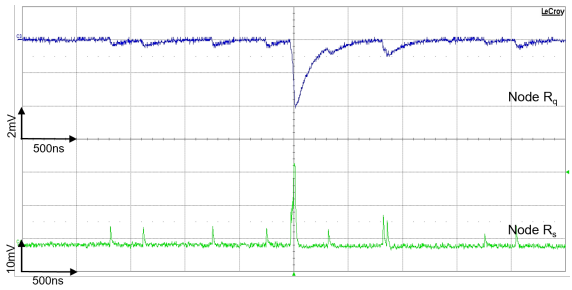
(d) At 300mV



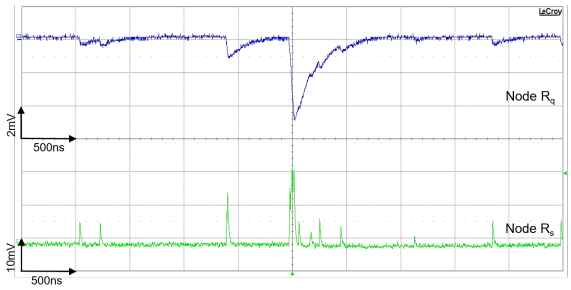
(e) At 400mV



(f) At 500mV



(g) At 600mV



(h) At 700mV

Figure 5.11. Transient response of the 1×1 pixel sensor showing voltage and current mode output waveforms at different overvoltages.

For AC measurements, a Phillips Wideband 6954 amplifier and a LeCroy 610Zi oscilloscope are employed to record the waveforms. The experiment is conducted in a dark environment to analyze the quenching behavior under varying overvoltages and temperature conditons. At each bias voltage, over 1,000 waveforms are recorded to obtain statistically significant results. Both voltage and current outputs are measured, with node R_q representing the voltage output and the node R_s corresponding to the current output, as detailed in Figure 5.10. The oscilloscope is configured to ensure accurate waveform capture and clear differentiation between voltage and current outputs. The upper curve at node R_q is set to 1 MΩ DC coupling with low bandwidth mode enabled to minimize high-frequency noise interference while the lower curve at node R_s is also set to low bandwidth mode but with 50 Ω DC coupling. The trigger mode is set to DC, with the upper curve at node R_q triggered on the negative edge and the lower curve at node R_s triggered on the positive edge, ensuring the precise observation of avalanche triggering.

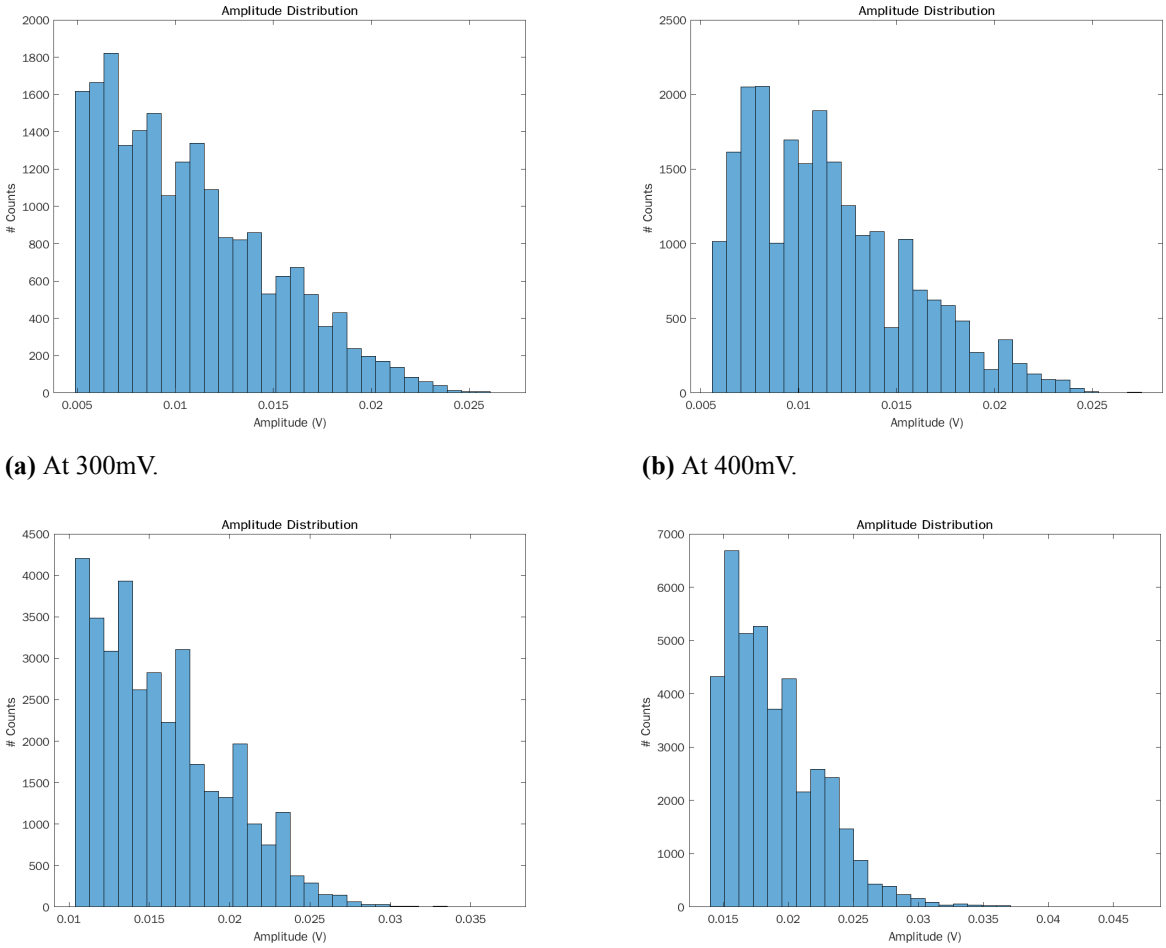
The waveform, presented in Figure 5.11, illustrates the behavior of the 1×1 pixel sensor at room temperature for overvoltages ranging from 50 mV to 700 mV. As expected, at the nodeR_q, the bias voltage initially drops to the breakdown voltage when the avalanche triggering occurs and then begins its recovery phase. However, rather than exhibiting a smooth and well-defined quenching, the waveform demonstrates irregular fluctuations. The expected behavior for an ideal SPAD is a single, distinct pulse that indicates the detection of an individual photon, but in this case, multiple small pulses appear before the primary pulse, disrupting the anticipated signal shape. It suggests that the avalanche is not fully developed over the pixel but rather localized in a small area.

Examining the current output at the node R_s provides a more detailed view of the pulse structure. The presence of mini pulses preceding the main pulse is an indication of inconsistencies in avalanche quenching. Rather than a single, well-defined pulse, the main pulse itself appears fragmented, composed of multiple sub-pulses. As the overvoltage increases, these mini pulses become more frequent, leading to an increased dark count rate. Still, no formation of a coherent pulse consistent with a fully developed avalanche is observed. The presence of numerous small pulses at short intervals, each with varying amplitude, indicates an adverse effect on avalanche triggering and an increase in the probability of pulse pile-up.

5.2.3 Amplitude distribution

The amplitude distribution is analyzed to assess the uniformity of the SPAD at different overvoltages. At least 1000 waveforms are recorded at various overvoltage settings, all performed at room temperature under dark conditions. A peak-finding algorithm implemented in MATLAB is used to identify the waveform peaks that exceed the noise threshold level.

The selected peaks are then used to generate amplitude histograms, allowing the study of how the output signal distribution behaves across different overvoltage values, ranging from 300 mV to 600 mV, as illustrated in Figure 5.12.



(a) At 300mV.

(b) At 400mV.

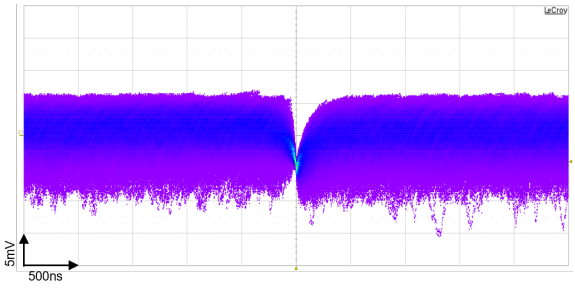
(c) At 500mV.

(d) At 600mV.

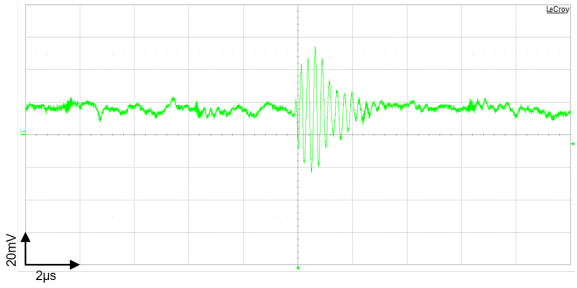
Figure 5.12. Amplitude distribution of recorded waveforms of 1×1 pixel sensor (01E15) at different overvoltages ranging from 300 mV to 600 mV.

Ideally, in single-pixel detection measurements, a Gaussian distribution reflecting a uniform response in the avalanche process is expected. However, the distributions presented do not follow any standard pattern in all overvoltage values, showing deviations from the expected Gaussian profile. The plots reveal that the amplitudes are not fixed, and the gain appears to be quite low. This could be due to a phenomenon of edge breakdown, where the avalanche process seems to be happening in a very tiny part, showing inconsistent signal output.

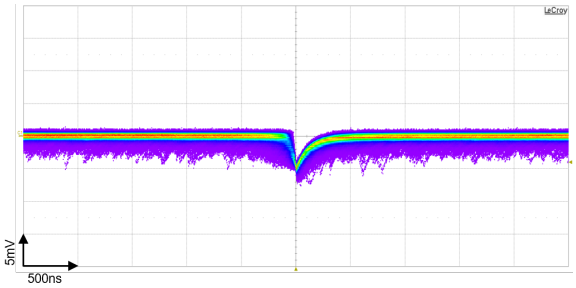
5.2.4 Persistence waveform analysis



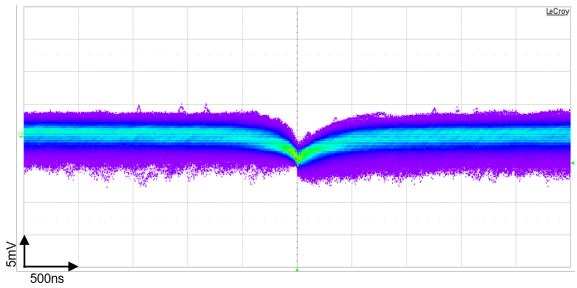
(a) Type 01E05 - at room temperature.



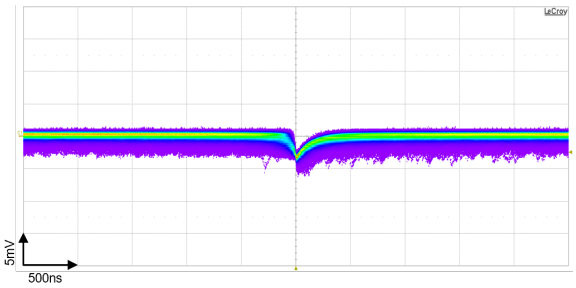
(b) Type 01E05 - at 5°C.



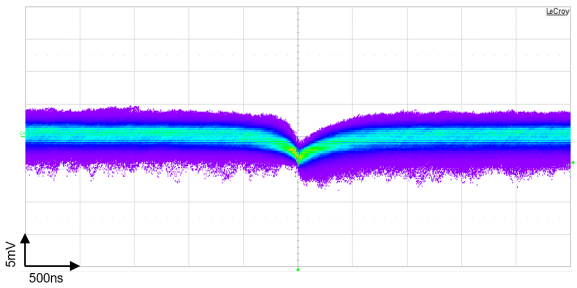
(c) Type 01E10 - at room temperature.



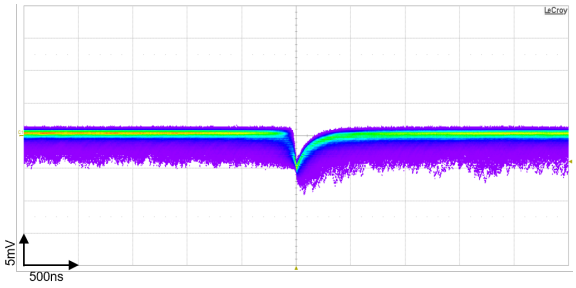
(d) Type 01E10 - at 5°C.



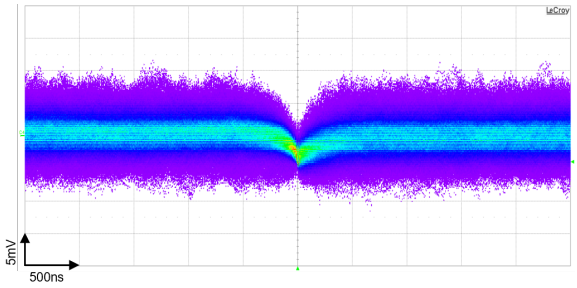
(e) Type 01E15 - at room temperature.



(f) Type 01E15 - at 5°C.



(g) Type 01E20 - at room temperature.



(h) Type 01E20 - at 5°C.

Figure 5.13. Proof of principle of quenching mechanism and noise levels at room temperature and 5°C of all single pixel variations (01E05, 01E10, 01E15 and 01E20) in persistent mode.

To further investigate the behavior of the sensor, the analysis is conducted at different temperatures in the persistent mode while adjusting the breakdown voltage accordingly, as the breakdown voltage is dependent on temperature. The temperature is maintained by placing the PCB in the light-tight temperature-controlled oven. Figure 5.13 compares four single-pixel sensor variations, namely *01E05*, *01E10*, *01E15*, and *01E20*, and provides insight into how temperature affects the quenching process. The left side of the figure shows measurements taken at room temperature, while the right side presents data collected at 5 °C.

Ideally, lower temperatures should reduce noise by decreasing the thermal activity of charge carriers and improving quenching behavior. However, the results deviate from this expectation, revealing unexpectedly high noise levels at lower temperatures. One crucial factor influencing sensor performance is the guard ring value, which varies across sensor types. The *01E05* sensor, with the smallest guard ring value of 0.5 μm, exhibits the highest noise levels, making it the least efficient variation. At room temperature, the noise is already prominent, and at 5°C, a well-defined quenching pulse is not clearly visible. Moving to sensors with higher guard ring values, the quenching pulse improves, yet noise remains an issue at lower temperatures. This observation suggests that, while temperature variations play a role in noise behavior, sensor design factors such as the guard ring size significantly impact overall performance.

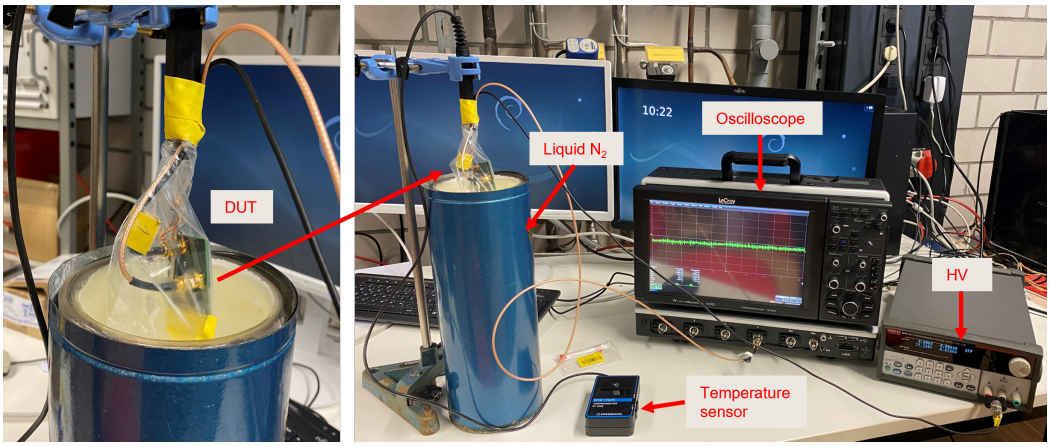


Figure 5.14. Liquid nitrogen measurement setup.

The results are particularly interesting under extreme cooling conditions, such as those shown in Figure 5.15, where the sensor is tested in a liquid nitrogen environment. The experimental setup, shown in Figure 5.14, consists of a cylindrical vessel filled with liquid nitrogen, with the PCB suspended above the liquid. A GTH 175 PT temperature sensor is used to obtain accurate temperature readings in the liquid medium. Given that the GTH 175 PT is a high-precision Pt1000 probe, its reliability in extreme conditions ensures that the recorded temperatures are accurate. Contrary to expectations, the persistence mode output at extremely low temperature, displayed in Figure 5.15, reveals a waveform dominated by

noise. This result is unexpected, as extreme cooling should typically suppress noise due to the reduced thermal agitation of charge carriers. Instead of observing a cleaner quenching pulse, the data indicate an increased presence of random noise.

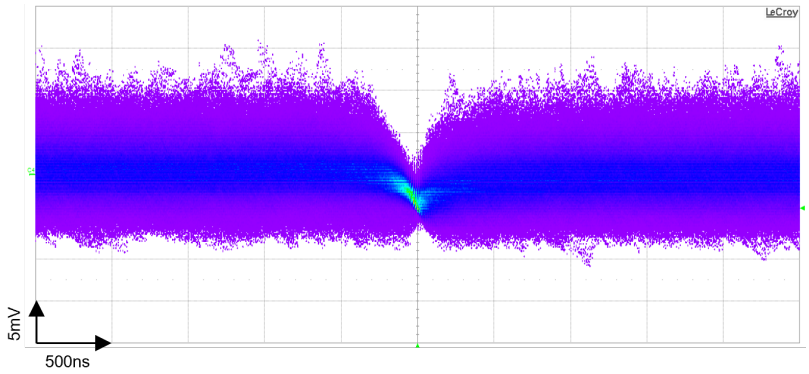


Figure 5.15. Persistent mode output at -52°C.

From the analysis described, the presence of mini pulses without a fixed amplitude, increased avalanche propagation speed, and persistently high noise levels at low temperatures, including those not improving under liquid nitrogen conditions, indicates potential issues with the sensor's performance. These irregularities might be attributed to premature or uneven breakdown events within the sensor. One possible cause for these observations could be edge breakdown, which typically results from local maxima of the electric field near the boundaries of the sensor's active area. Such a strong field can cause the sensor to prematurely trigger an avalanche with inconsistencies, leading to unpredictable behavior and noise.

After observing the high dark counts, the measured breakdown voltage should also be considered now, which comes out to be more than double the simulation results. These results suggest that further investigation is necessary to identify the issues affecting the sensor's performance.

5.2.5 Time distribution

The time distribution analysis provides insights into the underlying noise mechanisms within the SPAD device. By studying the time differences between consecutive pulses, known as interarrival times (Figure 5.16), it becomes possible to separate diffusion-driven events, which occur at short time intervals, and thermally generated events, which are widely spaced pulses.

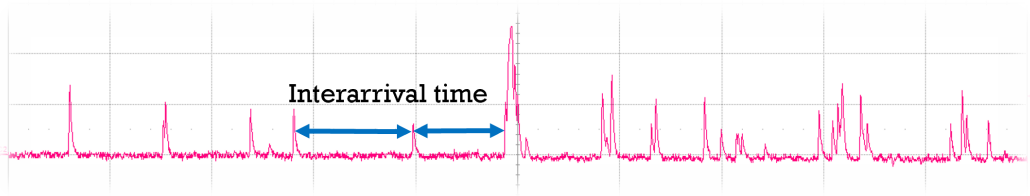
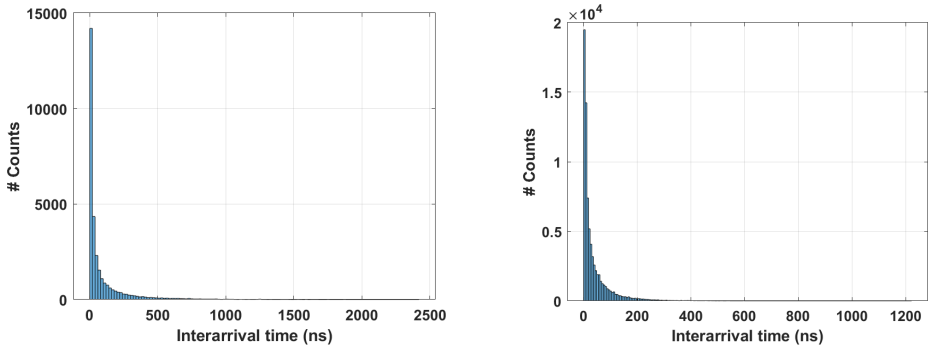


Figure 5.16. Interarrival time definition.

The analysis is performed by recording at least 1000 waveforms under dark conditions. The measurements are carried out at different operating temperatures (5°C, 15°C, 25°C, and 35°C) with an overvoltage setting of 400 mV and the results are shown in Figure 5.17.



(a) At 5°C. **(b)** At 15°C.

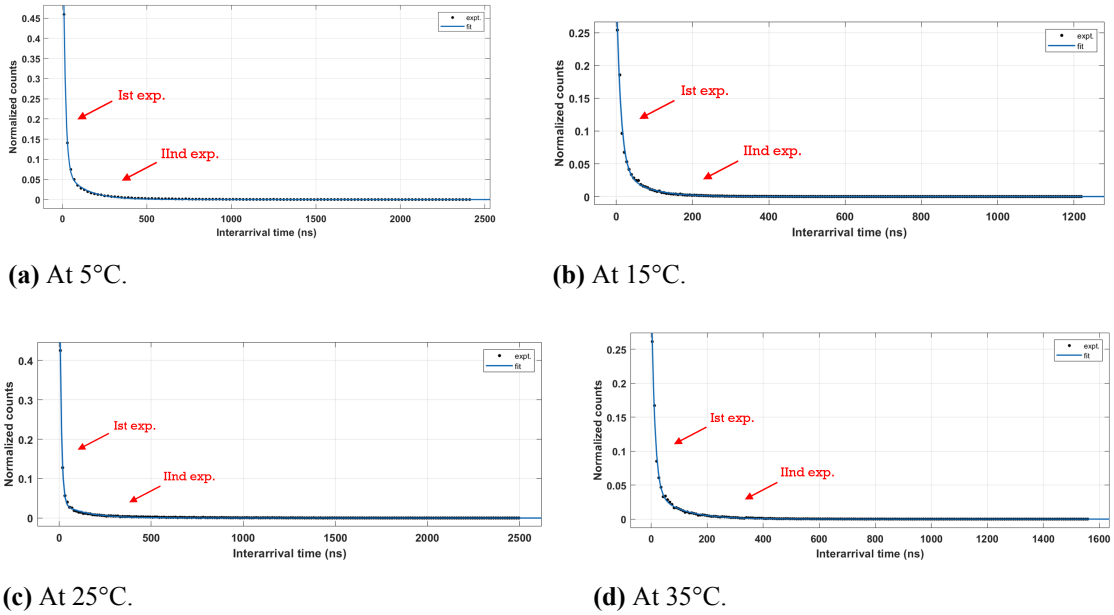
(c) At 25°C. **(d)** At 35°C.

Figure 5.17. Measured interarrival time histogram at different temperatures.

In SPADs, photon arrivals or dark events appear randomly and uniformly over time and can be described statistically as a Poissonian process. As a result, the number of events within a specified time frame follows a Poisson distribution, while the times between these events, the interarrival times, are exponentially distributed. The pattern seen in the measured data at each temperature thus shows an exponential distribution, which aligns with the Poisson nature of photon arrivals.

The data from these plots are further analyzed by fitting them to a double exponential

function, defined as $A\exp(-bt) + C\exp(-dt)$. This fitting, shown in Figure 5.18, helps examine two distinct components within the dark pulses. The first term with time constant b accounts for the mini pulses that are frequent and closely spaced, representing diffusion-driven events, and the second term with time constant d captures the pulses that are more widely spaced, indicative of thermal generation effects.



(a) At 5°C.

(b) At 15°C.

(c) At 25°C.

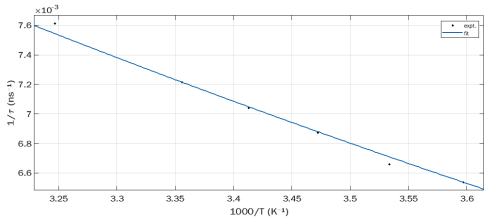
(d) At 35°C.

Figure 5.18. Double exponential fitting with the fit function = $A\exp(-bt) + C\exp(-dt)$ at different temperatures.

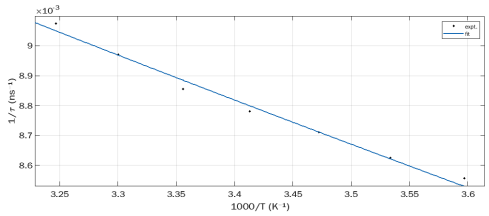
The time constants b and d from the fitting results are extracted and analyzed. Both diffusion and thermal time constants are plotted against temperature. While the diffusion time constant b showed non-monotonous behavior and so is not further analyzed, the thermal time constant d demonstrates significant findings. As per the Arrhenius equation,

$$\tau = Nexp(-E_{act}/kT), \tag{5.9}$$

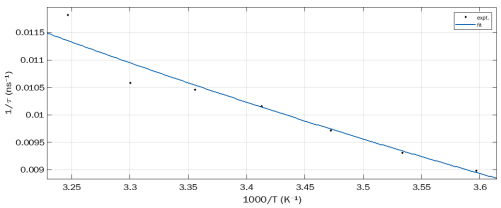
which models the temperature dependency of the charge carrier processes within the sensor, the thermal time constants are analyzed. In this equation, τ is the time constant associated with the thermal process, E_{act} represents the activation energy which tells the location of the trap energy levels and should be less than E_g, T is the absolute temperature, k is the Boltzmann's constant, and N is the pre-exponential factor that tells about the trap densities.



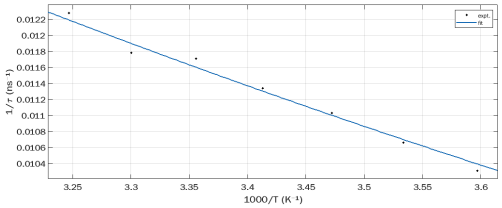
(a) At 300 mV of overvoltage.



(b) At 400 mV of overvoltage.



(c) At 500 mV of overvoltage.



(d) At 600 mV of overvoltage.

Figure 5.19. Arrhenius plot of thermal time constant at different overvoltages.

The results, plotted in Figure 5.19 using a single exponential fit, both activation energies and trap densities are extracted. Notably, as the temperature decreases, the emission time for charge carriers to escape from trap states increases, suggesting a higher energy barrier at lower temperatures. The extracted activation energies are 0.408 eV, 0.169 eV, 0.677 eV, and 0.4549 eV for overvoltages of 300mV, 400mV, 500mV, and 600mV, respectively along with varying trap densities, 0.028 cm^{-3}, 0.0156 cm^{-3}, 0.102 cm^{-3}, and 0.053 cm^{-3}. Ideally, the activation energy and trap densities do not change with the variation of overvoltage, however, the results show potential non-uniformities affecting sensor performance.

5.2.6 Light emission test

A light emission test (LET) experiment has been conducted to investigate the location where avalanche breakdown is triggered in the sensor, following the major need after observing the unexpected results in the measurement. Using a self-built microscope-based camera set-up, the experiment is aimed at capturing the emission patterns from the sensors when subjected to bias above breakdown voltage.

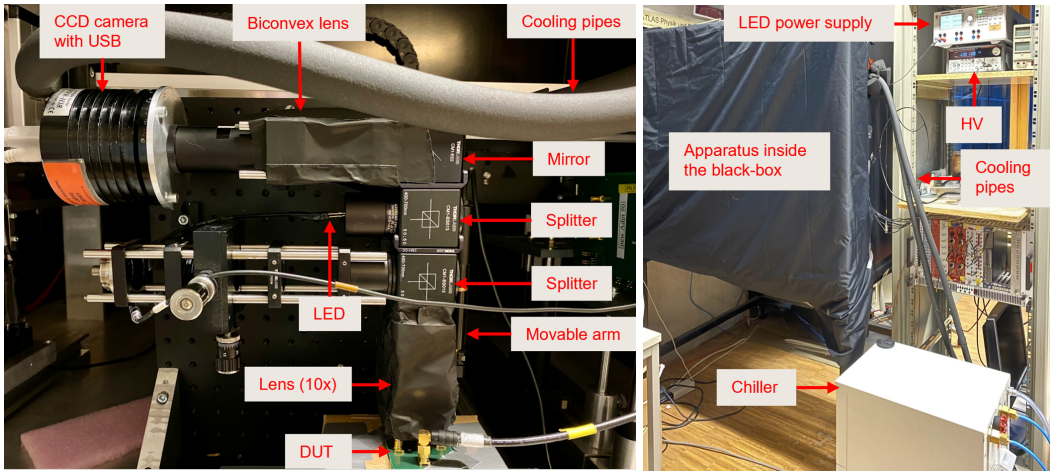
Experimental setup

The setup consists of an adjustable optical path equipped with various components, shown in Figure 5.20, including 50/50 optical splitters, lenses, mirrors, collimators, and a CCD camera. All of these components are arranged on an aluminum base assembled from two planes to provide a stable and vibration-free platform. The whole assembly is enclosed in a dark box to eliminate any ambient light interference, thereby ensuring accurate and reliable

measurements.

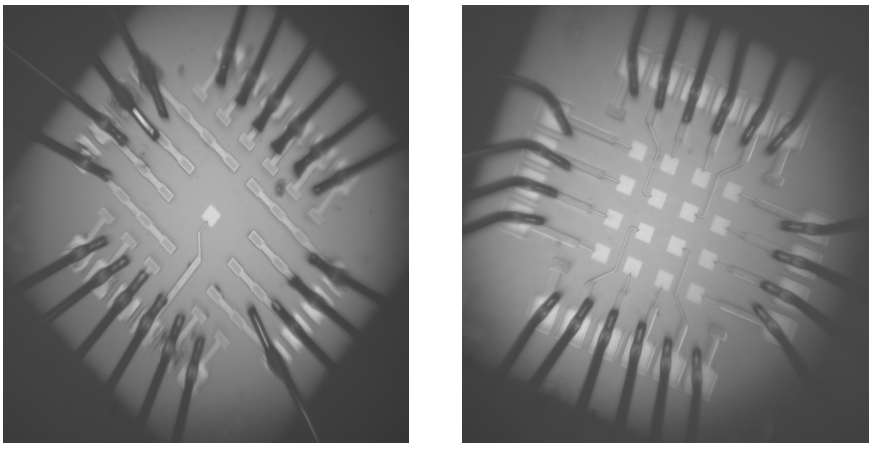
A Starlight Xpress SXVR-H18 CCD camera is chosen for its low-noise performance, making it well-suited for capturing light emissions. To minimize the noise from the camera sensor, the camera temperature is set to -10°C. Additionally, an external cooling system is implemented, maintaining a temperature of 5°C via liquid loop pipes to prevent overheating of the camera caused by the dark current.

Achieving precise focus on the sensor is critical for capturing detailed images of the emitted photons. A 10× objective lens, mounted on a remote-controlled movable arm provided by a Thorlabs APT-DC controller, is used to fine-tune the focus with sub-micron precision. To aid in the focusing process, a white light-emitting diode (LED) served as an external light source. This LED is carefully covered with black foil to minimize any unwanted interference with the CCD camera's imaging [60].



(a) Internal setup. **(b)** External setup.

Figure 5.20. Experimental setup used for Light Emission Test.



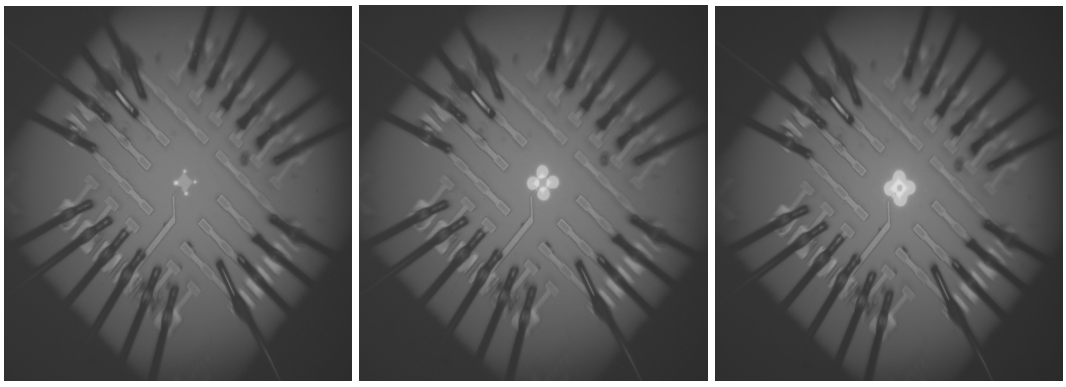
(a) 1×1 pixel. **(b)** 4×4 pixel.

Figure 5.21. Focused basic picture of sensors without the exposure.

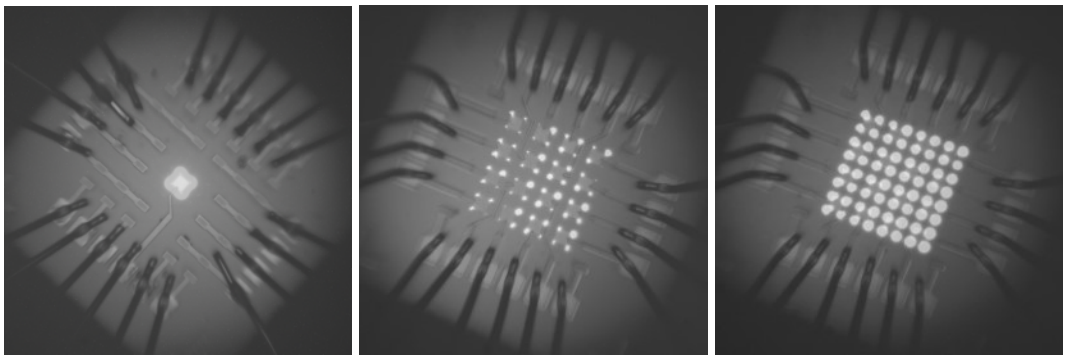
Experimental procedure

The initial step involves focusing the sensor using the LED light while in a zero-bias voltage state. The remote-controlled, movable arm is used to adjust the lens, ensuring optimal calibration of the sensor structure. The crucial parameter in this step is the distance between the objective lens and the sensor surface, as achieving the best image clarity depends on precise focusing. In the camera control application, the exposure time is set to 1 second, and a continuous image capture mode is used during the focusing process. This allowed real-time monitoring of sensor alignment and clarity. The focused 1×1 and 4×4 pixel(s) are shown in Figure 5.21.

Once the sensor is properly calibrated and focused, the LED light is turned off to conduct the light emission test in a dark environment at room temperature. The exposure time for capturing images is set to 30 minutes to ensure sufficient photon collection. The sensor is biased well above the breakdown voltage, enabling observation of the emission patterns.



(a) 1×1 pixel at V_{ov} = 4V. **(b)** 1×1 pixel at V_{ov} = 11V. **(c)** 1×1 pixel at V_{ov} = 16V.



(d) 1×1 pixel at V_{ov} = 21V. **(e)** 4×4 pixels at V_{ov} = 4V. **(f)** 4×4 pixels at V_{ov} = 16V.

Figure 5.22. Light emission test results for 1×1 and 4×4 pixel(s).

Figure 5.22 presents the light emission patterns at different overvoltage values. Initially, for a 1×1 pixel, the overvoltage is set to 4 V, where photon emission is primarily observed at the corners of the sensor. As the voltage is increased to 11 V and subsequently to 16 V,

photon emission appears to gradually expand towards the middle of the sensor. At 21 V of overvoltage, an overvoltage considered beyond typical operational limits from a SPAD, a significant emission is observed originating from the centre of the sensor. It is hypothesized that the emission observed in the corners could have contributed to an optical flow effect, making the light appear to migrate toward the center. The same criteria are visible in 4×4 pixels at overvoltages of 4 V and 16 V.

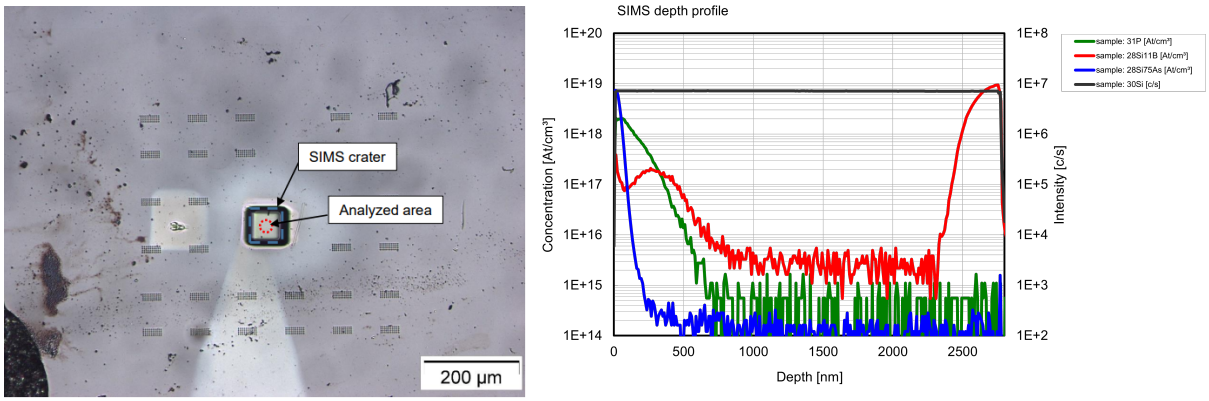
The experimental results indicated that early avalanche breakdown occurs predominantly at the corners of the p-n junction. This suggests that the electric field intensity is higher at the junction corners compared to the central region. The primary function of the guard ring is to regulate the electric field distribution by reducing its intensity at the edges. The unexpected behavior observed in the experiment suggests that there may have been deviations in the doping concentrations or profiles during the fabrication process. These deviations could result in an unintended increase in the electric field at the corners, leading to premature breakdown.

In addition to doping irregularities, the physical structure of the sensor, particularly the corner geometry, plays a crucial role in field distribution. Ideally, the corners should be sufficiently rounded to prevent field crowding, which can otherwise result in localized electric field intensification. If the corners are not adequately rounded during the fabrication process, it could contribute to the high electric field concentration at the corners, observed in the experiment.

5.3 Secondary ion mass spectroscopy

Secondary Ion Mass Spectrometry (SIMS) is an analytical technique used to estimate the doping concentration of the fabricated device. A focused ion beam bombards a sample, ejecting many secondary ions. A mass spectrometer analyzes these ions, permitting measurement of precise depth profiling of implanted elements and extracting concentration variations.

The SIMS analysis is conducted to investigate the depth profiles of boron (B), phosphorus (P), and arsenic (As) in a structured silicon 1×1 pixel sensor (*01E20*). The purpose of this study is to compare the expected doping profiles derived from TCAD process simulations with the actual doping profiles measured in fabricated sensors. The reason behind this test is to confirm if there is any mismatch of doping profiles between the processed sensors and the simulated ones, as guessed from the measurement results.



(a) Sensor details for SIMS. **(b)** SIMS depth profile.

Figure 5.23. Secondary Ion Mass Spectroscopy: showing details of the analysis and the resulting doping concentration.

The analysis is done by SGS Institut Fresenius, Dresden. The samples are measured at the center of the silicon structure, ensuring accurate profiling of the dopants. The analysis is conducted using a Cameca 7F instrument with Cesium (Cs^+) as the primary ion species, set to an energy of 15 keV and a primary ion current of 15 nA. The scan area is set to $75\,\mu$m $\times$ $75\,\mu$m, and the analyzed secondary ions included $^{28}Si^{11}B$, ^{31}P, $^{28}Si^{75}As$, and ^{30}Si. The mass resolution, defined as $M/\Delta M$, is set to 4000 to ensure high-precision measurements [61]. The analyzed area is $35\,\mu$m, and the depth scale is calibrated using crater depth measurements. To remove the layers on the sample, an HF lift-off process is applied before the SIMS measurement. The doping concentrations are quantified based on reference implantations, ensuring reliable results. The position of the SIMS crater and the analyzed area on the sample with the depth profile measurement are shown in Figure 5.23.

5.3.1 Comparison with original parameters

The SIMS-measured doping profile is imported into the TCAD process simulation to compare it with the original simulated doping parameters. The data is imported after simulating the geometry of the sensor and used in place of the original corresponding ion implantation step. Figure 5.24 presents the depth-dependent doping concentration of total, boron, phosphorus, and arsenic after importing the SIMS data into TCAD. Due to the resolution of SIMS being limited to $1 \times 10^{14} cm^{-3}$, a relatively constant profile can be observed in the middle of the plot. Also, SIMS overestimates the depth of the sensor; therefore, the comparison is done at the same depth level which shows the magnitude difference at the beginning of the SIMS-measured and TCAD-imported plots. Figure 5.25 illustrates the net doping concentration of both the original and SIMS-measured profiles. A noticeable variation exists between these

profiles, with the junction appearing closer to the surface in the SIMS-measured profile than in the original simulation. Additionally, the peak doping concentration in the SIMS profile is lower than the expected values from the initial simulation.

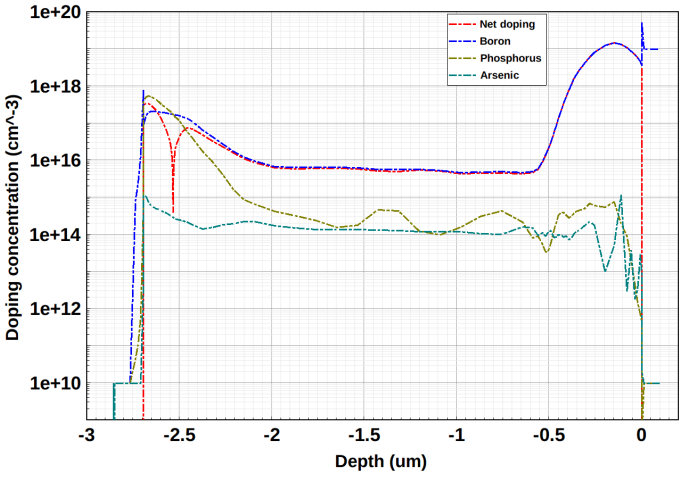


Figure 5.24. SIMS simulated total and material-wise doping concentration.

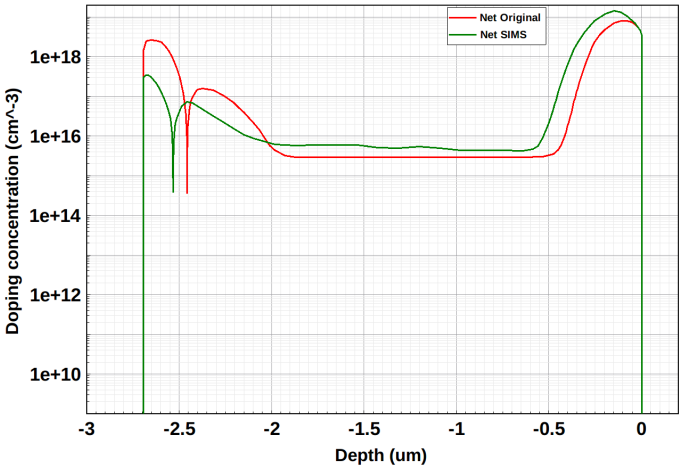
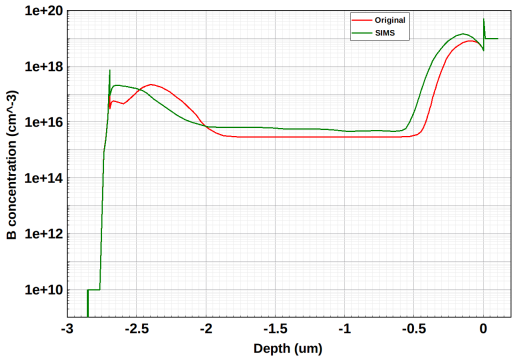


Figure 5.25. Net comparison of the original and SIMS imported doping concentration.

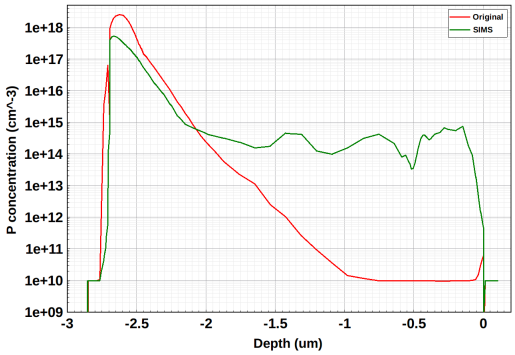
5.3.2 Comparative analysis of simulation and experimental data

Further analysis is shown in Figure 5.26, where the individual doping profiles of boron, phosphorus, and arsenic from the original simulation and the SIMS measurement are compared.

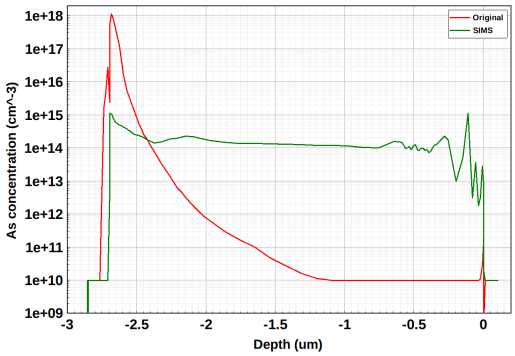
This comparison highlights discrepancies between the expected and measured profiles, indicating variations in implantation parameters during fabrication. The deviations suggest that either different implantation energies or doses were used or unexpected diffusion effects occurred during processing.



(a) Boron doping concentration.



(b) Phosphorus doping concentration.



(c) Arsenic doping concentration.

Figure 5.26. Comparison of simulated original and SIMS-imported doping concentrations.

To understand the discrepancies and identify the correct implantation parameters used in fabrication, modifications are made to the original TCAD doping parameters. The initial simulation assumed the implantation parameters shown in Table 5.1.

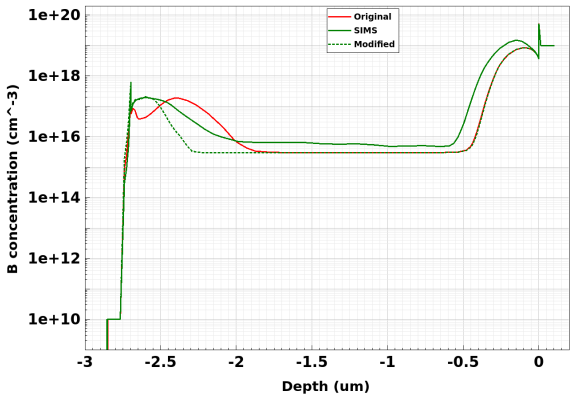
Implantation	**Energy (*keV*)**	**Dose (cm^{-2})**
Boron	95	0.6×10^{13}
Phosphorus	70	4×10^{13}
Arsenic	30	0.4×10^{13}

Table 5.1. Original implantation parameters.

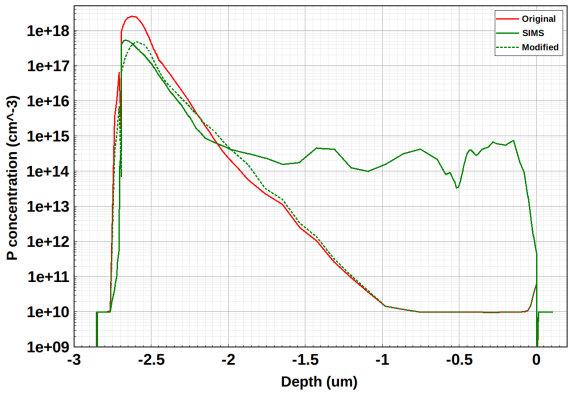
Implantation	**Energy (*keV*)**	**Dose (cm^{-2})**
Boron	35	0.3×10^{13}
Phosphorus	100	0.8×10^{13}
Arsenic	50	0.001×10^{13}

Table 5.2. Modified implantation parameters.

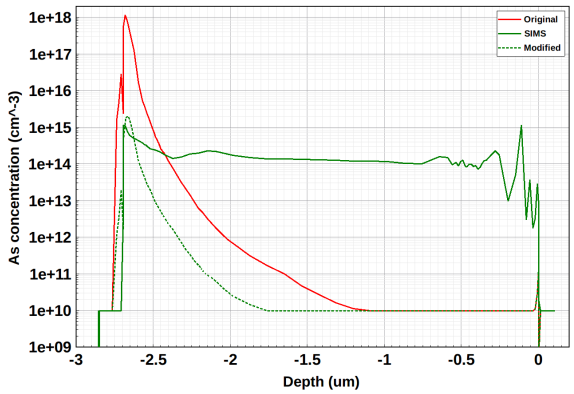
However, after analyzing the SIMS profiles, the optimized doping parameters that better matched the SIMS-measured data are shown in Table 5.2.



(a) Boron doping concentration.



(b) Phosphorus doping concentration.



(c) Arsenic doping concentration.

Figure 5.27. Comparison of original, SIMS and modified doping concentrations.

The differences between the original and optimized parameters highlight significant variations in the actual fabrication process shown in Figure 5.27. The fabricated sensors have approximately three times less energy and half the dose for boron, 1.5 times more energy and four times less dose for phosphorus, and twice the energy with a significantly lower dose for arsenic assumed. These changes in implantation conditions significantly altered the doping profile, resulting in deviations from the expected TCAD simulation. Since arsenic is primarily used for contacts rather than forming the p-n junction, its deviations in implantation parameters do not significantly impact the overall device performance. However, boron and phosphorus play a critical role in defining the p-n junction characteristics. The lower implantation energy and dose for boron resulted in a shallower junction, which directly influences the electrical properties of the sensor. Similarly, the phosphorus implantation depth increased due to higher implantation energy, shifting the junction location.

5.3.3 IV comparison

The modified doping profiles directly impact the electrical characteristics of the sensor. As boron and phosphorus determine the junction behavior, any deviation in their concentration and depth affects the breakdown voltage and overall device performance.

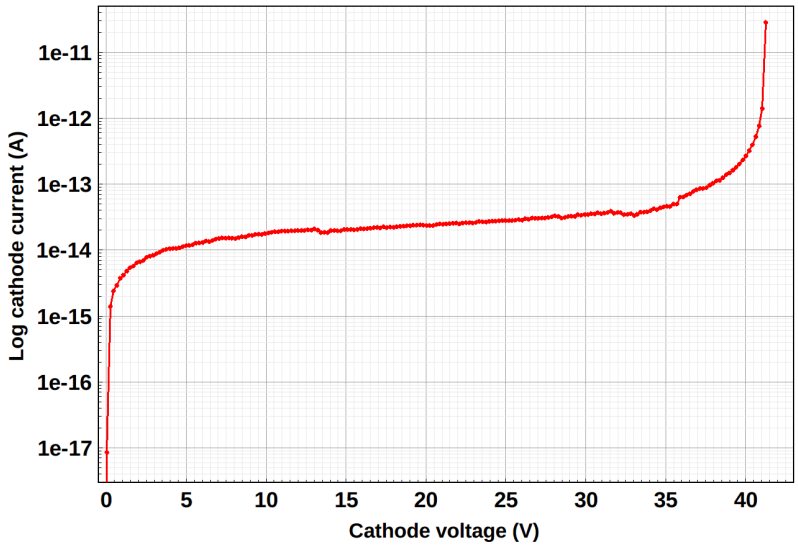


Figure 5.28. Logarithmic IV curve of modified parameters of doping concentration.

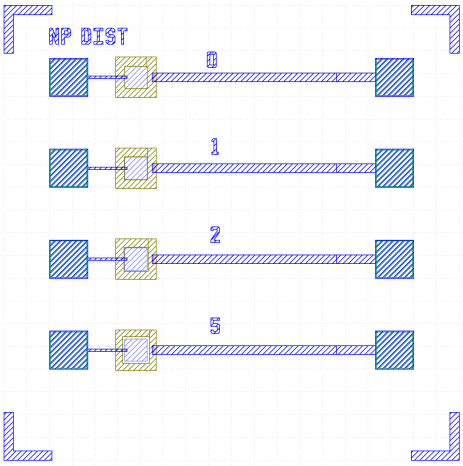
To validate the impact of the SIMS-based doping profile on device performance, electrical simulations are performed in TCAD using the modified doping profiles shown in Figure 5.28. The simulated breakdown voltage using the adjusted SIMS doping parameters is found to be 42 V, which closely aligns with the experimentally measured breakdown voltage from laboratory tests. This agreement suggests that the fabricated sensors may have incorrect implantation parameters, leading to deviations from the expected electrical characteristics.

The consistency between the SIMS-based TCAD simulation and the measured breakdown voltage suggests that the implantation process differs from the designed specifications. The results emphasize the importance of integrating SIMS profiling into the fabrication workflow to detect and correct deviations in the doping process, ensuring consistency between simulation and experimental outcomes.

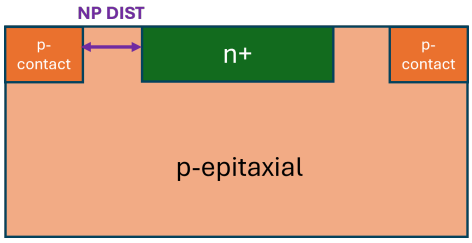
5.4 Findings from test structures

A few test structures fabricated on the wafer are examined by manufacturers to ensure their adherence to design specifications, focusing particularly on the impact of structural variations within the sensor elements. One significant study involves the spacing between the n+ region, which forms part of the pn-junction, and the p-contact. This spacing is critical as it influences the breakdown voltage of the sensor.

Figure 5.29 shows the test structure of n+ - p-contact distance where the lateral distance between them varies from 0 μm to 5 μm, and a visual representation of the structure for the clarification of the distance. The tests are performed at the wafer level, examining the current-voltage characteristics.



(a) GDS structure of n+ - p-contact distance variation.



(b) Visual representation showing n+ - p-contact distance for understanding.

Figure 5.29. A test structure layout showing the n+ - p-contact distance variation.

The results, shown in Figure 5.30, indicate that very small distances, like 0 μm and 1 μm, result in a low breakdown voltage around 6 V. When the distance increases to 2 μm, the breakdown voltage rises to about 26 V. Further extending to 5 μm, it reaches up to 38V, which aligns with our case. For reliable sensor operation, the gap between the n+ region and p-contact should be more than 5 μm according to this sensor design as per the manufacturers, to avoid the leakage effect. According to the observance of changing breakdown voltage with the lateral distance, it suggests that the doping of n+, which is phosphorus-doped, is correct. There are no apparent inconsistencies in the doping process

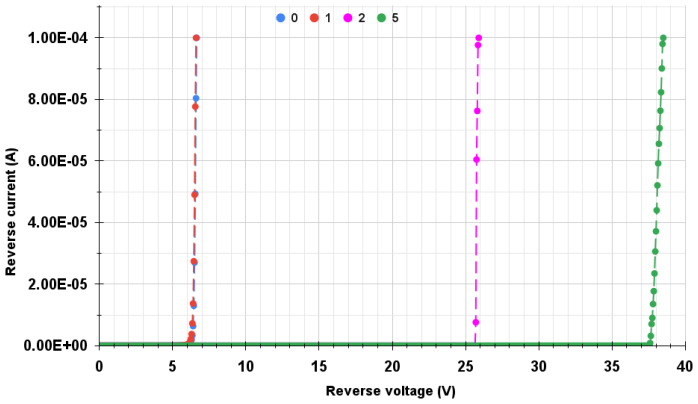
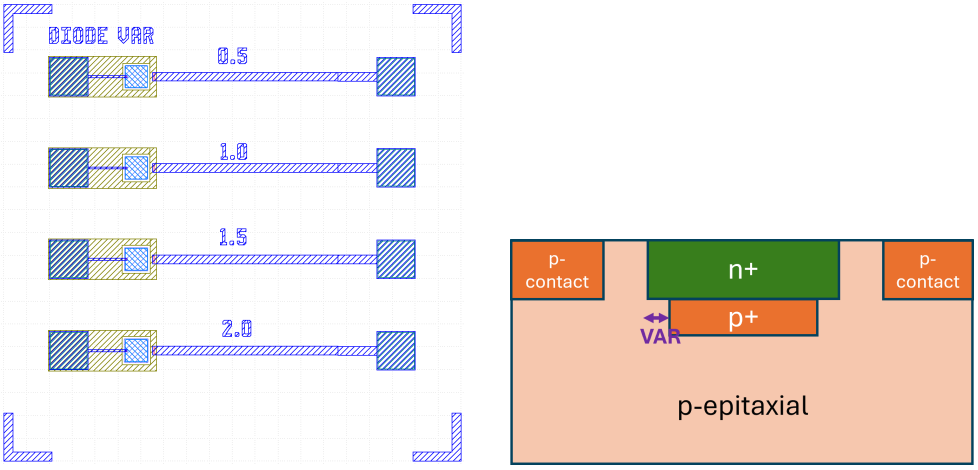


Figure 5.30. Reverse I-V curve of n+ - p-contact distance variation done on wafer level.

that could lead to unexpected performance.

In another test structure, variations in the widths of guard rings are analyzed, with values ranging from 0.5 μm to 2 μm. Figure 5.31 illustrates the test structure for these variations, detailing the spatial arrangement within the sensor design. Here, guard rings are formed virtually by adjusting the lateral overlap of the n+ and p+ regions.



(a) GDS structure showing the guard ring width variation.

(b) Visual representation of guard ring width variation for understanding.

Figure 5.31. A test structure layout showing the guard ring width variation.

Figure 5.32 presents the IV characteristics obtained from these guard ring variations, tested at the wafer level. Interestingly, despite the variations in guard ring thickness, the breakdown voltage remains almost constant across all tests, around 38 V. This outcome aligns with previous lab measurements of the sensors, where breakdown voltage does not vary with changes in guard ring dimensions. Typically, adjusting the guard ring's thickness should influence the sensor's edge effects, impacting the electric field distribution and consequently the breakdown voltage. The consistent breakdown voltage across variations suggests a potential issue with the p+ implantation process. Specifically, it indicates that the

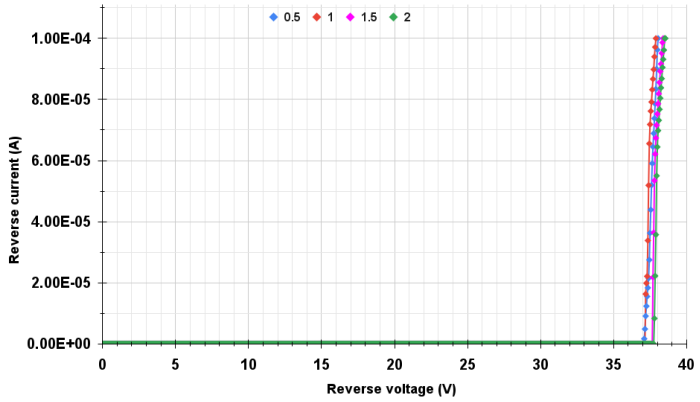


Figure 5.32. Reverse I-V curve of guard ring width variation done on wafer level.

boron implantation, which defines the guard rings and their interaction with the surrounding pn-junction, is a source of malfunctioning of the junction. This observation leads to the conclusion that there may be inconsistencies in the p+ implantation process. Such a flaw could affect the overall performance of the sensors, underscoring the need for potentially revising the implantation technique to ensure that expected electrical behavior is achieved.

5.4.1 IV curve characterization study of variations over wafer with a wafer prober setup

For an immediate evaluation of potential deviations in fabrication quality, several test wafers are analyzed alongside previously assessed samples. The CM300xi wafer probe station is employed for this purpose. This semi-automatic probe station incorporates an advanced digital microscope system to ensure precise probe-to-pad contact and alignment, essential for reliable DC analysis. The station maintains a controlled environment, free from moisture, light, and electromagnetic interference, facilitated by the Velox probe control software, which supports comprehensive wafer alignment and mapping. The wafer prober setup is depicted in Figure 5.33.

For the analysis, 34 diagonally arranged dies from the wafer are selected, each containing 8 subdies shown in Figure 5.34 and Table 5.3 with sensor designations. The chosen subdies are single-pixel sensors, half of which are covered and the other half exposed to light, the variation explained at the beginning of the chapter. The microscope's light is used for focusing the sensor and ensuring the alignment of probes on the pads. Once the focusing is achieved, the light is turned off to conduct the measurements in dark conditions. The wafer chuck temperature is set with an external temperature-controlled system.

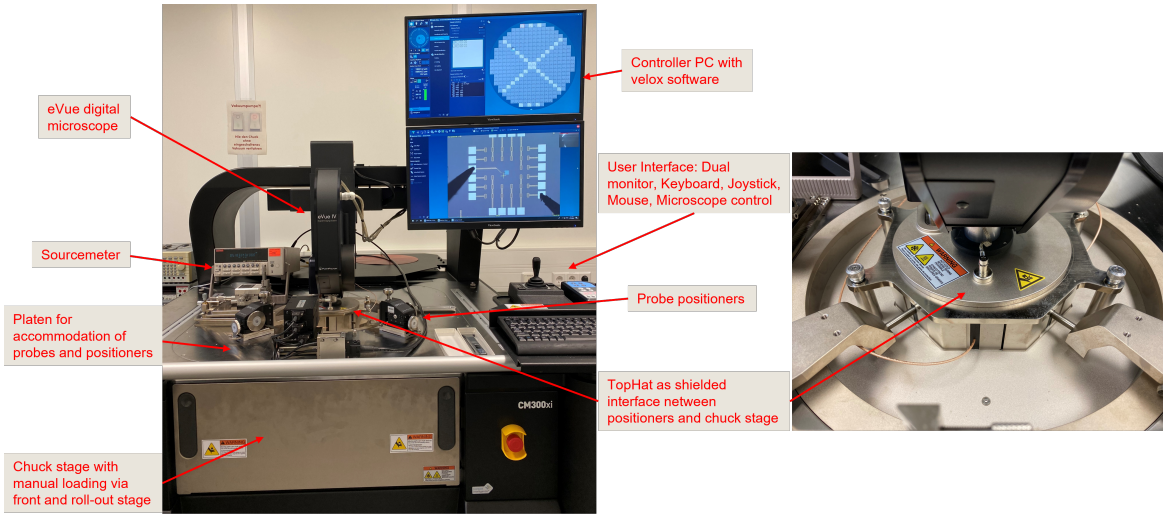
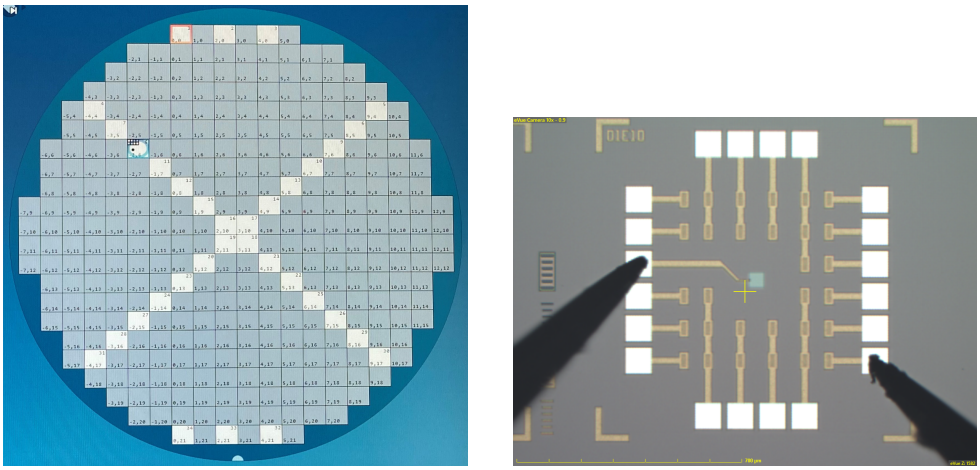


Figure 5.33. Wafer prober setup.



(a) Wafer mapping of selected dies. **(b)** Probes positioned on the pads.

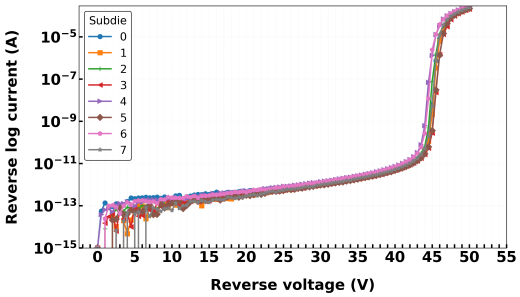
Figure 5.34. Wafer mapping and probe positioning on the selected dies .

Sub-dies	Sensors	Sub-dies	Sensors
0	01E10	4	01E05
1	01E20	5	01E15
2	01C10	6	01C05
3	01C20	7	01C15

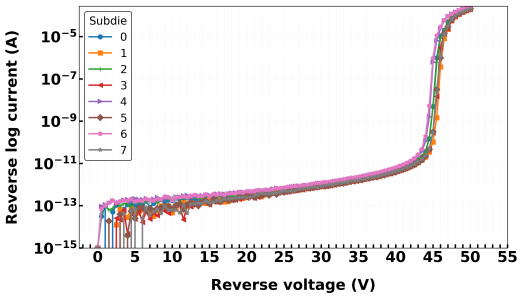
Table 5.3. Sub-dies with the sensor designations.

Figure 5.35 presents the measurement results for selected dies, namely die numbers 5, 15, 20, and 34, done at a temperature of 25°C. For each sub-die, the current is measured while varying the bias voltage from 0 to 50 V in 1 V of step-size. Each sub-die measurement in the voltage sweep takes approximately 80 seconds to complete.

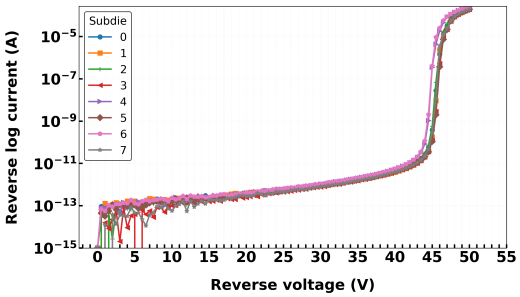
The observed breakdown voltages are around 45 V, which are slightly higher and suggest subtle inconsistencies that might be present in the manufacturing process. The sub-dies 4 and 6, corresponding to the least value of the guard ring from the variations of 0.5 μm, have at least 1 V of voltage shift compared to other sub-dies. This slight increase in the breakdown voltage implies minor variations that could impact the overall performance of the sensor. These results indicate that, although they are fabricated in the same batch as previous wafers, there might be slight discrepancies that do not align with simulated expectations, thereby highlighting the need for closer attention in manufacturing practices to ensure uniformity and predictability in the production of sensors.



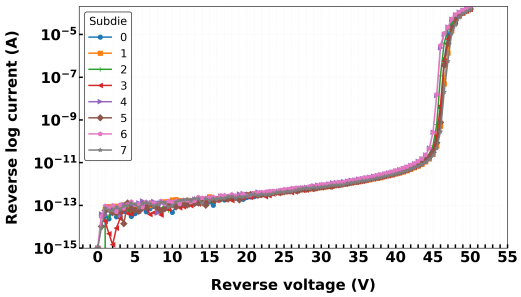
(a) Die 5.



(b) Die 15.



(c) Die 20.



(d) Die 34.

Figure 5.35. Reverse IV curve of a few dies showing a breakdown of around $45V$.

Chapter 6

Approach to the new sensor design development

As the measurement studies revealed significant issues during production, a systematic revision of the ion implantation parameters used for making pn junction is undertaken. By using TCAD simulations, the implantation parameters are redefined and the design approach is reconsidered to enhance production accuracy. This step ensures that the fabrication process aligns with the necessary performance and reliability standards.

6.1 Impact of boron (p+) implantation parameters

This section delves into the influence of boron (p+) implantation on the sensor's performance since boron implantation performed in making a p-n junction has shown almost no impact on the breakdown voltage extracted in Chapter 5 from the test structures. Three critical aspects are examined here: the doping concentration profile, the electric field distribution profile, and the current-voltage characteristics. Each of these factors plays a pivotal role in determining the operational effectiveness and reliability of the sensor. By performing several implantations at different energies, it is possible to achieve the desired doping distribution, such as a uniformly doped region.

6.1.1 Doping concentration profile

The variation in boron implantation energy is investigated to understand its impact on the SPAD's doping profile. As described in Chapter 4, the higher the energy of the ions implanted, the deeper they penetrate the material following a Gaussian distribution.

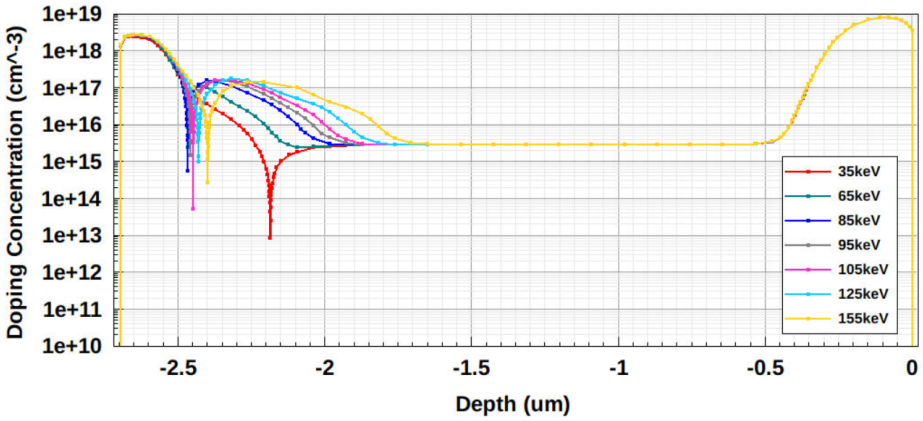


Figure 6.1. Net doping concentration with varied boron implantation energy at an unchanged dose as originally used. The left and right sides of the x-axis are the top and bottom of the sensor, respectively. The phosphorus is kept the same as the original (70 *keV*). The dip shows the point where p-n junction occurs.

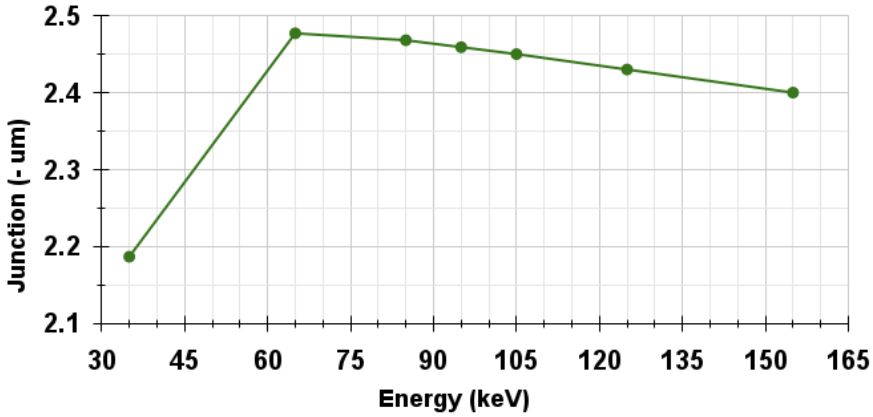


Figure 6.2. Junction depth from the surface at different boron implantation energies.

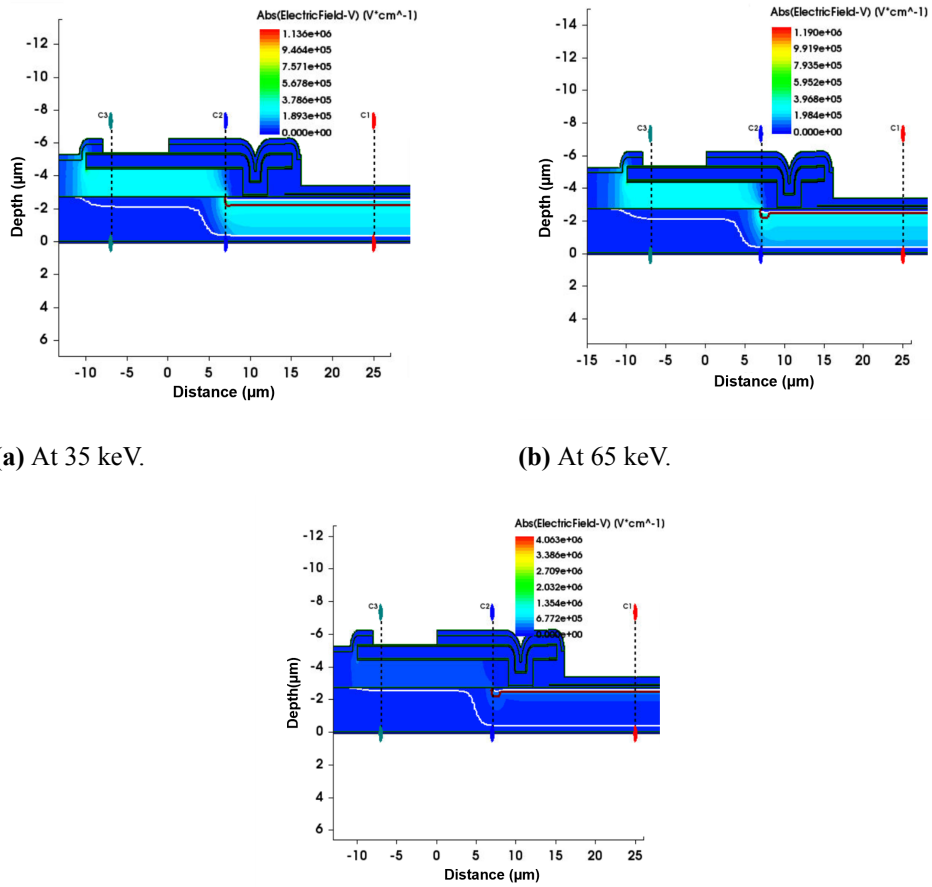
In the sensors, boron (p-type dopant) and phosphorus (n-type dopant) are implanted to form the p-n junction. Initial process simulations had a boron implant energy of 95 *keV* and phosphorus energy of 70 *keV* (with corresponding doses) to achieve the doping profile. Now, the implantation energy of boron varies from 35 *keV* to 155 *keV*, with 60 *keV* of energy less and more than the originally used 95 *keV*, where boron's role in sensor performance is evaluated. Figure 6.1 illustrates the net doping concentration profile against depth for various boron energies. It is observed that increasing the implantation energy results in a deeper junction, as expected. However, an unusual behavior is observed at the lowest energy setting of 35 *keV*, where the junction appears deeper than anticipated, as lower implantation energies produce shallower junctions. In this case, due to the low energy of boron implantation, the boron profile (p+) completely overlaps with the phosphorus implant, causing the

p-n junction to form between the n+ region and p-epitaxial, rather than between the intended p+ and n+ regions.

The energies between 85 *keV* and 125 *keV* keep the distribution similar to that achieved at 95 *keV*. At 155 *keV*, the junction is a bit deeper, which might lessen the sensor's ability to detect blue light efficiently due to the increased depth of light absorption.

Figure 6.2 plots the junction depths against the energy used for implantation, extracted from the doping concentration profile. This graph is particularly useful to visualize where exactly the junction depth is situated and how it varies with changes in implantation energy. It is clearly seen that for values between 85 *keV* and 125 *keV*, the junction depth is almost similar to that at 95 *keV*, consistently lying around -2.45 μm.

6.1.2 Electric field distribution profile



(a) At 35 keV.

(b) At 65 keV.

(c) At 85 keV.

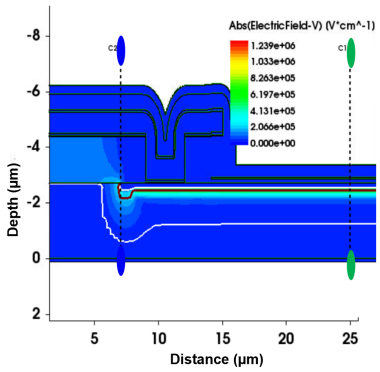
Figure 6.3. Electric field profile at lower energies.

The electric field distribution within a sensor is a critical factor in determining the breakdown voltage and minimizing the edge effects. It is primarily influenced by the doping

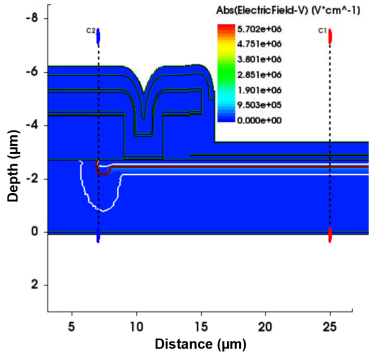
concentration, which varies with different boron implantation energies.

Figure 6.3 displays the electric field profile along different cuts across the sensor for energies less than 95 keV to examine the distribution keeping the same overvoltage.

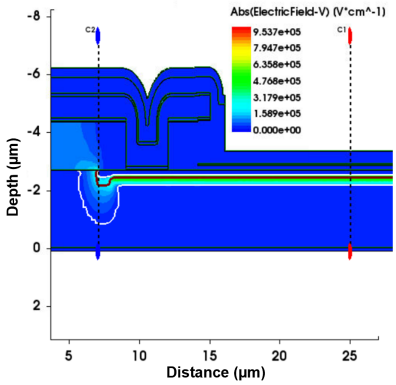
For the lower energies of 35 keV, 65 keV, and 85 keV, where the doping extends laterally more than usual, cuts are made at three distinct points (C1, C2, and an additional C3) to observe how the electric field behaves with such expansion. For higher energies, more than 85 keV, the electric field profile is shown in Figure 6.4 by keeping the same overvoltage, where the cuts are made at two distinct points (C1 and C2), corresponding to not too much lateral depletion depth. The distribution along the cuts is depicted in Figure 6.5.



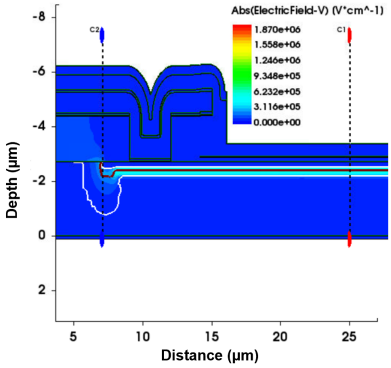
(a) At 95 keV.



(b) At 105 keV.



(c) At 125 keV.



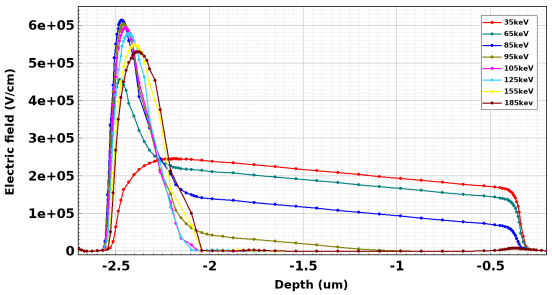
(d) At 155 keV.

Figure 6.4. Electric field profile at higher energies.

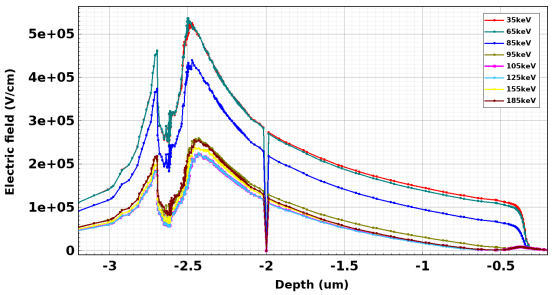
Figure 6.5 compares the electric fields at the different cuts made in Figures 6.3 and 6.4 across varying energies. From the cut C1, at the middle, the plot reveals that at 35 *keV*, the electric field in the middle is considerably lower, and the implantation depth is deeper, leading to a thicker depletion zone. As the energy increases to 65 and 85 *keV*, the depletion zone remains thick, which is not ideal. However, starting from 95 *keV* up to 125 *keV*,

the depletion zone thickness decreases, and the electric field strength begins to align more optimally within the central region. Beyond 125 *keV*, however, the depletion zone begins to increase and the peak electric field decreases again.

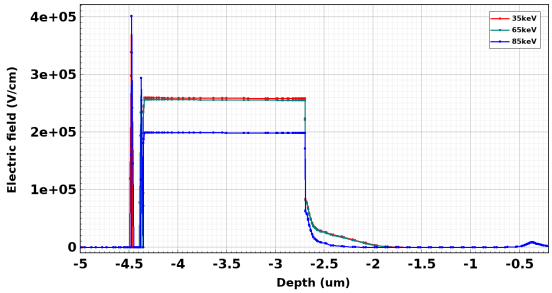
From the cuts C2 and C3, it can be observed that for lower energies, the electric field is higher at the edges rather than the middle, which could potentially lead to detrimental edge effects. As the implantation energy increases to 85 *keV*, the peak electric field begins to stabilize more towards the centre of the sensor structure (in cut C1), which is a desirable outcome for sensor performance. From 95 *keV* onwards the electric field profile shows a pronounced peak in the middle, significantly reducing the field intensity at the edges.



(a) Cut C1 from Figures 6.3 and 6.4.



(b) Cut C2 from Figures 6.3 and 6.4.



(c) Cut C3 from Figures 6.3 and 6.4.

Figure 6.5. Electric field distribution at different boron implantation energies where the cuts are made from Figures 6.3 and 6.4.

The analysis indicates that implantation energies around 95 *keV*, 105 *keV*, and 125 *keV* are the most suitable for achieving a balanced electric field distribution and maintaining an optimal depletion zone thickness.

6.1.3 Current-Voltage characteristics

In evaluating the effects of boron implantation energies on sensor performance, the simulated current-voltage characteristics across a range of energies from 35 *keV* to 185 *keV* are analyzed. This study aimed to understand how different energy levels influence the sen-

sor's breakdown voltage and overall electrical behavior. Figure 6.6 presents the IV curves for different boron energies, highlighting that 105 *keV* and 125 *keV* are particularly effective, highly converging to the electrical equations with a smooth behavior, more stable, and producing favorable IV curve slopes.

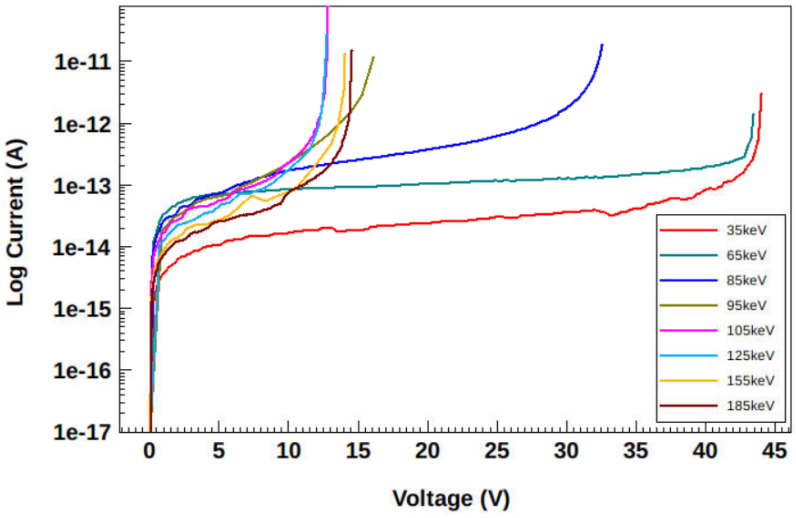


Figure 6.6. Simulated reverse IV curve at different boron energies.

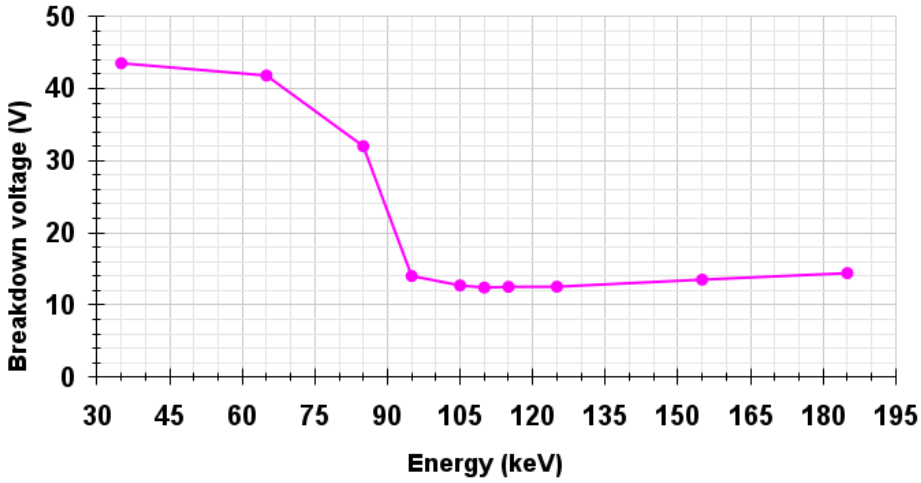
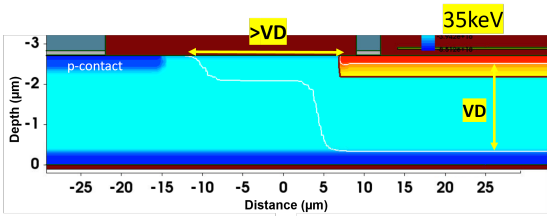


Figure 6.7. Breakdown voltage at different boron energies.

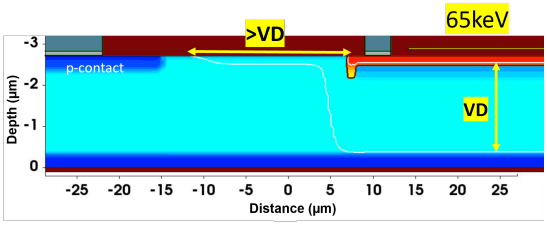
The breakdown voltages for these boron implantation energies are depicted in Figure 6.7. A trend is observed where breakdown voltage decreases with increasing energy up to 105 *keV*, then rises slightly at higher energies, although not as significantly as at lower energies. Energies up to 85 *keV* maintain breakdown voltages above 30 *V*, aligning for 35 *keV* with laboratory measurements of the sensors, while energies above 85 *keV* drop below 15 *V*. Therefore, the optimal range to choose from this analysis is between 95 *keV* and 125 *keV*, which shows a lower breakdown voltage trend of less than 30 *V*.

6.1.4 Space charge region lateral distance

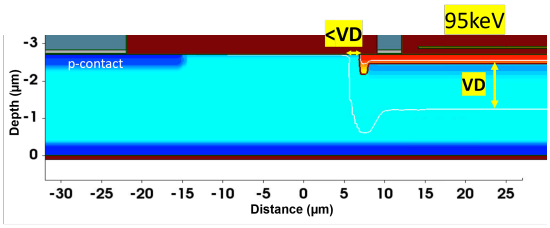
The process of ion implantation significantly impacts the lateral expansion of the depletion region. The lateral depletion width should be minimized in comparison to the vertical depletion width (VD) [62]. This design princple is essential to prevent unintended avalanche triggering at the lateral edges and to maintain an efficient multiplication zone.



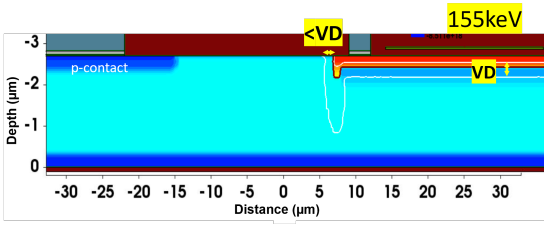
(a) At 35 keV of boron implantation energy.



(b) At 65 keV of boron implantation energy.



(c) At 95 keV of boron implantation energy.



(d) At 155 keV of boron implantation energy.

Figure 6.8. Comparison of space charge region lateral distance at low and high boron implantation energies at constant overvoltage.

Figure 6.8 examines the lateral distances at both low and high implantation energies. It reveals that for lower energies, specifically 35 *keV* and 65 *keV*, the implantation is more laterally spread, leading to a broader lateral space charge region that extends significantly toward the edge of the active region. This can introduce the non-uniform electric field distribution, especially near the junction corners, which are known to be susceptible to edge breakdown if not properly implanted. This lateral spread risks reaching into non-sensitive edge regions

In contrast, higher energies such as 95 *keV* and 155 *keV* produce deeper and more vertically confined profiles. As a result, the lateral spread of the space charge region is sharply reduced. The simulations show that in these cases, the lateral depletion width is limited to a narrow zone near the edge of the junction, preserving a uniform high-field region vertically, also explained in Section 6.1.2, where avalanche breakdown is intended to occur.

Another parameter called the aperture ratio [62], defined as the proportion of the photo detector's surface area that is sensitive to incident light, is a key metric in determining how much of a pixel contributes to actual pixel detection. A large lateral depletion region, as seen at lower implantation energies, reduces the light-sensitive area, thus lowering the aperture

ratio. The high aperture ratio is achieved effectively at higher energies (between 95 *keV* and 155 *keV*), where the lateral spread is compact.

This observation indicates that lower energy levels do not achieve the ideal doping profile necessary for optimal sensor performance, thus suggesting a shift toward higher energies in subsequent production cycles to ensure the space charge region is appropriately constrained.

6.2 Finalized boron implantation energies

Based on comprehensive simulations and tests including doping concentration profiles, IV characteristics, breakdown voltages, electric field distributions, and the lateral expansion of the space charge region, it becomes evident that lower implantation energies, particularly those below 95 *keV*, are not optimal for achieving expected outcomes.

At 35 keV, the doping profile significantly deviates from expectations, showing unusual behaviors such as high electric fields at the edges of the junction. These observations coincide with edge breakdown issues noted during lab measurements, indicating that such low energy levels might be used in production. The IV curve data at this energy level indicate a breakdown voltage of around 40 *V*, closely matching laboratory measurements, which also recorded similar values of around 42 *V*. This agreement between simulated and experimental data further implies that the 35 *keV* boron implantation energy might be the reason for the malfunctioning of the junction in the fabricated sensors.

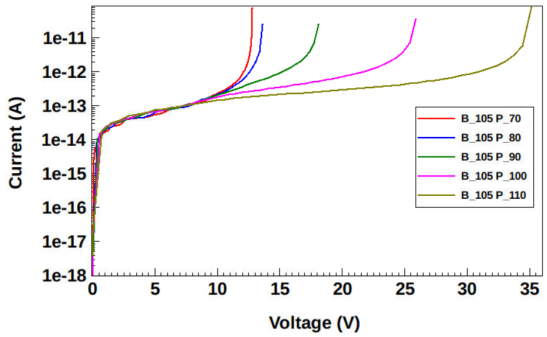
Additionally, energies such as 65 *keV* and 85 *keV* also failed to meet the performance criteria based on the analyses. Therefore, higher energies, specifically ranging from 95 *keV* to 125 *keV*, have been identified as optimal. This range of implantation energies creates effective outcomes in terms of electric field distribution, breakdown voltage, and doping depth that align with the requirements.

6.3 Impact of phosphorus (n+) implantation energies

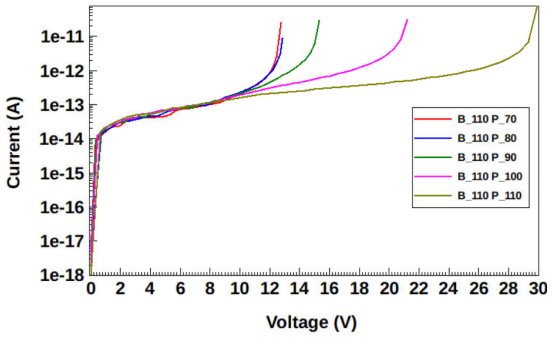
Understanding how implantation energies for phosphorus (n-type dopant) and boron (p-type dopant) affect the current-voltage characteristics and the breakdown behavior is important for optimizing sensor performance. In the following, two scenarios are analyzed: one in which the phosphorus implant energy is varied (with boron constant) and another in which boron implant energy is varied (with phosphorus kept constant). In each case, the effects on the IV curve, particularly the breakdown voltage, are discussed.

B (keV)	P (keV)	B (keV)	P (keV)	B (keV)	P (keV)
105	70	110	70	115	70
105	80	110	80	115	80
105	90	110	90	115	90
105	100	110	100	115	100
105	110	110	110	115	110

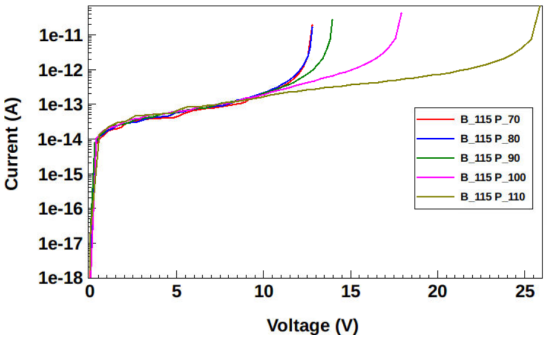
Table 6.1. Variations considered for analysing IV behavior for the first scenario.



(a) With B = 105 *keV* and P = variation.



(b) With B = 110 *keV* and P = variation.



(c) With B = 115 *keV* and P= variation.

Figure 6.9. IV curve at different phosphorus energies by keeping boron constant.

In the first scenario, the boron implantation parameters are held constant while the phosphorus implant energy is varied, where the chosen variations are shown in Table 6.1. Phosphorus, as an n-type dopant, primarily defines the depth and profile of the n-type side of the junction. Increasing the phosphorus implantation energy consistently leads to a higher breakdown voltage, as seen in the simulated IV characteristics in Figure 6.9.

A lower phosphorus energy keeps the n-type dopant closer to the surface, creating a very abrupt and highly doped junction region. That sharper junction produces higher electric field intensity for a given voltage, so the device will avalanche at a lower reverse voltage, i.e., a reduced breakdown voltage.

This expected behavior is confirmed by simulation, for example, devices with phosphorus implant energies of 100–110 keV exhibited noticeably higher breakdown voltages than those with 70 *keV*, under identical boron doping conditions.

P (keV)	B (keV)	P (keV)	B (keV)	P (keV)	B (keV)
70	105	80	105	90	105
70	110	80	110	90	110
70	115	80	115	90	115

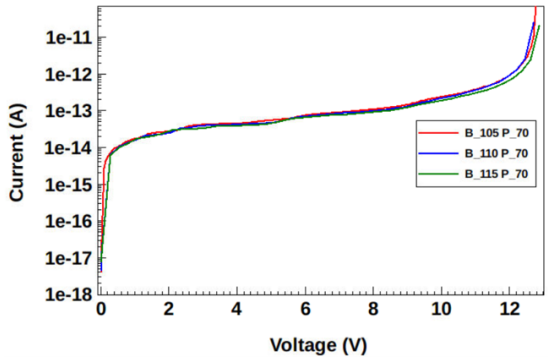
P (keV)	B (keV)	P (keV)	B (keV)
100	105	110	105
100	110	110	110
100	115	110	115

Table 6.2. Variations considered for analysing IV behavior for the second scenario.

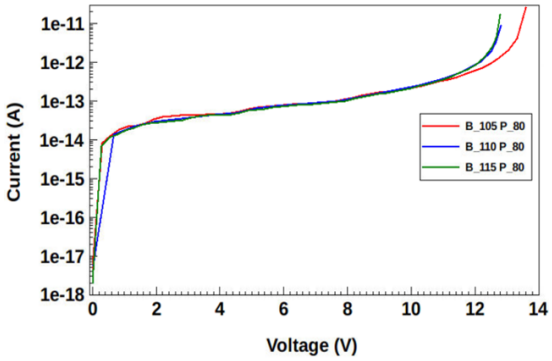
In the second scenario, when the phosphorus implantation energy is held fixed and the boron implant energy is varied, the depth and concentration of the p-type region change, as shown in Table 6.2. This has a noticeable impact on the reverse IV characteristics in Figure 6.10, especially the breakdown point. Increasing the boron implantation energy (making the p-type junction deeper) causes the breakdown voltage to decrease.

Simulations showed that at higher phosphorus implantation energy, the effect of varying the boron energy becomes more pronounced. In simple words, when the n-type side is deep, the p-side profile can dominate where the junction fields peak.

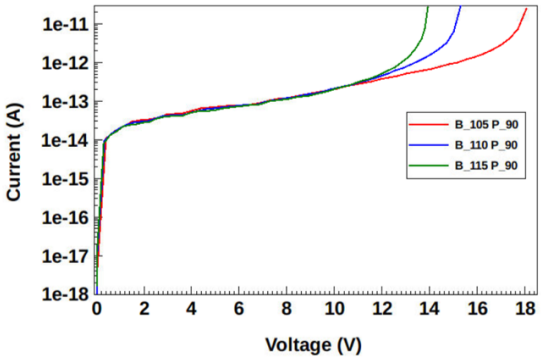
For example, at a phosphorus energy of 110 *keV*, a shift in the boron energy from 105 *keV* to 115 *keV* might produce a large change in the breakdown voltage (since the n-side is not the limiting factor), whereas at a shallow phosphorus energy (70 *keV*), the breakdown voltage is low and dominated by the n-side, so tweaking the boron depth has a smaller effect. This insight emphasizes that co-optimization of both implants is required. The best sensor performance is achieved when both the p+ and n+ implants are tuned such that the electric field is uniformly distributed.



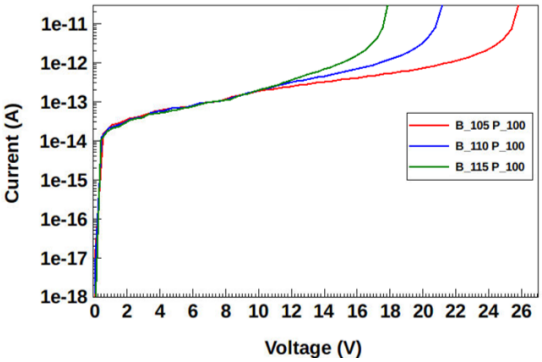
(a) With P = 70 *keV* and B = variation.



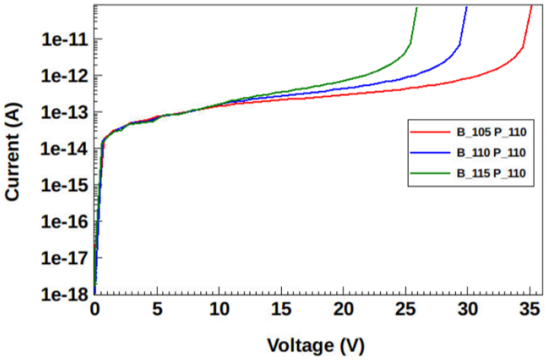
(b) With P = 80 *keV* and B = variation.



(c) With P = 90 *keV* and B = variation.



(d) With P = 100 *keV* and B = variaton.



(e) With P = 110 *keV* and B = variation.

Figure 6.10. IV curve at different phosphorus energies by keeping boron constant.

6.4 Final simulated implantation parameters for the production

To finalize the optimal doping parameters for the sensor's production, a comprehensive analysis of boron (B) and phosphorus (P) implantation energies was conducted. This analysis aimed to identify the best combinations that balance breakdown voltage performance with sensor functionality. Figure 6.11 illustrates the breakdown voltages calculated for various

combinations of boron and phosphorus energies.

The parameter sets are divided into four categories based on constant boron energy levels. Each section explores different phosphorus energies to understand their impact on the sensor's breakdown voltage.

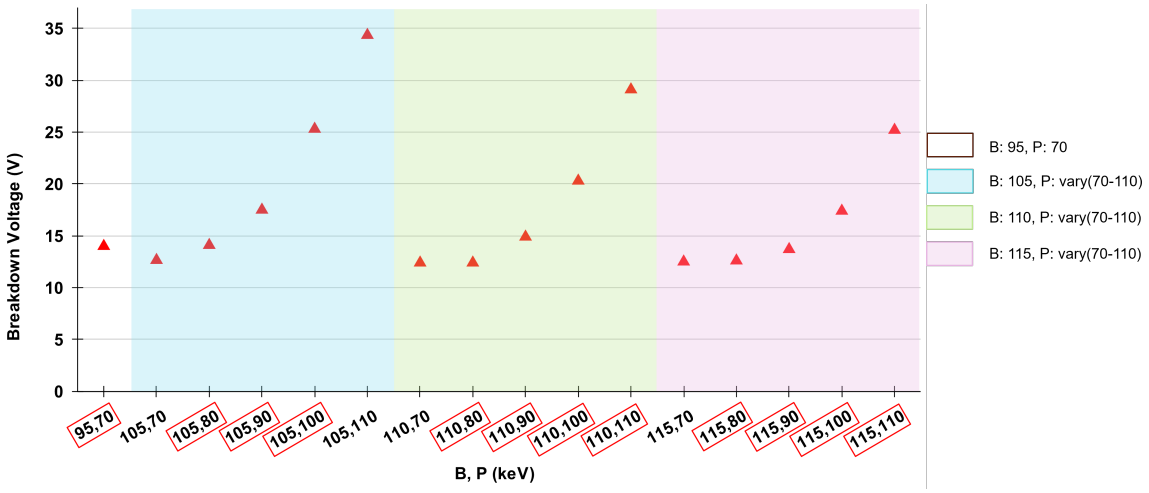


Figure 6.11. Comparison of breakdown voltages for all the tested implantation energies of boron and phosphorus. Chosen parameter sets are highlighted with red boxes.

1. B = 95 *keV* and P = 70 *keV* (white section): This original set, also a reference parameter set, is chosen to be retested due to unsatisfactory performance in initial sensor batches, aiming to verify if repeating the parameters would yield different results.

2. B = 105 *keV* with variations in P (blue section): Within this category, five phosphorus energy variations are tested, and B = 105 *keV*, P = 70 *keV*, and B = 105 *keV*, P = 110 *keV* are excluded. The exclusion of P = 70 *keV* is due to its similarity in analysis results to the original parameters. P = 110 *keV* is excluded because the phosphorus energy surpasses boron, which theoretically could disrupt the junction balance, and its breakdown voltage of 35 *V* is higher than desired.

3. B = 110 *keV* with variations in P (green section): Similarly, in this set, five variations in phosphorus energy are tested, and the combination of B = 110 *keV* and P = 70 *keV* is excluded due to their similarity to the base parameters. The remaining phosphorus variations offer a broader range of breakdown voltages and the potential for improved performance.

4. B = 110 *keV* with variations in P (pink section): Again, in this set, five phosphorus energy variations are tested, and the combination of B = 115 *keV* and P = 70 *keV* is also excluded for reasons similar to the other sections, focusing on higher phosphorus energies to potentially enhance performance.

The selected parameter sets are the highlighted red boxes on the x-axis that cover a range of breakdown voltages from 13 V to 30 V. The non-highlighted parameter sets are excluded due to their theoretical inefficacy or duplication of previous results.

Due to the limited number of wafers available, only 12 variations of boron and phosphorus implantation energies can be tested. This constraint necessitates the exclusion of certain combinations that either replicate previous results or are considered infeasible based on theoretical and practical considerations. This range provides a diverse array of configurations to test and analyze, aiming to identify the most effective settings for optimal sensor performance in low-light detection applications.

6.5 New wafer layout

With the new parameters selected, there has also been a redesign of the wafer layout depicted in Figure 6.12. Previously, as detailed in Chapter 5, the configurations included flip-chips and wire-bonded variants with 1×1, 2×2, and 4×4 pixel(s) arrangements, with guard rings and with or without anti-reflective layers.

No. of pixels	Covered / Exposed	GR	Corner radius	Designation
1/4	C/E	$0.5\mu m$	$2\mu m$, 2×GR, 4×GR	1C05R2, 4C05R2, 1E05R2, 4E05R2, 1C05C2, 4C05C2, 1E05C2, 4E05C2, 1C05C4, 4C05C4, 1E05C4, 4E05C4
1/4	C/E	$1.0\mu m$	$2\mu m$, 2×GR, 4×GR	1C10R2, 4C10R2, 1E10R2, 4E10R2, 1C10C2, 4C10C2, 1E10C2, 4E10C2, 1C10C4, 4C10C4, 1E10C4, 4E10C4
1/4	C/E	$1.5\mu m$	$2\mu m$, 2×GR, 4×GR	1C15R2, 4C15R2, 1E15R2, 4E15R2, 1C15C2, 4C15C2, 1E15C2, 4E15C2, 1C15C4, 4C15C4, 1E15C4, 4E15C4
1/4	C/E	$2.0\mu m$	$2\mu m$, 2×GR, 4×GR	1C20R2, 4C20R2, 1E20R2, 4E20R2, 1C20C2, 4C20C2, 1E20C2, 4E20C2, 1C20C4, 4C20C4, 1E20C4, 4E20C4

Table 6.3. Different configurations of sensors.

However, the updated layout eliminates the flip chip options, focusing solely on wire-

bonded variations configured in 1×1 and 4×4 pixel(s) matrices. The different combinations are listed in Table 6.3, where each row specifies the number of pixels, whether the pixel is covered or exposed (with or without window), the guard ring width, and the corner radius of the pixel. The last column lists the designation codes for each configuration. The dimensions of the reticle and sensor remain unchanged at 9.6 $\times$ 9.6 mm^2 and 1 mm^2.

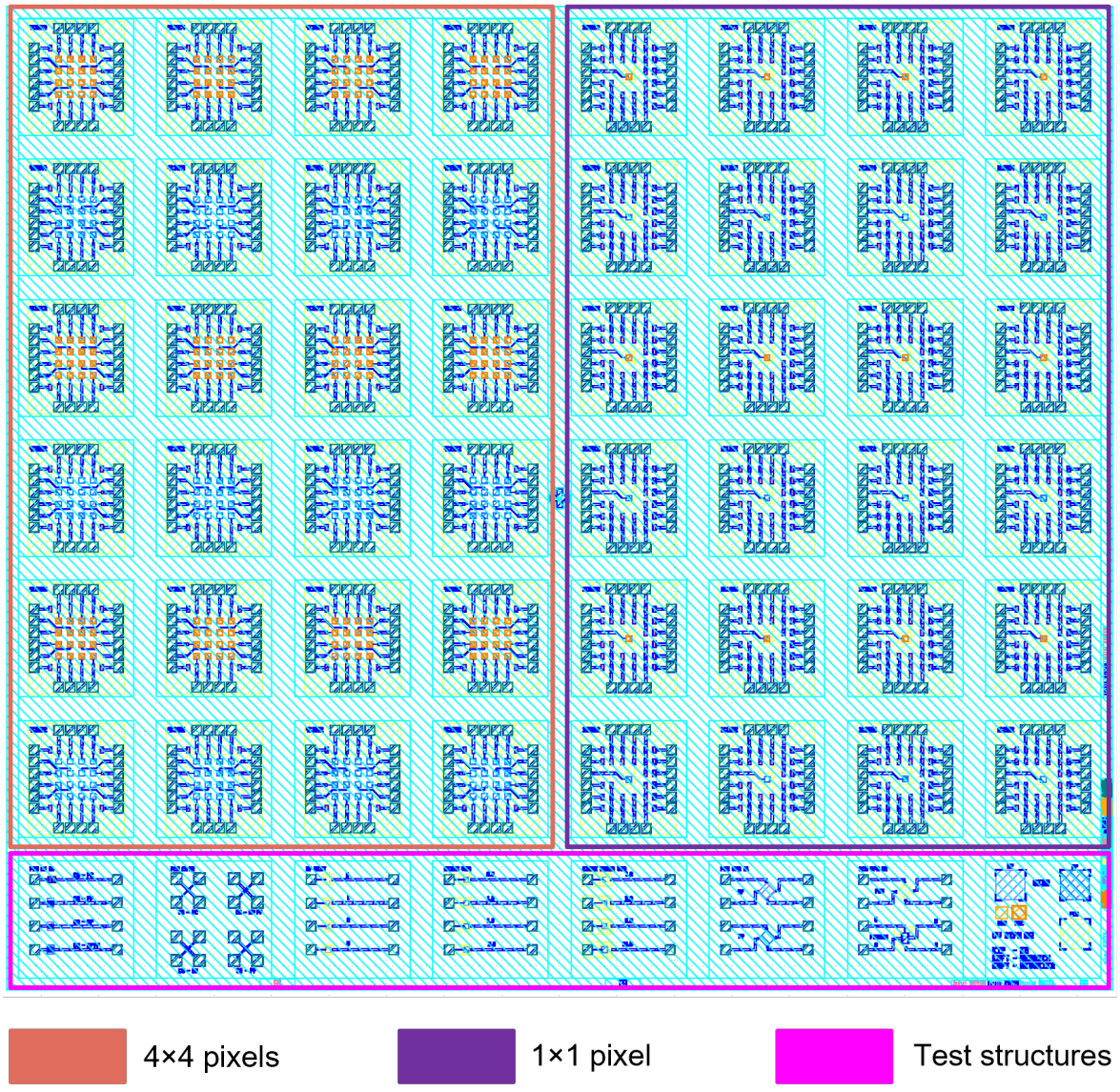
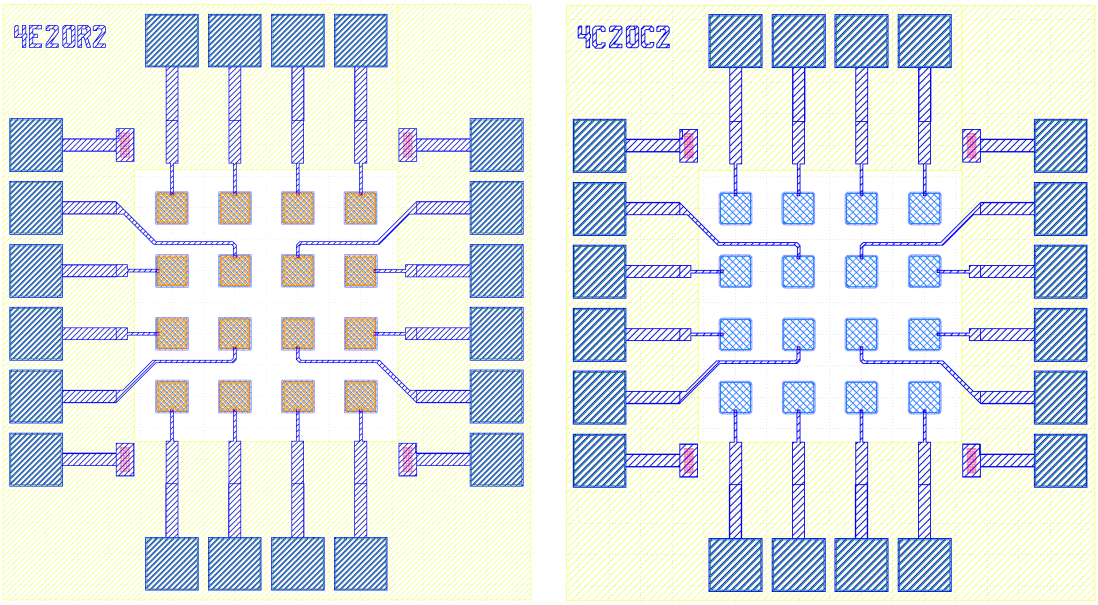


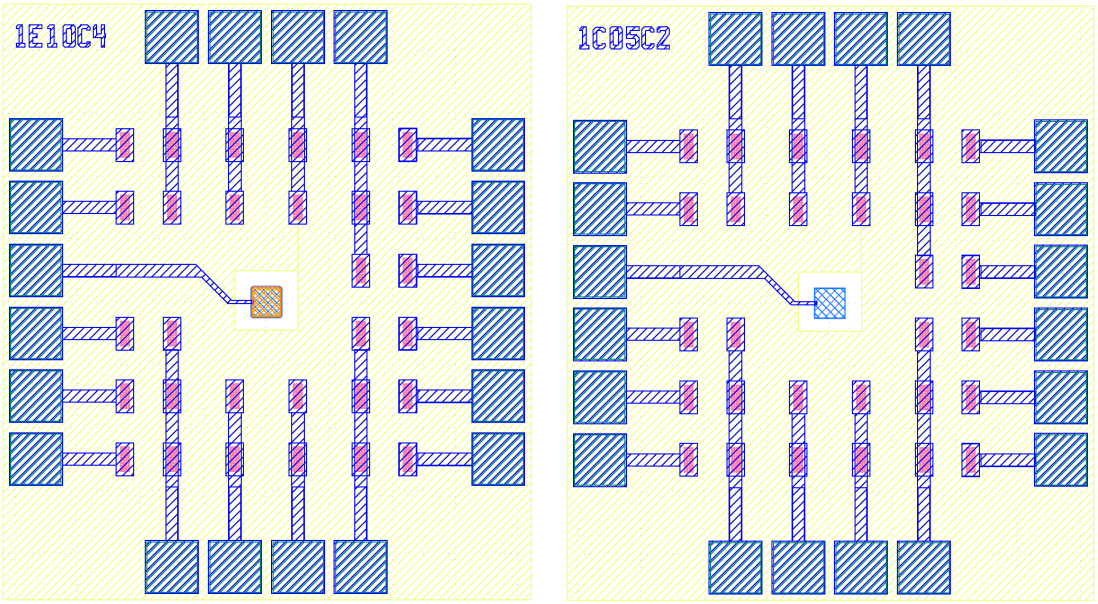
Figure 6.12. New layout of reticle showing 1×1, 4×4 pixel matrices and test structures at the bottom.

Figure 6.13 presents the GDS layout of a 4×4 pixel variant with a variation of presence or absence of window opening and Figure 6.14 shows the same variation in a 1×1 pixel.



(a) Type 4E20R2 (window opening). **(b)** Type 4C20C2 (no window opening).

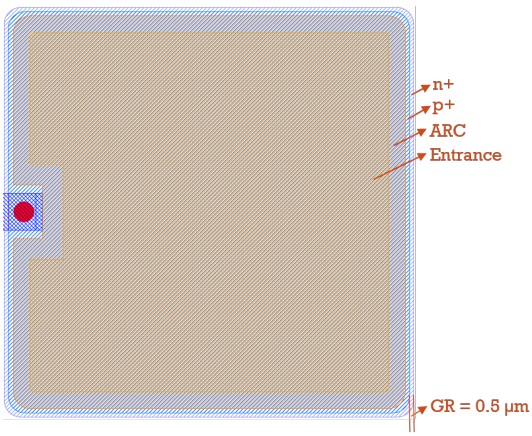
Figure 6.13. GDS structure of 4×4 pixel(s) matrices in the presence and absence of the window opening.



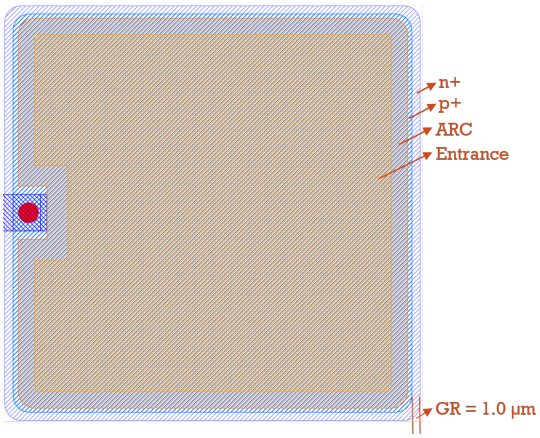
(a) Type 1E10C4 (window opening). **(b)** Type 1C05C2 (no window opening).

Figure 6.14. GDS structure of 1×1 pixel in the presence and absence of the window opening.

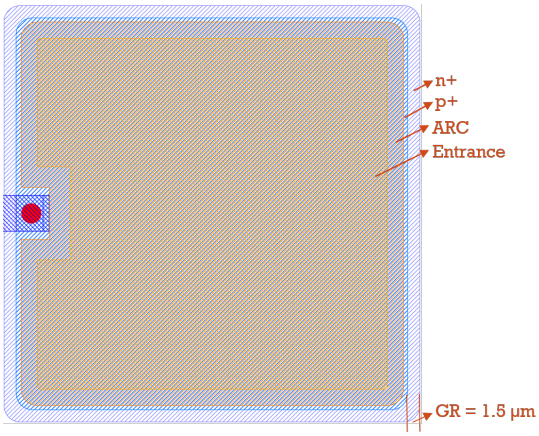
The four variations in guard ring widths (0.5 μm, 1.0 μm, 1.5 μm, and 2.0 μm) are shown in Figure 6.15, where different layers of the pixel are also shown.



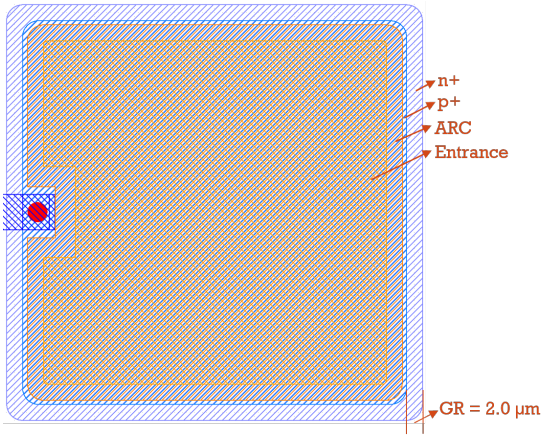
(a) Type 1E05R2 (GR 0.5 μm).



(b) Type 1E10R2(GR 1.0 μm).

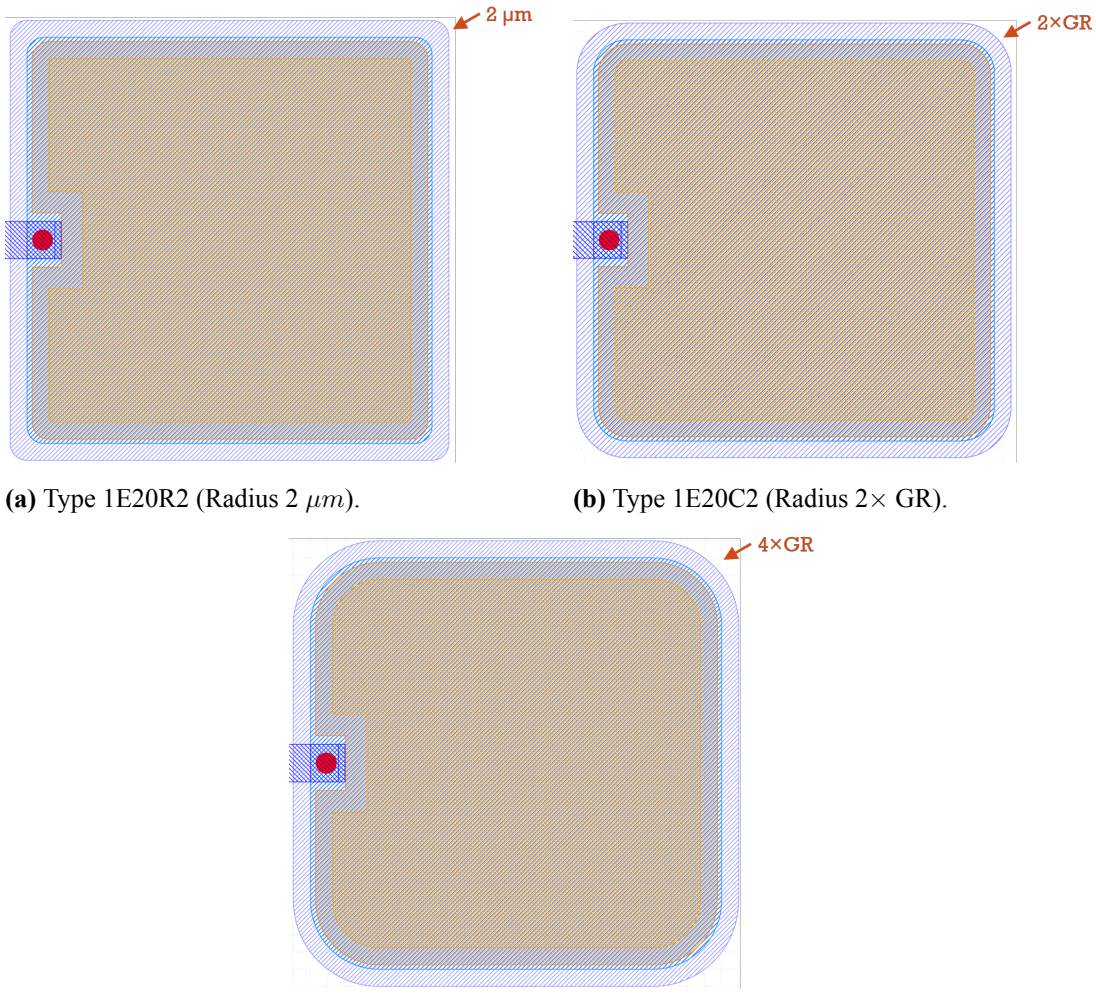


(c) Type 1E15R2 (GR 1.5 μm).



(d) Type 1E20R2 (GR 2.0 μm).

Figure 6.15. GDS structure showing variations in guard ring (GR) structures.



(a) Type 1E20R2 (Radius 2 μm).

(b) Type 1E20C2 (Radius 2× GR).

(c) Type 1E20C4 (Radius 4× GR).

Figure 6.16. GDS structure of active area showing different corner roundings.

The new reticle variation incorporates three different corner roundings to minimize edge effects. These variations include 2 μm radius, twice the guard ring width, and four times the guard ring width depicted in Figure 6.16.

This variety in design parameters across the reticle adaptations aims to assess and optimize sensor performance under different geometric conditions within the manufacturing process.

Chapter 7

Summary

This thesis aimed at the development and optimization of a novel SPAD designed for enhanced detection of blue and NUV light, using SOI wafer technology. Through the combination of simulation, development, and testing, the work demonstrates a complete development cycle for a SPAD sensor, from initial concept to refined design.

The initial phase of the research involved simulating the design and confirming the feasibility of the SPAD using TCAD Synopsys tools. The design involved precise control over the doping concentrations and depths to create a thin SPAD, a p-n junction with high and uniform electric fields using virtual guard rings to avoid premature breakdown. The simulation configured a shallow p-n junction optimized for sensitivity to the blue and NUV wavelengths. Optical simulations were also refined to maximize photon absorption and minimize reflections from the surface, thus boosting specified light sensitivity. The simulations resulted in a breakdown voltage of 15.25 V and a photon detection probability of 44 % at 405 nm of wavelength at an overvoltage of 1 V.

Following the simulation phase, after development, the fabricated SPADs were tested under various conditions to verify the simulation predictions. Initial tests confirmed the sensors' functionality to operate in Geiger mode with an effective passive quenching mechanism, thus providing a "proof of concept" validation of the design. However, more experimental results revealed a significant discrepancy between simulation and measurement. In particular, the measured breakdown voltage of the SPAD was on the order of 40 V, much higher than the 15.25 V predicted by simulation. Dynamic characterization revealed high dark counts, even at lower temperatures with mini-breakdown pulses, indicating an early edge breakdown and non-uniformities of an electric field in the active region of SPAD despite the presence of guard rings. Further tests, such as LET and SIMS, were employed. LET pinpointed that the corner regions of the pixel were initiating the early avalanche breakdown, while SIMS provided a measurement of the doping profiles in the fabricated structures, revealing a mismatch in the doping profiles of the simulation and fabricated sensors. In other

words, the actual ion implantation dose and energy parameters in the fabricated device differed from those predicted in the simulation, leading to a higher breakdown voltage.

With the knowledge of the doping discrepancy, adjustments were made to the doping parameters, and the device design was updated. Simulations of the revised doping profile with a modification in the corner geometry of SPAD indicated that these refinements would align the breakdown and electric field behaviors closely with initial expectations, following 12 variations of doping profile parameters submission. These measures aim to eliminate edge-related early breakdowns and ensure a uniform electric field distribution across the active area, thereby enhancing device reliability. This iterative process of simulation, measurement, and redesign represents an outcome of the thesis. The development of sensors with the newly submitted doping parameters is currently underway, with expectations set for the successful validation of the improved structure.

List of Figures

List of Tables

Bibliography

[1] Andreino Simonelli. Photomultiplier tubes, evolution uses and future perspectives. Presentation, DRD4 WG1 Meeting on PMTs.

[2] R. H. Haitz, A. Goetzberger, R. M. Scarlett, and W. Shockley. Avalanche effects in silicon p-n junctions. *Journal of Applied Physics*, 34(6), Jun 1963.

[3] Massimo Ghioni, Angelo Gulinatti, Ivan Rech, Franco Zappa, and Sergio Cova. Progress in silicon single-photon avalanche diodes. *IEEE Journal of Selected Topics in Quantum Electronics*, 13(4):852–862, 2007.

[4] O. Stenzel. *Light-matter interaction*. Springer, 2022.

[5] G.F. Knoll. *Radiation Detection and Measurement*. John Wiley Sons, INC., USA, 3 edition, 2000.

[6] R. D. Evans. *The Atomic Nucleus*. Tata McGraw-Hill, New Delhi, 1955.

[7] Claude Leroy and Pier-Giorgio Rancoita. *Silicon Solid State Devices and Radiation Detection*. World Scientific Publishing Co. Pte. Ltd.., Singapore, 2012.

[8] C. Kittel. *Introduction to Solid State Physics*. John Wiley Sons, INC., USA, 8 edition, 2005.

[9] Paul A. Tipler. *Physics for Scientists and Engineers*. W. H. Freeman and Company, New York, 5 edition, 2004.

[10] Donald A. Neamen. *Semiconductor Physics and Devices: Basic Principles*. McGraw-Hill, New York, NY, 4 edition, 2012.

[11] Andrew S. Grove. *Physics and Technology of Semiconductor Devices*. John Wiley & Sons, Inc., New York, wiley international edition edition, 1967.

[12] S. M. Sze. *Semiconductor Devices: Physics and technology*. John Wiley Sons, INC., USA, 2 edition, 2002.

[13] S. O. Kasap. *Principles of electronic materials and devices.* McGraw Hill, New York, 4 edition, 2018.

[14] H. Kolanoski and N. Wermes. *Particle Detectors: Fundamentals and Applications*. Oxford University Press, 2020.

[15] S. O. Kasap. *Principles of electronic materials and devices*. McGraw-Hill, New York, 3 edition, 2007.

[16] A. Lacaita, F. Zappa, S. Bigliardi, and M. Manfredi. On the bremsstrahlung origin of hot-carrier-induced photons in silicon devices. *IEEE Transactions on Electron Devices*, 40(3):577–582, 1993.

[17] F. Zappa et al. Principles and features of single-photon avalanche diode arrays. *Sensors & Actuators*, 140(1):103–112, 2007.

[18] X. Qian, W. Jiang, A. Elsharabasy, and M. J. Deen. Modeling for single-photon avalanche diodes: State-of-the-art and research challenges. *Sensors (Basel, Switzerland)*, 23(7):3412, 2023.

[19] A. G. Chynoweth. Ionization rates for electrons and holes in silicon. *Physical Review*, 109(5):1537–1540, Mar 1958.

[20] E. Charbon, M. Fishburn, R. Walker, R. K. Henderson, and C. Niclass. SPAD Based Sensors. In F. Remondino and D. Stoppa, editors, *TOF Range-Imaging Cameras*, volume 9783642275, pages 11–38. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg, 2013.

[21] Claudio Piemonte. A new silicon photomultiplier structure for blue light detection. *Nuclear Instruments and Methods in Physics Research Section A*, 568:224–232, 2006.

[22] M. M. Vignetti, F. Calmon, P. Lesieur, F. Dubois, T. Graziosi, and A. Savoy-Navarro. A novel 3d pixel concept for geiger-mode detection in soi technology. pages 166–169, 2016.

[23] Pengfei Sun, Edoardo Charbon, and Ryoichi Ishihara. A flexible ultrathin-body single-photon avalanche diode with dual-side illumination. *IEEE Transactions on Electron Devices*, 64(3):1124–1128, Mar. 2017.

[24] D. N. Yaung, B. C. Hsieh, C. C. Wang, and J. C. Liu. High performance 300 mm backside illumination technology for continuous pixel shrinkage. In *Proc. International Electron Devices Meeting (IEDM)*, pages 175–178, 2011.

[25] Lin Qi, K. R. C. Mok, Mahdi Aminian, Edoardo Charbon, and Lis K. Nanver. Uv-sensitive low dark-count pureb single-photon avalanche diode. *IEEE Transactions on Electron Devices*, 61(11):3768, Nov 2014.

[26] M. S. Oh, H. K. Kong, and H. S. Lee. Backside-illumination 14 μm pixel qvga time-of-flight cmos imager. In *Proc. IEEE 10th International New Circuits and Systems Conference (NEWCAS)*, pages 325–328, 2012.

[27] Myung-Jae Lee, Augusto Ronchini Ximenes, Preethi Padmanabhan, Tzu-Jui Wang, Kuo-Chin Huang, Yuichiro Yamashita, Dun-Nian Yaung, and Edoardo Charbon. High-performance back-illuminated three-dimensional stacked single-photon avalanche diode implemented in 45-nm cmos technology. *IEEE Journal of Selected Topics in Quantum Electronics*, 24(6):1–9, 2018.

[28] Myung-Jae Lee, Pengfei Sun, Gregory Pandraud, Claudio Bruschini, and Edoardo Charbon. First near-ultraviolet- and blue-enhanced backside-illuminated single-photon avalanche diode based on standard soi cmos technology. *IEEE Journal of Selected Topics in Quantum Electronics*, 25(5):1–6, 2019.

[29] T. Al Abbas, N. A. W. Dutton, O. Almer, S. Pellegrini, Y. Henrion, and R. K. Henderson. Backside illuminated spad image sensor with 7.83µm pitch in 3d-stacked cmos technology. *2016 IEEE International Electron Devices Meeting (IEDM)*, pages 8.1.1–8.1.4, 2016.

[30] L. Baudis, M. Galloway, A. Kish, C. Marentini, and J. Wulf. Characterisation of silicon photomultipliers for liquid xenon detectors. *Journal of Instrumentation*, 13:P10022, Oct 2018.

[31] F. Carnesecchi. Light detection in darkside-20k. *Journal of Instrumentation*, 15:C03038, 2020.

[32] Alessandro Razeto and Nicola Rossi. Challenges for dark matter direct search with sipms. *Frontiers in Physics*, 11, 2023.

[33] J. Aalbers et al. Darwin: towards the ultimate dark matter detector. *Journal of Physics G: Nuclear and Particle Physics*, 50(1):013001, 2023.

[34] Paulo Crespo, Georgy Shakirin, Fine Fiedler, Wolfgang Enghardt, and Andreas Wagner. Direct time-of-flight for quantitative, real-time in-beam pet: A concept and feasibility study. *Physics in Medicine & Biology*, 52(23):6795, 2007.

[35] A. N. Otte, J. Barral, B. Dolgoshein, J. Hose, S. Klemin, E. Lorenz, et al. A test of silicon photomultipliers as readout for pet. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 545(3):705–715, 2005.

[36] M. G. Bisogni, A. Del Guerra, and N. Belcari. Medical applications of silicon photomultipliers. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 926:118–128, 2019.

[37] Synopsys. *Sentaurus Process User Guide*. Synopsys. `https://www.synopsys.com/manufacturing/tcad/process-simulation.html`.

[38] Synopsys. *Sentaurus Device User Guide*. Synopsys. `https://www.synopsys.com/manufacturing/tcad/device-simulation.html`.

[39] C.-T. Sah, R. Noyce, and W. Shockley. Carrier generation and recombination in p-n junctions and p-n junction characteristics. *Proceedings of the IRE*, 45(9):1228–1243, Sep. 1957.

[40] W. Shockley and W. T. Read. Statistics of the recombinations of holes and electrons. *Physical Review*, 87(5):835–842, Sep. 1952.

[41] S. M. Sze and K. K. Ng. *Physics of Semiconductor Devices*. John Wiley & Sons, Inc., Hoboken, NJ, USA, Oct. 2006.

[42] Synopsys. *Sentaurus Visual User Guide*. Synopsys. `https://www.synopsys.com/manufacturing/tcad/calibration-and-services.html`.

[43] J. P. Colinge and C. A. Colinge. *Physics of Semiconductor Devices*. Springer, New York, NY, 1 edition, 2002.

[44] 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. *Optics Express*, 23(10):13200–13209, 2015.

[45] M. Singh Tyagi. Zener and avalanche breakdown in silicon alloyed p-n junctions—ii: Effect of temperature on the reverse characteristics and criteria for distinguishing between the two breakdown mechanisms. *Solid-State Electronics*, 11:117–128, 1968.

[46] R.VanOverstraetenandH.DeMan. Measurement of the ionization rates in diffused silicon p-n junctions. *Solid-StateElectronics*, 13(5):583–608, 1970.

[47] Y. Okuto and C. Crowell. Threshold energy effect on avalanche breakdown voltage in semiconductor junctions. *Solid-State Electronics*, 18(2):161–168, Feb 1975.

[48] T. Lackner. Avalanche multiplication in semiconductors: A modification of chynoweth's law. *Solid-State Electronics*, 34(1):33–42, 1991.

[49] M. Valdinoci et al. Impact-ionization in silicon at large operating temperature. In *International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pages 27–30, Sep 1999.

[50] S. Reggiani et al. Investigation about the high-temperature impact-ionization coefficient in silicon. In *Proceedings of the 34th European Solid-State Device Research Conference (ESSDERC)*, pages 245–248, Sep 2004.

[51] M. C. Vecchi and M. Rudan. Modeling electron and hole transport with full-band structure effects by means of the spherical-harmonics expansion of the bte. *IEEE Transactions on Electron Devices*, 45(1):230–238, 1998.

[52] S. Reggiani et al. Electron and hole mobility in silicon at large operating temperatures—part i: Bulk mobility. *IEEE Transactions on Electron Devices*, 49(3):490–499, 2002.

[53] A. Panglosse, P. Martin-Gonthier, O. Marcelot, C. Virmontois, O. Saint-Pé, and P. Magnan. Modeling, simulation methods and characterization of photon detection probability in cmos-spad. *Sensors*, 21(17):5860, 2021.

[54] W.G. Oldham, R.R. Samuelson, and P. Antognetti. Triggering phenomena in avalanche diodes. *IEEE Transactions on Electron Devices*, 19(9):1056–1060, 1972.

[55] S. O. Kasap. *Optoelectronics and Photonics*. Pearson, USA, 2 edition, 2013.

[56] N. Otte. The silicon photomultiplier – a new device for high energy physics, astroparticle physics, industrial and medical applications. In *Proceedings of the SNIC Symposium*, Stanford, California, April 2006.

[57] Ferenc Nagy, Gyula Hegyesi, Gábor Kalinka, and József Molnár. A model based dc analysis of sipm breakdown voltages. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 849:55–59, 2017.

[58] G. Zhang, D. Han, C. Zhu, and X. Zhai. Turn-on and turn-off voltages of an avalanche p–n junction. *Journal of Semiconductors*, 33(9):094003, 2012.

[59] S. Cova, M. Ghioni, A. Lacaita, C. Samori, and F. Zappa. Avalanche photodiodes and quenching circuits for single-photon detection. *Applied Optics*, 35(12), April 1996.

[60] Tiancheng Zhong. *Irradiation Studies for the Mu3e Tile Detector*. Phd thesis, 2023.

[61] SGS Institut Fresenius. Sekundärionen-massenspektrometrie.

[62] Cristiano Niclass and Mineki Soga. Single photon avalanche diode with second semiconductor layer buried in epitaxial layer, February 2016. U.S. Patent.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to Prof. Dr. Hans-Christian Schultz-Coulon for giving me the opportunity to work under his supervision. His unwavering support, guidance, and encouragement have been instrumental throughout my PhD journey.

I am sincerely thankful to Prof. Dr. Peter Fischer for agreeing to be my second referee. His support and valuable feedback during the course of my doctoral studies have been highly appreciated.

My special thanks go to Dr. Wei Shen, who guided me during the initial years of my PhD. His mentorship, patience, and technical guidance helped me navigate several initial challenges and laid a solid foundation for this research.

I would also like to express my gratitude to Dr. Konrad Briggl for his support and guidance during the laboratory measurements and for sharing his expertise and helpful suggestions, which contributed to the progress of my work. I am grateful to Dr. Vera Koleva Stankova for her kindness and for supporting me in the final phase of my PhD project by giving valuable feedback. Thanks to both for reading the main chapters of my thesis and for providing insightful comments and suggestions that helped improve the quality of this dissertation.

I would also like to thank Dr. Ralf Achenbach for teaching me bonding techniques and assisting me in cleanroom practices to enrich my practical experience.

My appreciation extends to the members of the F8 and F11 groups for creating a pleasant and supportive working environment, which made my research experience truly enjoyable.

A very special thanks to my dear friend Dr. Tamasi Kar, with whom I shared the entire PhD journey. Thank you for your friendship, for all the laughter and memorable trips, and for your support and advice in both academic and personal life. It would not have been the same without you.

This thesis is dedicated to my family, whose unwavering love, trust, and support have been the foundation of my achievements. Your belief in me has always been my greatest strength and motivation.

Lastly, I would like to express my heartfelt thanks to my fiancé, Varun Dua, and to my

future in-laws for their constant support and belief in me. Varun, I am forever grateful for your love, patience, understanding, and for being my strongest support through every phase of life. You are truly the best!