Deep-submicron CMOS Single Photon Detectors and Quantum Effects

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof. ir. K. C. A. M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op dinsdag 17 mei 2011 om 12.30 uur
door

Mohammad Azim KARAMI

Master of Science Electronics
University of Tehran, Tehran, Iran
geboren te Tehran, Iran.

Dit proefschrift is goedgekeurd door de promotor:

Prof. dr. E. Charbon

Samenstelling promotiecommissie:

Rector Magnificus	voorzitter
Prof. dr. E. Charbon	Technische Universiteit Delft, promotor
Prof. P. Magnan	Institut Supérieur de l'Aéronautique et de l'Espace, Toulouse, Frankrijk
Prof. dr. ir. A.J.P. Theuwissen	Technische Universiteit Delft
Prof. dr. L.K. Nanver	Technische Universiteit Delft
Dr. H. Valk	NXP Semiconductors, Nijmegen
Dr. W. Snoeys	CERN, Geneva, Zwitserland
Prof. dr. P.J. French	Technische Universiteit Delft, reservelid

Cover design by Maziyar Pahlevan.

ISBN: 978-94-6113-035-8

Printed by Sieca Repro, Delft, the Netherlands

To my mother and the memory of my father

Acknowledgments

The praise is due to Allah, most gracious, most merciful, the lord of the worlds. A praise equal to the praise of his close angels, and his prophets.

I am greatly obligated to express my sincere thanks to my supervisor Prof. Edoardo Charbon, for his mentorship and guidance during my Ph.D. years. His patience and advises helped me a lot in my education.

With pleasure, I thank the committee members of my Ph.D., for reading my thesis and for their valuable comments.

I wish to thank Dr. Hyung-June Yoon for helping me with his great experience as a professional in the field of image sensors. I also wish to say thanks to all lab members that contributed through good advice and their good spirits to this thesis. Dr. Marek Gersbach has helped me a lot for the first SPAD farm design and Dr. Yuki Maruyama has helped me with his advice on test and process. I am also grateful to Mauro, Claudio, Matt, Matteo, Shingo, Harmen, Chokalingam, and Mahdi.

Special thanks go to Minaksie, Rosario, Laura, Antoon, Ali and Javad for their kind help with all administrative issues, software support and their friendliness.

I learned silicon device processing in clean room 100 of Dimes (Delft institute of microsystems and nanolelectronics) under the supervision of Mrs. Silvana Milosavljevic and I should express my thanks to Charles, Mario, Tom, Hugo, Peter, Wim, Cassan, Jan, Jaber, Amir and everybody who helped me running my processes in the clean-rooms.

I should thank Hozanna Miro from TNW and to Mehmet Dundar in TU Eindhoven for the time he spent helping me through plasmonic tests.

I also wish to express my gratitude to all the people who helped me in EPFL, more particularly Babak Falsafi, Chantal, Hossein, Olivier Martin, Lucio, Cristiano, Estelle, and Zhong Zhong Ni with whom I did the first 90nm technology tape-out.

I should also thank my previous supervisors including Prof. Afzali-Kusha, Prof. Faraji-Dana, Prof. Afjei who taught me the fundamentals of electronics and electromagnetics.

My close friends in Delft, and Lausanne helped me a lot with their understanding including Moosa, Mohammad, Shahab, Hossein, and Hamed, and I wish the best in their life.

It is beyond my ability to thank my parents, the way they deserve. I am extremely thankful to my mother for her love and blessings, and for her patience in these years. May God bless the soul of my father who sacrificed himself for his family. Last but not least I should thank my brother Amin who supported me during all these years with his compassion and guidance.

Delft, *Mohammad Azim Karami*
May 2011

Contents

Chapter 1

Introduction

This chapter starts with presenting the scope of the thesis, gives an overview over the main photodetector types being used in today's industrial and research applications, followed by thesis motivation and contributions.

1.1 Thesis scope

This dissertation deals with the emerging field of single-photon imaging and single-photon image sensors. To make single-photon image sensors relevant and useful in many fields, from physics to biology, from entertainment to biomedical sciences, it is necessary to implement high-performance devices.

Performance can be measured in many ways depending on the application. For example in entertainment, 3D cameras based on time-of-flight must discriminate millimeter depths in a large number of pixels, e.g. in VGA formats. Single-photon detectors in these applications must be ultra-miniaturized and capable of resolving picosecond time delays in incoming photons. In another example, in medicine, the required timing resolution is even higher to be able to distinguish, for example, gamma events in coincidence. However, in these applications, miniaturization is not the top priority.

One performance measure that common to all these applications is noise. In this thesis we look at several noise sources characteristic of single-photon detection, but we do that in the context of miniaturization to achieve detectors in low-cost, commercially available CMOS technologies.

Quantum parasitic effects are the most important causes of noise in single-photon detectors. Thus, we turned our attention to tunneling noise and Random Telegraph Signal (RTS) noise which are described in this thesis comprehensively in the context of miniaturized detectors. The quantum parasitic studies are performed by presenting fundamental equations and comparing different designs.

This dissertation also demonstrates how single-photon detectors can be fabricated in commercial deep-submicron CMOS processes, so as to keep noise in check. Miniaturization is explored with sub-130nm designs that were the first to be successfully tested and published. By implementing different structures, the geometric trade-offs involved in the design of deep-submicron SPADs are being studied.

1.2 Single- and multi photon detectors

1.2.1 PMT and micro channel plates

A photomultiplier tube (PMT) is a vacuum tube consisting of an input window, a photocathode, focusing electrodes, an electron multiplier, and an anode. Figure 1.1 shows the cross-section of a photomultiplier [1]. A PMT is based on the external photoelectric effect, whereby the electrons in the valence band are emitted to the vacuum by absorbing energy of the photons. The photocathode is usually made of an alkali metal or from III-V compounds to ensure that photoelectrons generated on the cathode have enough energy to leave the cathode. The workfunction is defined as the difference between Fermi level of material and vacuum level.

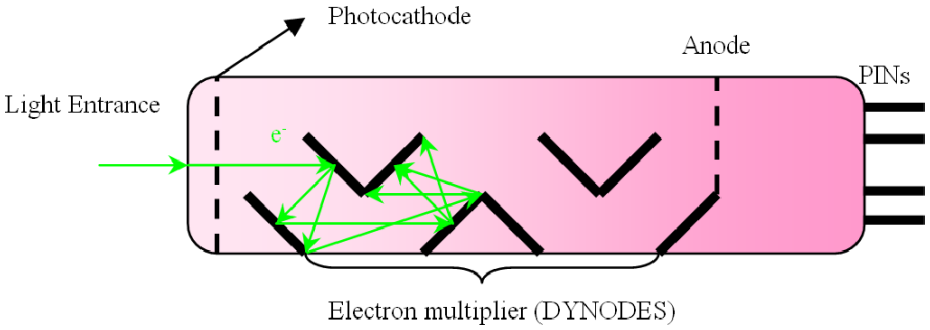

Figure 1.1: *Different parts of a photomultiplier tube. The light is generating an electron that will be multiplicated in different phases.*

When a primary electron released by an impinging photon strikes the surface of dynodes, a multiplied number of secondary electrons is emitted. A PMT with n dynodes generates δ^n electrons, if each dynode has a multiplication factor δ. The timing spread of impacting electrons on the anode, which controls the timing jitter of PMT, is determined by the dynode design. The last electrode in the PMT which converts the electron cloud to an electrical signal is the anode.

In order to prevent surface charge effects, charges should be removed from the anode quickly after their absorption. The dead time of PMTs is defined as the time

which takes to remove the charge from the anode surface. In most PMT designs the optimal detection efficiency of light is around 20 % in the visible range and falls below 1 % in near infra-red (IR) region of light. Changing photocathode materials boosts the detection efficiency, for instance GaAsP photocathodes increase the maximum quantum efficiency up to 45 % at 600 nm wavelength and GaAs can achieve 30% of detection efficiency at 900 nm wavelength [2].

Microchannel plates (MCPs) are secondary multipliers consisting of an array of millions of glass capillaries fused to form a thin disk, which replaces the dynodes and makes a MCP PMT. An electron, which enters the channel, hits the channel wall and produces the secondary electrons which are accelerated by the electric field. The accelerated secondary electrons hit the opposite wall and make additional secondary electrons. A single stage of MCP has a gain of 10^4 [3]. MCPs can achieve good spatial resolution and excellent timing jitter due to the spatial confinement of each photo generated electron cloud. The detection of the electron clouds can be through a phosphorous screen, which is imaged by a CCD or CMOS device [4].

The dark current in PMTs is the small amount of current that flows in the PMT even without the presence of any photon. The quantum efficiency and dark current increases by increasing the supply voltage of PMTs. Nowadays, PMTs are one of the technologies of choice for single-photon detection, due to their maturity and wide availability. However the disadvantages of this technology are their fragility, non-scalability, and weak spectral response in near IR [5].

1.2.2 PN photodiode

A photodiode is a reverse biased $p - n$ junction, whereas photons are absorbed everywhere with absorption coefficient α and an electron-hole pair is generated. An electron in the valence band moves to the conduction band by absorbing energy from a photon upon generation of an electron-hole pair. Wherever the electric field is present, electrons and holes are separated and transported in the opposite directions. Since the $p-n$ junction electric field is mostly in the depletion region, the photocarrier generation region is mostly confined to the depletion region. Figure 1.2 shows the cross-section of a $p - n$ junction and possible locations of electron-hole pair generation.

There are three possible locations where an electron-hole pair can be generated [6]:

1. In the depletion layer (region 1), where the electron and hole quickly drift in opposite directions because of high electric field in the region. Since recombination does not take place in the depletion region, each photocarrier adds e (considering generation rate of 1) to the current.

2. Far away from the depletion region (region 3). Here the pair recombines after some time without contributing to the signal.

3. Close to the depletion layer, where the electron and hole have a chance to diffuse to the depletion region by random diffusion (region 2). An electron coming from the p side or a hole coming from the n side can quickly be transported across the junction and therefore contribute the charge of e to the external circuit.

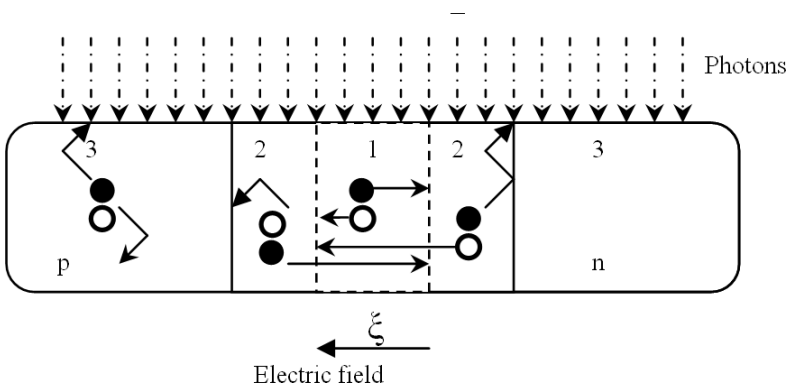

Figure 1.2: *Cross-section of a $p-n$ junction showing different regions where the photon can generate an electron-hole pair, and the direction each carrier takes.*

Photodiodes can be produced from different materials such as Si, Ge, and III-V semiconductor compounds. The direction of the incident light and the $p-n$ junction are often designed to be perpendicular to the $p-n$ surface, instead of parallel to it. The photodiode has an I-V relation as [6]:

$$i = i_s[\exp(\frac{eV}{k_BT}) - 1] - i_p \tag{1.1}$$

The term i_p, the photocurrent, is proportional to the photon flux. Photodiodes can have three different regions for operation: *open circuit* (photovoltaic), *short circuit*, and *reverse biased* (photoconductive mode).

1.2.3 PIN photodiodes

A $p-i-n$ diode is a $p-n$ junction with an intrinsic (lightly doped) layer sandwiched between p and n layers. The $p-i-n$ junction can also work on the three regions of bias operation mentioned for $p-n$ diodes.

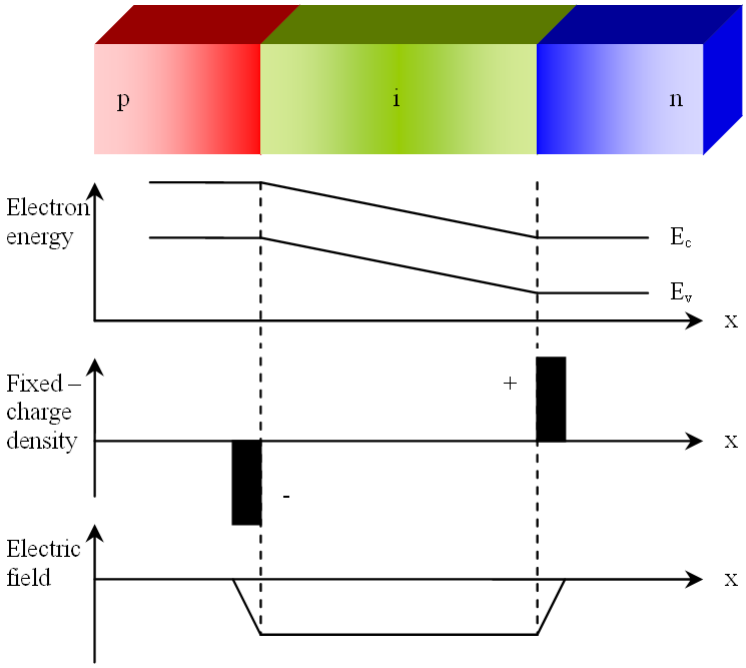

Figure 1.3: *The $p-i-n$ energy diagram, charge distribution and electric field distribution. The device can be illuminated parallel to the junction or perpendicular.*

Figure 1.3 shows the energy band diagram, charge distribution, and electric field distribution for a reverse biased $p-i-n$ diode. Using the intrinsic region has two advantages.

1. Increased depletion width: it can be translated in increasing the photon absorption depth which means more absorption in red and IR region of light.

2. Decreased junction capacitance, thereby reducing the RC delay constant.

The transit time of carriers drifting across the depletion region, and the *RC* time response are the main sources of the response time of photodiode detectors. The response time of $p-i-n$ photodiodes can be as low as a few hundreds of picoseconds [6]. It should be noted that in the case of silicon, the maximum responsivity occurs on wavelengths which are substantially shorter than the bandgap wavelength. This is due to the indirect-gap nature of silicon where the photon absorption transitions typically take place from valence to conduction band states which are above the conduction band edge.

1.2.4 Schottky-barrier photodiodes

Schottky barrier photodiodes are metal semiconductor photodiodes made by metal-semiconductor heterojunctions. A thin transparent metal plays the role of n or p layer in the $p - n$ junction. The Schottky-barrier band diagram is shown in figure 1.4.

Since n doped or p doped regions cannot be synthesized for every semiconductor material, the Schottky barrier devices are of interest for implementing photodetectors. The other reason for interest in Schottky photodiodes is that the depletion region in these devices is on the surface which decreases the surface recombination. Finally another reason for the popularity of these devices is the lower resistance of metal and hence less RC delay and larger bandwidth of operation, being a majority-carrier device [6].

1.2.5 Avalanche Photodiode (APD)

An avalanche photodiode is a strongly reverse-biased photodiode where the resulting electric field across the junction exceeds the critical value necessary to cause impact ionization. Figure 1.5 shows the principle of operation of APDs.

A photon which is absorbed at point 1, generates an electron-hole pair (electron in the conduction band and a hole in the valence band). The electron ionizes an atom in point 2 generating another electron-hole pair and so on. The energy needed for an electron to ionize an atom is supplied by relaxing the kinetic energy of the electron in a high electric field.

The ability of an electron or a hole to trigger ionization is quantified by ionization coefficient, α_e for electrons and α_h for holes. The ionization coefficients increase with electric field and decrease with temperature. The increase of the ionization coefficient by electric field is due to electron acceleration by electric field and its decrease by increasing temperature is due to an increasing frequency of collision that decreases the probability of a carrier gaining enough energy to cause ionization.

Two factors should be regarded simultaneously in an APD design. The first factor is maximizing photon detection probability which needs wider depletion regions. The second is to reduce the thickness of the multiplication region to minimize the possibility of localized uncontrolled avalanches (instabilities and microplasmas) being produced by electric field. A high uniform electric field should be designed in order to prevent the microplasmas.

The structure which meets both requirements is the separate absorption multiplication (SAM) structure. Nishida *et al.* [7] was among the first, employing separate absorption and multiplication regions in order to achieve low dark current. The photons will be absorbed mostly in the lightly doped large intrinsic region. The photoelectrons then drift across it under influence of a moderate electric field, and finally enter a thin multiplication layer with a strong electric field where the avalanche occurs.

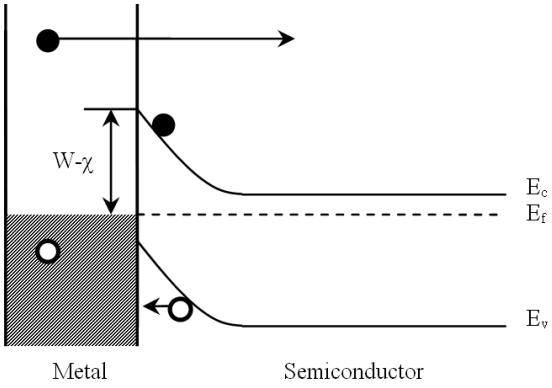

Figure 1.4: *Energy band diagram of a Schottky- barrier photodiode formed by depositing a metal on a n-type semiconductor. This photodetector is responsive to the photon energies larger than the Schottky barrier height* $h\nu > W - \chi$.

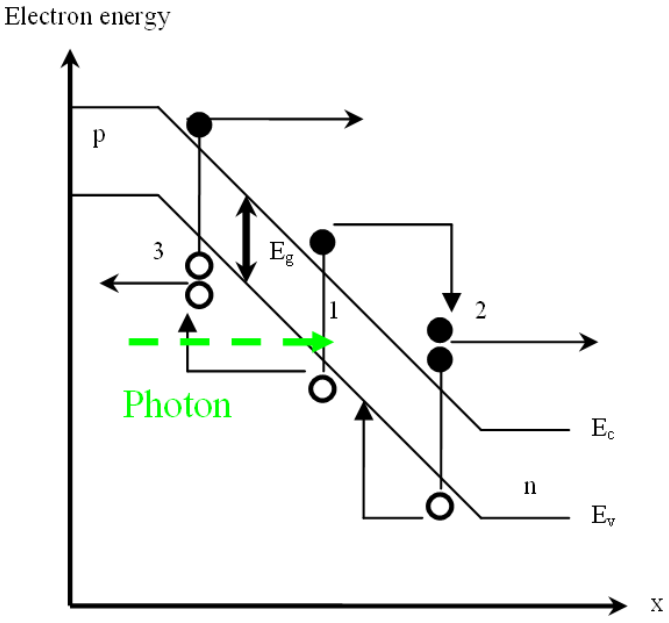

Figure 1.5: *The multiplication process in an APD.*

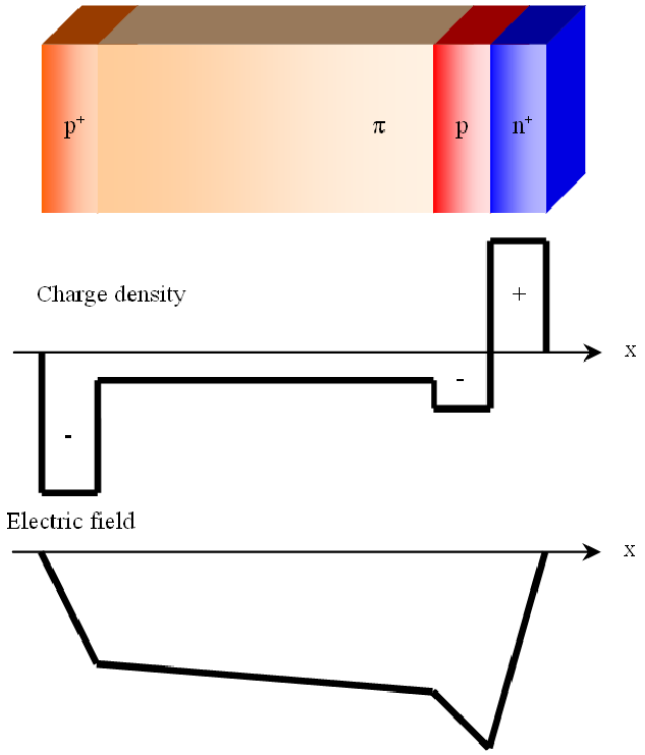

Figure 1.6: *Reach through $p^+ - \pi - p - n^+$ APD structure, the charge density distribution and electric field.*

Figure 1.6 shows this $p^+ - \pi - p - n^+$ APD structure. The photon absorption mostly takes place in wide π region, and then electrons drift to $p - n^+$ section where they experience a sufficiently strong electric field to cause avalanching. The reverse bias applied between p^+ and n^+ should be large enough for the depletion layer to reach through p and π regions, and to electrical contacts.

1.2.6 Quantum dot photon detectors

More recently, quantum dot based detectors have been introduced for use as single-photon detectors. This single-photon detector is based upon a transistor structure in which the conducting channel is close to a layer of quantum dots. When the separation of the quantum dots and channel is several nanometers, the resistance of the transistor is sensitive to a change in the occupancy of a single quantum dot by just a single electron. This combination allows the device to perform single-photon detection, since absorption of a photon creates carriers in

semiconductor, which after capture by a dot, produces a detectable change in resistance of channel of transistor [8].

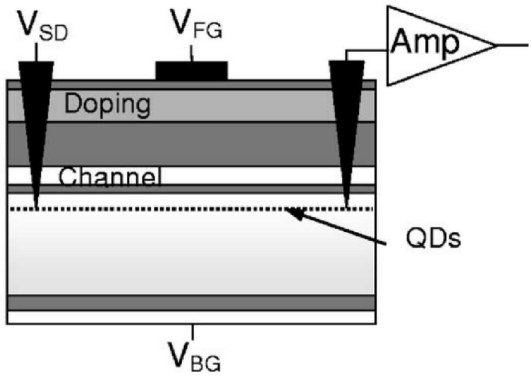

Figure 1.7: *Schematic of a quantum dot single-photon detector [9].*

A photon detection probability of 50-86 % has been reported for GaAs, and AlGaAs types of these quantum dot optically gated field effect transistors [10]. Figure 1.7 shows the cross-section of the detector. It is suggested that by avoiding avalanche process which can amplify both photo generated carriers and dark carriers, quantum dot photodetectors have lower noise. Since this type of single-photon counter is based upon a transistor, It is expected that the device can scale with technology and operate at low voltages (< 5 V).

1.2.7 Superconducting Single Photon Detectors (SSPD)

Superconducting single-photon detectors are nanostructure devices based on long stripes of an ultrathin superconducting film [11]. The nanostructure NbN superconducting single-photon detectors (SSPDs) are fast and reliable photon counters that have become a highly desired technology in recent years especially for working in infrared wavelengths [12].

Superconducting single-photon detectors operate based on the local suppression of superconductivity after absorption of a photon. Figure 1.8 shows how a photon can change the current characteristics of a superconducting nano layer. A superconductor is biased with a direct current near the critical current value above which the material would not be superconductor. When the photon is absorbed, a localized hot spot forms by absorption heating and the super current which diverts around the spot can be recognized electrically. The superconducting single photon detectors can also use NbN nanowires [13] as the medium of photon absorption.

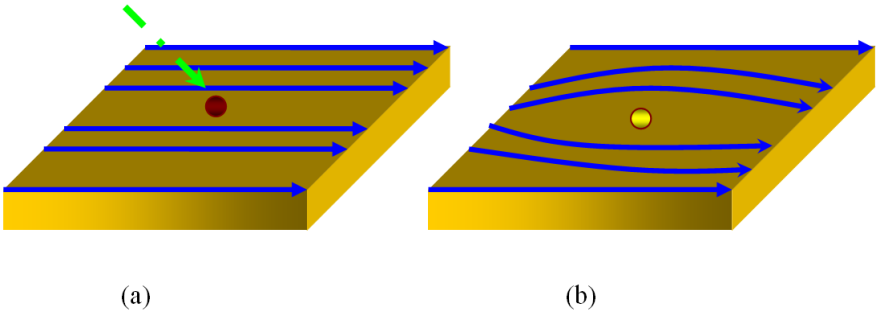

Figure 1.8: *Operation of a superconducting nanowire photon counter (a) A photon impacts a nanowire carrying a direct current near the critical superconducting current (b) A local hot spot forms, diverting current around it.*

The superconducting setup must be cooled till 2-4 K which is below the critical superconductivity temperature (T_c) of NbN. The advantage of these devices is the absence of leakage current.

1.2.8 Single-Photon Avalanche Diode (SPAD)

If an APD has the internal gain of M, for a single detected photon, the avalanche process produces M carriers [14]. Incidentally, the optimum value of M is the ratio $\frac{\alpha}{\beta}$, in which α is the ionization coefficient of electrons and β is the ionization coefficient of holes. Increasing the avalanche gain of M is achievable by widening the multiplication region. Multiplication region widening results in the response bandwidth decrease and increase in noise level. This mechanism is the reason why linear APDs have the maximum gain of $\frac{\alpha}{\beta}$ [5].

By further increasing the bias, an APD gain becomes virtually infinite. This biasing regime is known as Geiger mode of operation. In this case a single electron can start the avalanche process until an external circuit limits the avalanche current, thereby preventing the destruction of the device. This process is called quenching. Upon occurrence of an avalanche by a free carrier, a spike of current is generated. The combined time required to quench and recharge a SPAD is known as dead time.

The advances of SPADs was slow over the past 40 years until 2003, with the introduction of the first CMOS SPAD, when it accelerated to quickly reach nanometer feature size [14]. The main achievement of this miniaturization process was the reduction of the population of carriers induced in an avalanche. This led to several consequences. First, it reduced dead times from micro to nanoseconds, thereby increasing saturation intensities. Second, it decreased the probability of afterpulsing. Third, it has decreased optical and electrical crosstalk.

The simplest passive method for quenching is the use of a ballast resistance. Active methods are used when the dead time must be kept within precise limits and/or when the detector is large, thus causing a large parasitic capacitance. Active methods can include pull-up or pull-down transistors controlled by a feedback loop.

Figure 1.9 shows the gating and quenching mechanisms to control dead time. V_G brings the $p-n$ junction of the SPAD to the region below breakdown. When V_G returns to zero, V_R is used to bring back the $p-n$ junction above breakdown in a controlled time. When a SPAD detects a photon again, passive quenching is performed via MOS transistor M_Q that is controlled by bias voltage V_B.

After completing the quenching, the same transistor begins recharging to reach the biasing voltage in order to activate the device again. The dead time is thus forced upon the SPAD until sufficient relaxation time has elapsed, when the recharge is triggered again by the V_R signal. The bias voltage of V_b is used to realize high resistance in M_Q during quenching and low current for a long passive recharge to avoid interfering with the active recharge circuit [14].

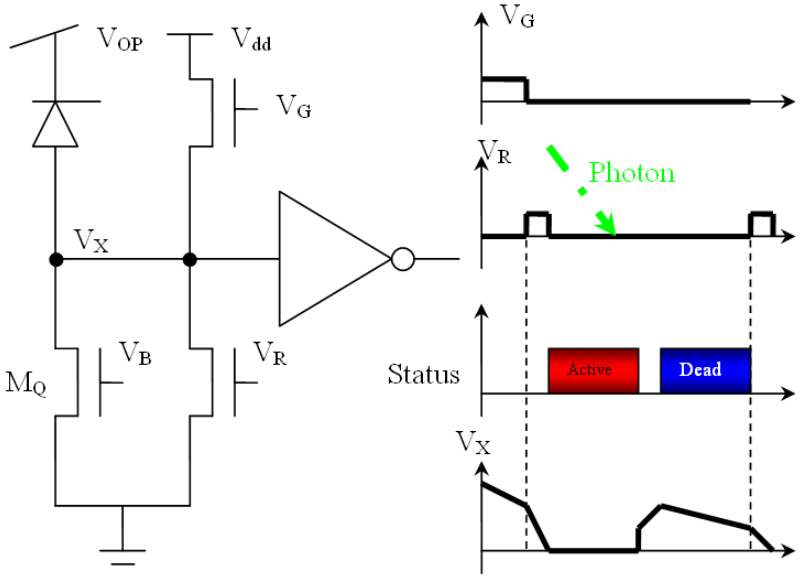

Figure 1.9: *Active circuit for gating and quenching of SPAD.*

1.3 Thesis motivation and contributions

This thesis addresses two main issues in the world of single-photon detection namely:

1. quantum parasitic effects

2. miniaturization

Several quantum parasitic effects will be quantified in novel ways. n-tub guard rings identified in SPADs for the first time are compared with conventional p-tub guard rings on the same technology, to quantify guard ring topology effect on the tunneling.

Random Telegraph Signal (RTS) behavior is reported and characterized in the dark count rate. RTS is observed in a SPAD fabricated in 0.8 μm CMOS technology and in four proton-irradiated SPADs designed and fabricated in 0.35 μm CMOS technology. To the best of our knowledge, this is the first time RTS is being reported in SPADs in CMOS technology. RTS characteristics are evaluated experimentally and verified theoretically with respect to bias and temperature.

Miniaturization is explored with the first single-photon avalanche diode designed and successfully tested in technologies smaller than 130nm. The proposed structures, implemented in 90nm standard CMOS technology, emerged from a systematic study aimed at miniaturization, while optimizing overall performance. The guard ring design is the result of an extensive modeling effort aimed at constraining the multiplication region within a well-defined area where the electric field exceeds the critical value for impact ionization.

Chapter 2

Single-Photon Avalanche Diode Fundamentals

This chapter reviews the SPAD types following by figures of merit of individual and arrays of SPAD. Last section introduces the main SPAD applications.

2.1 SPAD types

2.1.1 Reach-through devices

SPADs can be categorized in at least two different forms of design:

Reach-through devices: in these devices the substrate thickness is used as absorption region. The thick substrate, which can be divided to two parts of absorption and multiplication, can detect light in large wavelength ranges.

Planar devices: in these devices only a few percent of the substrate thickness is used as an absorption region. Most of the substrate is not depleted. The devices are more sensitive to blue and UV light since the absorption is mostly taking place close to the surface.

McIntyre and Webb introduced one of the first reach-through SPADs on silicon with a custom technology with ultra-low doped p-substrate silicon [15],[16]. Figure 2.1 shows the cross-section of the reach through SPAD.

The depletion layer of the device was 20 to 100 μm thick and the breakdown voltage was 100 to 500 volts. The SPAD active area was 50-500 μm and early edge breakdown effects were prevented by a p^+ enrichment and by reducing the silicon thickness over it by etching the wafer.

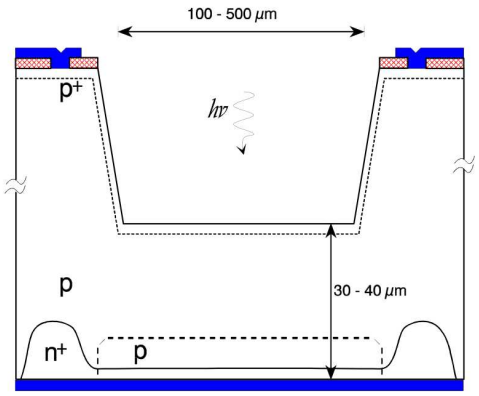

Figure 2.1: *Reach-through Single-photon detector. The device contains different absorption and multiplication regions [15].*

The first implementation of the device, was improved in terms of photon detection probability reaching 30 % at 500 nm and 70 % at 700 nm [16].

The dark count rate of these devices reached 150 Hz at room temperature due to an ultra-clean process employed in the fabrication phase. Since large diameter SPADs were used, the jitter introduced by the devices was in the order of 300 ps. The high bias voltage (300-400 V) dictates a special effort to design peripheral circuits. The devices fabricated in this technology were expensive and yield and uniformity were poor. Other approaches were proposed for making reach-through SPADs in the Lincoln Laboratories for the development of LIDAR systems [17],[18].

The device presented in [17], unlike previous reach-through devices, does not need special etching at the edges to confine the electric field. Rather, deep implantation structures confine the electric field and direct the generated charges to the avalanche area [19].

2.1.2 Early planar SPADs

Haitz was one of the first to implement SPADs in planar technologies [20]. The process steps for the realization of these devices are similar to Metal Oxide Semiconductor Field Effect Transistor (MOSFET) standard technologies.

The narrower absorption region reduces the photon detection probability, and a shift to the blue region of the spectrum. Reducing the junction diameter also reduces jitter and the dead time due to lower junction capacitance. Moreover, it reduces the breakdown voltage, which makes it easy to integrate the device in CMOS circuits. In these structures the multiplication region is under the n^+ layer and two lightly doped n^- guard rings are employed to overcome the problem of

premature edge breakdown.

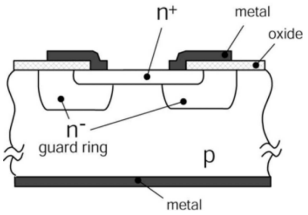

Figure 2.2: *Haitz 's APD using diffused guard ring as the first planar devices proposed as SPAD [20].*

The guard ring structure which Haitz employed in SPADs is used in most popular SPAD designs nowadays (figure 2.2). The junction breakdown occurs at 32 V in these devices.

2.1.3 Double epitaxial SPADs

Cova introduced a double-epitaxy SPAD [21] with an improved jitter response in 1989 (figure 2.3) [21],[22].

The photons which enter the device from the top surface are absorbed in p-type epitaxial layer. The electron-hole pairs are attracted by a low electric field to the n^+/p junction. The novelty of this SPAD is the use of a double-epitaxial structure, forming two diode junctions. The buried diode prevented photogenerated electrons to enter the multiplication region causing increased timing uncertainty.

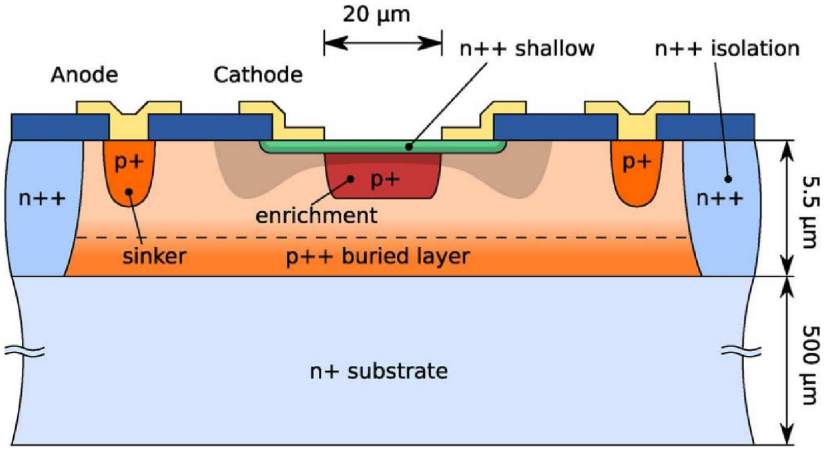

Figure 2.3: *Cross-section of double-epitaxy SPAD [23].*

The top epitaxial structure was used to form an ultraclean multiplication region for SPAD. The p^+ layer implanted below the top epitaxial layer was used to overcome large resistivity of this layer. Finally, a p^+ implant at the center of device was employed as guard ring to reduce edge breakdown effects. This device had a breakdown voltage of 50 V, a jitter performance of 55 ps Full Width Half Maximum (FWHM), and an improved dark count rate, compared with the previous devices mentioned in this section, which were in the order of 10 k counts per second or 10 kHz. Further improvements were reported for this device, to lower the jitter of device for TCSPC (time- correlated single photon counting) applications [24],[25].

2.1.4 SPAD in standard CMOS technology

Rochas used a commercial high-voltage standard CMOS process by Austrian Micro Systems (AMS) to implement a SPAD based on Haitz structure [26],[27]. Figure 2.4 shows the cross-section of SPAD presented by Rochas using standard technology.

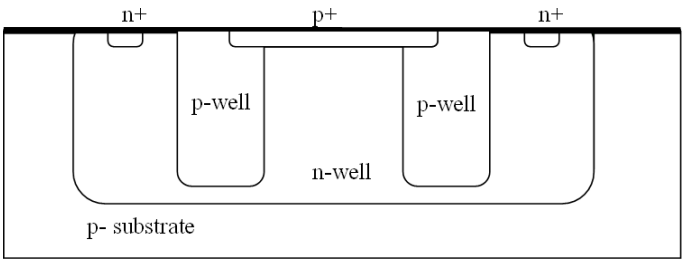

Figure 2.4: *Cross-section of a standard CMOS SPAD proposed by Rochas.*

Deep n-wells were employed for the SPAD separation from the substrate. These structures are analyzed in this thesis where the introduction of deep n-wells is studied in detail. The early device characterizations show that the photon detection probability peaks at 30 % at 450 nm wavelength. The shift in absorption wavelength, in comaprison with previous devices is due to shallow p^+ junction close to device surface. A dark count rate of 100 Hz for a 7 μm diameter SPAD was achieved. The dead time was 30 to 50 ns, the breakdown voltage was 25 V, and a FWHM jitter of 60 ps was reported.

Rochas CMOS SPADs show excellent performance, while the low fill factor is the main drawback.

2.1.5 SPAD implementation with using Shallow Trench Isolation (STI) layers

Shallow Trench Isolation (STI) was proposed by several authors as an alternate to guard rings. This layer is mainly used to prevent punch-through and latch-up in CMOS circuits. The trench layers are filled with silicon dioxide, isolating planar devices. STI is automatically created outside of the active areas containing transistors, diodes and contacts to wells and subtrate. Figure 2.5 shows the cross-section of a SPAD with STI.

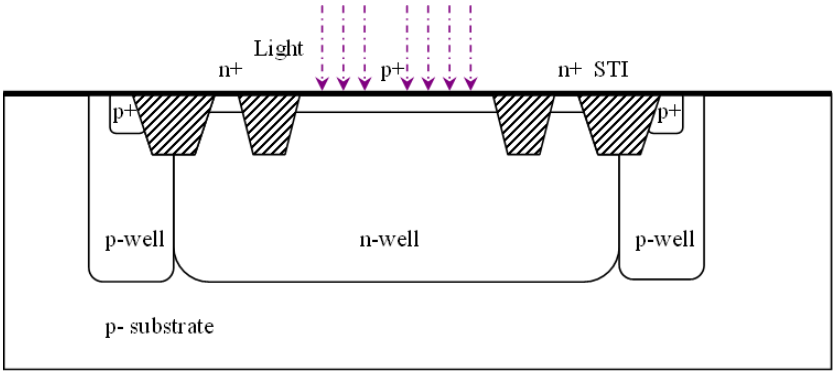

Figure 2.5: *Cross-section of SPAD using STI layer as guard ring.*

Because the breakdown field of silicon dioxide ($10^7 V/cm$) is about 30 times higher than the breakdown field of silicon ($3 \times 10^5 V/cm$)[28], STI guard rings can be made 30 times narrower than silicon guard rings.

2.1.6 Guard ring structures for Premature Edge Breakdown (PEB) prevention

SPADs must tolerate high avalanche currents with minimum impact on surrounding analog and digital circuits lying on the same substrate. Therefore, special care is needed to design structures for prevention of PEB. Guard rings have been used for decades to achieve several goals, including dark current reduction in photodiodes, $1/f$ noise reduction in oscillators, etc. The use of guard rings in SPADs for PEB prevention is known since Haitz. The main topologies introduced to create SPAD guard rings is reviewed in this section.

Figure 2.6 shows one of the first junction planarization methods used in high-field devices. The beveled junctions were cleaved so that only the planar junction surface remains [29],[30]. This technique cannot be used as a standard process for large number of pixels and cannot be included in CMOS standard processes because cleaving the device is incompatible with planar processes.

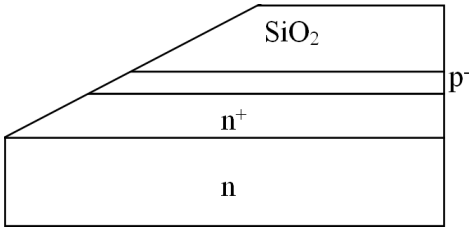

Figure 2.6: *Beveled junction. The depletion region is between n^+, and p^- [30].*

Figure 2.7 shows an extension of the previous guard ring. In the mesa guard ring, etching and subsequent filling with a dielectric physically planarizes and isolates adjacent junctions [31]. Arrays of devices using this technique are demonstrated, where fill factor and pitch are adversely impacted [5].

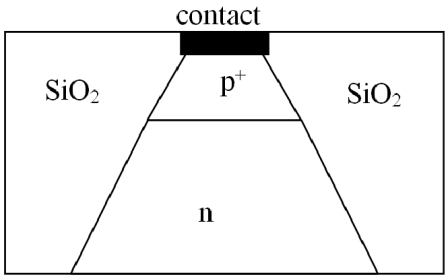

Figure 2.7: *Mesa isolation as the guard ring of SPAD [5].*

Figure 2.8 shows field-limiting rings which were used in high-voltage devices to prevent PEB in curved junction regions [32]. PEB prevention is accomplished by extending the depletion region to edges where most hot spots for electric fields are expected to be located.

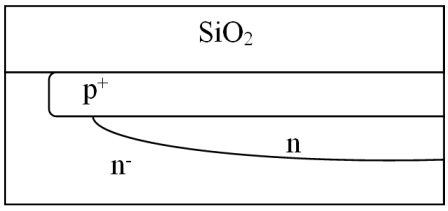

Figure 2.8: *Straddled junction as the guard ring for SPADs [5].*

Since the 1960s avalanche diodes have been using diffused-ring structures shown in figure 2.9 [33]. Edge breakdown is prevented using a low-doped guard ring at the edges of the junction where high electric field is expected. This method extends the region across which the electric field develops, thereby decreasing it at the edges. While diffused guard rings are compatible with standard processing steps, they occupy large space and lower the fill factor [5].

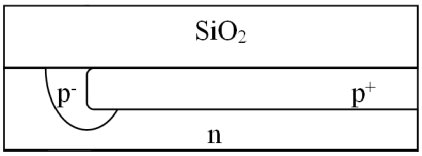

Figure 2.9: *Diffused guard ring as guard ring of SPAD [34].*

2.2 Figures of Merit (FOM) for individual detectors

2.2.1 Dark Count Rate (DCR)

The avalanche process in SPADs can start with any free carrier in the multiplication region. The spontaneous avalanche breakdowns started without light are called dark counts. The rate of avalanches which occurs spontaneously, is called dark count rate (DCR) and it is measured in Hz or cps. The different mechanisms that can cause dark counts are described in this section.

Shockley-Read-Hall generation: A free electron thermally excited from the valence band to the conduction band can cause an avalanche. However, due to the large (1.12 eV) and indirect bandgap of silicon, such transitions are extremely rare at normal working temperatures [35]. The traps and defects present in the depletion region introduce additional energy levels, acting as intermediate states between the valence and conduction band and thus significantly increase the rate of free carrier generation. This trap-assisted thermal carrier generation is known as Shockley-Read-Hall (SRH) process and it is studied in detail in this thesis [36]. When SRH generation is the dominant noise source of a SPAD, DCR doubles every 10 ^oC [35].

Band-to-band and trap-assisted Tunneling: Tunneling is defined as the penetration of electrons in confined potential barriers can occur between the valence and conduction band of a reverse biased $p-n$ junction. Tunneling probability is highly dependent on electric field intensity at the $p-n$ junction interface, and the thickness of the junction. Tunneling probability can also greatly increase by the presence of trap energy levels within the bandgap of the material. Tunneling generation rate is less dependent on the temperature in comparison with thermal SRH generation.

Multiple Excitonic generation: Multiple excitonic generation (MEG) is demonstrated in synthesized nanocrystals (quantum dots) including PbS, PbSe, PbTe, and Si. MEG is a process whereby multiple electron-hole pairs are generated upon absorption of a single photon in semiconductor nanocrystals [37]. MEG can be used to increase the solar conversion efficiencies in single-junction photovoltaic cells. The quantum origins of MEG is still being debated [38],[39]. Some studies have proposed impact ionization as the origin of multiple exciton (bound electron-hole pairs) generation. In excitonic generation process, electrons and holes in conduction and valence band are generated with different wavelengths.

Figure 2.10 shows different generation mechanisms affecting the behavior of a $p-n$ junction.

2.2.2 Photon Detection Probability (PDP)

The quantum efficiency of a photodetector(η) is defined as the probability that a single photon generates a photocarrier pair contributing to detector current. The quantum efficiency can be written as:

$$\eta = (1-R)\zeta[1-\exp(-\alpha d)] \tag{2.1}$$

Where R is the optical power reflectance of the surface, ζ is the fraction of electron-hole pairs contributing successfully to detector current, α is the absorption coefficient, and d is the photodetector absorption thickness.

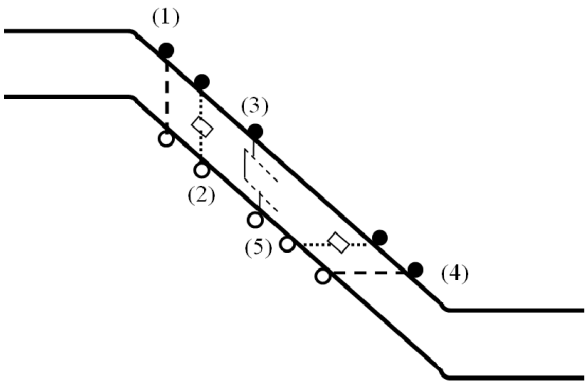

Figure 2.10: *Generation mechanisms in a SPAD (1) Thermal direct generation (2) SRH trap-assisted generation (3) Excitonic generation (4) Band-to-band tunneling (5) Trap-assisted tunneling.*

Photon detection probability (PDP) is defined as the probability of sensing a photon in a given wavelength. PDP and quantum efficiency are both functions of absorption coefficient and hence are functions of wavelength. PDP is calculated by decreasing the number of dark counts from total counts and dividing the result by the number of illuminated photons:

$$PDP(\lambda) = \frac{A - DCR}{P} \tag{2.2}$$

In this equation A is the avalanche breakdown counts being sensed by the SPAD in Hz, and P is the number of impinging photons in Hz at the wavelength λ. A photon goes through a number of phases before triggering a detectable avalanche:

1. Passing from the passivation layers located on top of the silicon surface is the first obstacle for photons. Since the technology trend is to increase the number of metal interconnects and passivation layers, the optical stack on top of a SPAD is becoming more complex. This complex stack with different refractive indices lowers the PDP. Special architectures can be made to increase penetration of photons in silicon and minimize the reflection from surface. Different anti-reflection coating materials can be used to increase light absorption in SPADs [40],[41]. Figure 2.11 shows the optical stack and 9 copper metal interconnect layers in 90nm CMOS technology. Surface reflections themselves can account to 40 % loss in detection probability [34]. Optical stacks can also create large swings in the spectral response with up to 30 % change in response within a 50 nm of spectral range [35].

   The quality of Chemical- Mechanical- Polishing (CMP) in device processing also determines thickness and planarity of silicon-dioxide layers located on top of substrate [42]. The variation of 4 micron is reported for a deep- submicron technology with 6-metal layers corresponding to a large variation in transmittance whithin different dies and different wafers [43]. This nonuniformity can contribute the PDP variation in different devices on the same run in deep-submicron technologies.

2. Similarly to APDs, the minority free carriers generated outside the depletion region can cause an avalanche if carriers reach the SPAD multiplication region. Recombination of photogenerated carriers outside the depletion region is the main obstacle, preventing them to trigger an avalanche [35].

   Therefore, for increasing the probability of avalanche initiation by photons, the multiplication region of SPAD should be located close to silicon surface and extend significantly into the substrate. The depletion region of SPAD should also be as wide as the absorption length for the targetted photon wavelengths.

3. The avalanche sensing should also be implemented appropriately. The readout circuit and quenching circuits should limit the current passing

through device in order to reduce the probability of device failure after avalanche.

2.2.3 Spectral response

Different applications require different spectral light response. While most of silicon absorption in reach-through devices occurs in the red region the spectrum due to the bandgap of silicon, absorption in deep-submicron SPADs mostly occurs in blue part.

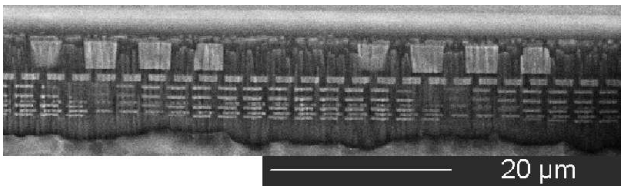

Figure 2.11: *Scanning Electron Microscopy of optical stacks in 90nm CMOS technology. The optical stack consists of 9 metal layers and 7 passivation layers decreasing the photon absorption.*

The depletion layer width of a double sided abrupt junction with the assumption of full depletion, and at thermal equilibrium can be calculated using the following equation:

$$W_{dep} = \sqrt{\frac{2\epsilon_s \phi_{bi}}{q}(\frac{1}{N_a} + \frac{1}{N_d})} \tag{2.3}$$

In this equation ϵ_s is silicon permittivity, ϕ_{bi} is the built-in potential, q is the charge of the electron, N_a is the acceptor concentration in p layer and N_d is the donor concentration in n region. In order to calculate the depletion depth as a function of reverse biasing, the following equation can be used:

$$W_{dep} = \sqrt{\frac{2\epsilon_s(\phi_{bi} + |V_r|)}{qN}} \tag{2.4}$$

in which V_r is the reverse bias voltage, and N is calculated using following equation:

$$\frac{1}{N} = \frac{1}{N_a} + \frac{1}{N_d}. \tag{2.5}$$

Higher bias voltages should be applied in order to increase photon detection probability according to the previous equations. The depletion region should cover silicon absorption length completely to achieve the whole spectral light sensitivity.

2.2.4 Timing jitter

The statistical timing variation of processes from photon absorption until avalanche sensing, makes up the overall timing jitter. Considering different electric field values in the depletion region, the timing jitter depends on the position where the avalanche is triggered. Approximately, the timing uncertainty of avalanche current is in the range of free carrier transit time across the junction at saturation velocity [44], which is 10 ps per micrometer of depletion depth [45]. The lateral position of the carrier which initiates the avalanche breakdown can also influence the response time [46],[47].

Due to lateral and vertical dependence of timing uncertainty on photon absorption location, SPADs with narrower and smaller depletion regions show better jitter responses [48],[49]. Since avalanche build-up has a nearly Gaussian distribution, which is the dominating source of jitter in SPAD systems, the timing jitter is usually given as full-width-half-maximum (FWHM) of the photon arrival time distribution [35].

Figure 2.12 shows the histogram of the response of a SPAD implemented in 90nm technology.

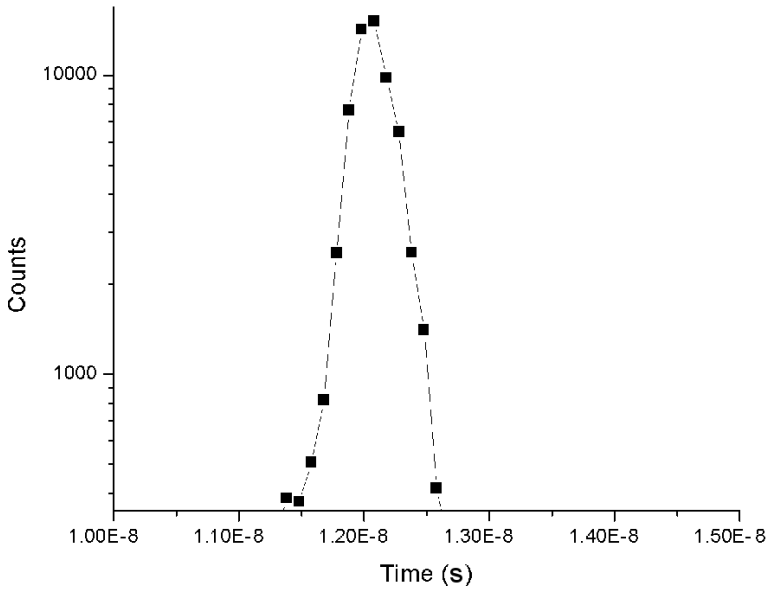

Figure 2.12: *Timing histogram of a SPAD, representing 368 ps FWHM jitter.*

The timing jitter is measured by creating a histogram of time difference

between an optical source (laser) and the electrical pulse representing photon sensing by the SPAD. In some cases, a slow tail of lower amplitude after the gaussian shape of jitter appears, which is due to the diffusion of other free carriers absorbed beneath the depletion region. These carriers generate electron-hole pairs and the minority carriers may diffuse back into avalanche regions and thus trigger an avalanche breakdown.

Due to the low speed of diffusion, low amplitude timing jitter induced by diffusion can be significantly larger than the main jitter from the avalanche build-up process. In addition, since the mean penetration depth of photons in silicon is strongly wavelength dependent, the characteristics of the diffusion tail vary significantly with changing wavelength.

The uncertainty of sense amplifier's threshold can also contribute to jitter. In reality, avalanche amplitudes may vary, either due to incomplete recharging, or due to secondary effects, such as temperature variations. The timing jitter is important in time correlated single-photon counting (TCSPC) experiments as well as in applications where time gating is used to reduce noise effect.

The following steps can be followed in order to reduce the SPAD jitter:

1. Reducing the volume of depletion region to minimize uncertainty in primary carrier position.

2. Designing SPAD with highly localized electric field. The retrograde doping devices can also prevent charges generated far from the depletion region to diffuse to the depletion region [50]. Figure 2.13 shows the cross-section of a SPAD with retrograde doping.

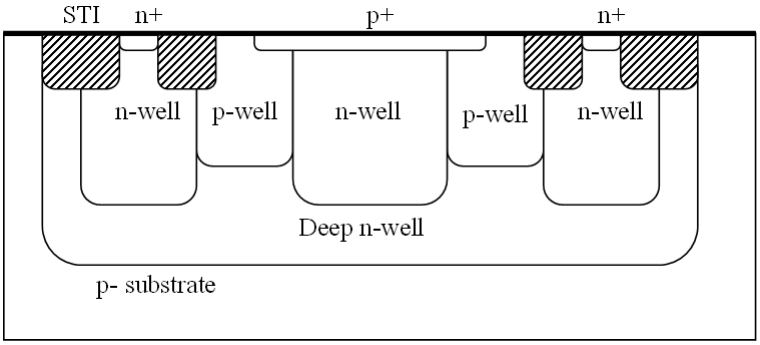

Figure 2.13: *Retrograde doping introduction for jitter reduction in SPADs.*

3. Using active recharge techniques which have identical avalanche amplitudes. Some special circuits had been proposed to high-pass filter avalanche pulses improving precision [51].

4. Performing time-to-digital conversion (TDC) on chip to prevent jitter from inter-chip interconnects, such as wire bonds [52].

2.2.5 Afterpulsing

In the avalanche process, each electron or hole accumulates enough energy to bring another electron from valence to conduction band by relaxing its energy in the collision. This flow of high-energy electrons can fill traps or change the configuration of a defect. The modified trap or defect itself causes secondary avalanches, thus triggering a process known as afterpulsing [53]. The probability of afterpulsing can be reduced, by reducing the number of traps, or by reducing the carrier flow; this reduces the probability of trap or defect deformation. Limiting the current flow can be carried out by reducing the parasitic capacitance of the junction or by active quenching [35].

Reducing the trap density is achievable by using advanced and cleaner device processes. Another way of restricting afterpulsing probability is to increase dead time, thus allowing trap relaxation.

Table 2.1 summarizes the noise sources in SPADs and their correlation to thermal and bias behaviors. The afterpulsing can be distinguished from other noise sources of SPADs by evaluating the noise autocorrelation function.

2.2.6 Dead time

SPADs are biased above breakdown to operate in Geiger mode. The voltage above breakdown is known as excess bias voltage (V_e). Upon avalanche build-up and propagation, a circuit senses the current and lowers voltage across the SPAD to prevent destruction of the device. This function is followed by a recharge phase that brings the SPAD to the same bias condition prior to the avalanche. The time, when SPAD is under lower biasing conditions and therefore unable to detect photons is defined as dead time. Dead time includes both quenching time and recharge time.

2.3 Figures of merit for SPAD arrays

2.3.1 Fill factor

Fill factor (FF) is defined as the ratio between light sensitive area and total area of a pixel. While SPADs using diffusion guard rings have low fill factors, STI-based SPADs have the potential for FF improvement [45],[41].

Passive quenching results in higher FFs, compared with complex active quenching. It is demonstrated that using microlenses integrated on top of SPAD pixels may increase fill factor by focusing light to absorption region [54]. However chief ray angle constraints arise on the optics to be used with these devices.

2.3.2 Crosstalk

Crosstalk is defined as the probability that photon detection in a pixel is caused by photon detection or dark counts in any surrounding pixels. Two different mechanisms are regarded as crosstalk sources.

Electrical crosstalk: during the avalanche process, a large number of carriers are flowing through the multiplication region. Each free carrier generated inside one device can move to the other depletion region and cause avalanche breakdown. In this case, electrical crosstalk may occur. Physical isolations can be used to prevent electrical crosstalks. Moreover, electrical isolations can be built using a potential barrier. In triple-well SPADs, each SPAD is located in a separate deep n-well surrounded by p-type substrate forming the potential barrier. Crosstalk may also be generated by supply line interference (resistance, capacitance, or inductive) [55].

Optical crosstalk occurs when a SPAD emits secondary photons due to electroluminescence during avalanches [56] that can be detected by other SPADs. Since electrolumiescene is related to current flowing through a SPAD, limiting the current reduces the probability of optical crosstalk. In CMOS SPADs the effect of direct crosstalk was analyzed in [57],[58].

Figure 2.14 shows an optical crosstalk measurement setup for two SPADs, designed far to prevent electrical crosstalk [59]. The figure shows possible paths a photon generated by an avalanche can take to trigger avalanches in other SPADs. Different studies are carried out to analyse optical crosstalk. Trenches coated with metals are proposed between detectors to reduce crosstalk [60],[61]. In [62], it was shown that most of crosstalk is originated from light reflection in the substrate back-side.

	SRH	Tunneling	Afterpulsing
Type	Uncorrelated	Uncorrelated	Correlated
Distribution	Poisson	Poisson	Exponential
Excess bias voltage	Strong dependence	Very strong dependence	Strong dependence
Temperaure	Strong dependence	Weak dependence	Inverse dependence

Table 2.1: *Summary of SPAD noise sources.*

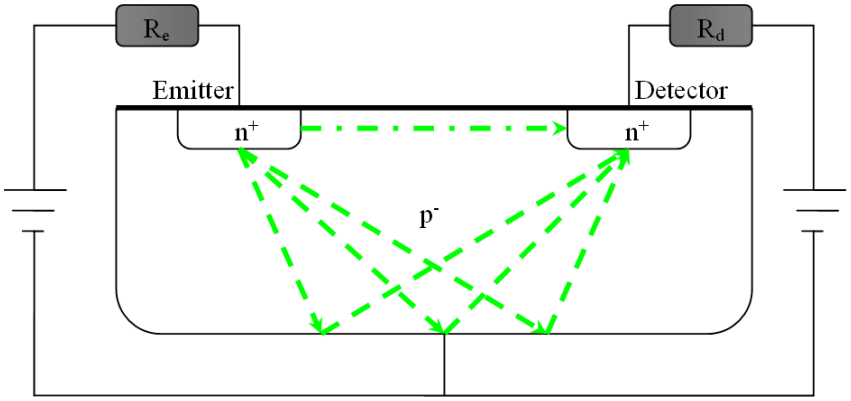

Figure 2.14: *Test setup for optical crosstalk evaluation. Light generated in an avalanche, can travel with or without reflections through the substrate [59].*

2.3.3 SPAD scaling

Scaling of electronic devices follows Moore's law since the 1960s. Moore predicted that transistor density would double every 18 months. Different reasons are suggested for device scaling namely increasing yield, increasing speed, and raising frequency of operation. Scaling involves a reduction of voltages, currents, vertical and lateral dimensions. Device scalings had forced doping implantation densities to rise. The design of avalanche photodiodes in deep-submicron CMOS technologies involves additional challenges. In order for a photodetector to operate in Geiger mode planar multiplication region with uniform electric field over the entire active region is required.

Higher noise levels in deep-submicron technologies are due to (1) *higher doping levels*, (2)*reduced annealing steps*, and (3) *STI*.

Higher doping levels increase DCR due to increased band-to-band tunneling. The increase in parasitic capacitance has an influence on afterpulsing and dead time. Furthermore, the strength and duration of annealing and drive-in diffusion steps are reducing in state-of-the-art processes. Moreover shallow dopings and small feature sizes of deep-submicron processes do not allow the use of annealing, as the lateral diffusion of implantation should be avoided. However, efficient annealing should be applied in order to reduce concentration of impurities and to reduce traps and defects present in the lattice, which increases trap assisted generation and tunneling.

Lower quantum efficiencies and PDP in deep-submicron technologies are due to (a) lower depletion widths and (b) increased optical stack thickness.

Breakdown and excess bias voltages are decreasing in deep-submicron technologies which can be translated in a even lower depletion layer depths.

2.3.4 History of SPAD scaling in CMOS technology

The first CMOS SPAD was introduced by Rochas in 0.8 μm technology [34]. The second generation of SPADs, demonstrated in 0.35μm high-voltage CMOS technology, originated in our research group [63]. The scaling trend continued in 180nm technology by Faramazpour and Niclass [64],[65]. It was also demonstrated that SPADs in 130nm imaging technology can have reasonable characteristics by Gersbach and Richardson [66],[50].

The following table compares the main characteristics of SPADs in different CMOS technologies.

Figure of merit	0.8 micron [58]	0.35 micron [63]	180nm [64]	130nm [67]	Unit
Timing jitter	82	80	-	200	ps
DCR (@ room temperature)	350 ($V_e = 5$ V)	750 ($V_e = 3.3$ V)	70000	12 ($V_e = 0.6$ V)	Hz
Active area	38	78	78.5	50	μm^2
V_{bd}	25.5	17.4	10.2	14.4	V
Dead time	< 40	40	30	100	ns
PDP @ 460 nm	26($V_e = 5$ V)	40($V_e = 3.3$ V)	5.5($V_e = 2$ V)	17.5($V_e = 0.6$ V)	%
EM spectrum	380-900	350-1000	360-700	350-1000	nm
Technology	CMOS	CMOS	CMOS	CMOS	-

Table 2.2: *Summary of SPAD reference measures.*

2.3.5 Challenges

STI is present in most CMOS processes with feature sizes below 0.25 μm, in regions not covered by the source or drain implantations. In some processes, there is an option to prevent STI formation in certain locations. Trench isolations were introduced to reduce the bird's beak effects and to increase the isolation capabilities [68]. The fabrication process of STI involves reactive ion etching (RIE) which causes damages to the sidewalls.

RIE includes surface residues, contamination of heavy metals, dislocations, stacking faults and mechanical stress [69]. Traps, defects, and dislocations are formed by STI and can increase the SPAD noise. It is proposed that cobalt impurities introduced by the cobalt silicide process are causing noisy pixels in CMOS active pixel imagers [70]. Therefore, and due to very high DCR levels of STI-bound SPADs [71], deep-submicron SPADs must be properly isolated from STI for reasonable noise performance. Chapter 4 describes design and implementation considerations we adopted in the design of 90nm and 65nm CMOS SPADs for reducing noise effects.

2.4 Single-photon avalanche diode applications

There has been a dramatic increase in interest toward single-photon detectors applications in the past decade [72]. Single-photon detectors now support a wide variety of applications at the frontiers of science and engineering from time correlated single-photon counting (TCSPC) to quantum information processing (QIP) [73],[74].

Single-photon detectors have found uses in diverse areas, including single-molecule dynamics [75], quantum communication systems [76],[77], medical imaging [78], and biometrics [58]. Providing information regarding individual photon-arrival events is a unique property of SPADs, consequently, several types of data can be derived from such detectors including:

1. A digital output corresponding to photon arrival, within the exposure window.

2. Time density of photon arrival events, corresponding to impinging photon flux.

3. Photon times-of-arrival statistics.

2.4.1 Ladar and long-range 3D imaging

Optical range finding offers significant advantages over radio-frequency (RF) and ultrasound (US) range finding schemes. Due to the shorter optical wavelengths, laser radars (ladars) can resolve fine features of objects, rather than just detecting them [79], [80]. A ladar consists of a pulsed or modulated light source, illuminating an object with a narrow beam. A scanning system focuses returning photons from different directions of object onto a SPAD pixel in rapid succession, measuring the photon time of flight [5]. Several closely related techniques allow for the collection of photons arriving within a limited time-gate, to form an image from objects [81]. A SPAD should work in the daylight with large background noise to be suitable to enter these applications. Moreover real-time, time-of-arrival processing is necessary with high accuracy measurements [82].

2.4.2 High-resolution 3D imaging

Three-dimensional non-stereoscopic imaging has been demonstrated using active-illumination and an array of SPADs [58]. For a 3D imaging system entitled for sub-millimeter range precision, to attain a 0.8 mm depth resolution, the time uncertainty of 5.3 ps is necessary. This estimate includes the jitter of the pulsed light source that of the SPAD, and the uncertainty introduced by the time-to-digital converter. Due to high processing load of data, necessary for high accuracy measurements, such applications are amenable to deep-submicron commercial technologies, such that processing can be performed on chip.

2.4.3 Single-photon detection for molecular imaging and spectroscopy

Behavior of individual molecules surrounded by a multitude of molecules, can be detected by single-molecule spectroscopy [83]. To probe the molecule, a light beam, typically from a laser source, is used to trigger an electronic transition specific to the target molecule. The resulting optical transition is detected by a time-resolved image sensor. The indirect detection of this absorption, via fluorescence can be used to determine certain properties of the fluorescent molecule and of its surroundings. SPADs have shown to be successful in applications involving fluorescence life time imaging microscopy (FLIM), fluorescence correlation spectroscopy (FCS), Forster Resonance Energy Transfer (FRET), etc [84].

2.4.4 Time-domain diffuse optical tomography

Time-domain diffuse optical tomography (DOT) uses light to determine the three-dimensional structure of highly-scattering living samples. Some applications of DOT include optical mammography [85] and brain imaging [78]. Since light is strongly scattered in living tissue, there are almost no ballistic (non-scattered) photons in tissues thicker than 1 cm [86]. Instead, photons diffuse through the tissue, thereby deteriorating their spatial resolution. In order to analyze the composition of the tissue sample, effects of absorption and scattering can be discriminated. Whereas absorption lowers the intensity of the optical signal, scattering smears it in time domain. That includes the main principle for analyzing DOT data, whereas using DOT results in a 3D model of the specimen [87].

2.4.5 Automotive applications

Automotive is one of the most demanding application domains in terms of reliability, robustness, cost, and environmental constraints. Operating constraints such as temperature range and background light levels impose significant challenges to TOF rangefinders. Operation range of sensors can be from -40^oC to 125^oC and the background-light rejection should reach 100klux sun light [45].

It is expected that environmental sensing in cars will considerably increase road safety [88],[89]. A 3D camera installed on a car will protect vulnerable road users, such as pedestrians, bicyclists, motorcyclists, etc. Moreover, front detection can be enhanced with a number of additional measurement views. While long-range sensors can assist drivers in lane change maneuvers, short-range front and rear detectors can be used for autonomous and semi-autonomous parking aids. Finally airbag deployment can be more efficient by detecting the occupant positions prior to accidents [45].

2.4.6 Human-computer interfaces, and machine vision

3D image sensors can perform fast acquisition and processing of real-time depth images. Detection and recognition of user gesture in the three dimensional space

is possible using this functionality [90]. The concept can be used in video gaming, videoconferencing and interactive home entertainment devices [45]. For example, by detecting the movement of arms and body in 3D, real physical activities and sports can be simulated. Depth sensing by 3D imaging, combined with grayscale or color imaging can result in more reliable image processing. The applications of machine vision which can incorporate 3D imaging can include robotics and safety in manufacturing [45].

2.4.7 Quantum information processing

Efficient single-photon detection is essential in photonic quantum information processing systems. With the availability of pseudo-single-photon and photon-pair sources, the success of quantum cryptography essentially depends on the ability of detecting single-photons [91]. Since the 1550 nm telecommunication band is the most attractive wavelength from the viewpoint of fiber transmission, InGaAs/InP SPADs are useful for quantum communication [91] and quantum state measurement [92],[93].

Chapter 3

Quantum Effects in Single-Photon Avalanche Diodes

Carrier transport equations in solid-state electronic devices are derived using two different areas of quantum and statistical mechanics [94]. The role of quantum mechanics is calculating the carrier density by introducing the two terms of *1) density of states 2) filling factor of the carriers*.

Given the chemical composition and atomic arrangements in materials, quantum mechanics can compute the electron density by defining the terms of *effective mass*, and *bandgap*. In addition, *Fermi levels* are defined in transport equations to separate the electrons taking role in the conduction process, from bonded electrons [95].

Beside the role of quantum mechanics in carrier transport calculation, some particular quantum behaviors affect single photon counter responses [6]. In this chapter two main quantum effects are discussed being observed in SPADs namely:

Tunneling phenomena: this effect is observed in deep-submicron SPADs implemented in this thesis. The tunneling noise is caused by band-to-band and trap-assisted mechanisms and is analyzed for a new SPAD introduced in $0.35\mu m$ CMOS technology.

Bistability and multistability phenomena: this effect is mainly caused by defects and traps contributing in trap-assisted Shockley Read Hall (SRH) generation, and trap-assisted tunneling, analyzed in $0.8\mu m$ CMOS SPADs and Proton irradiated $0.35\mu m$ SPADs. This effect is known as Random Telegraph Signal (RTS) in solid-state electronic devices and it was observed by us for the first time.

3.1 Tunneling phenomena

The penetration of an electron through a potential barrier is known as tunneling [28]. The tunneling behavior can occur in finite height and thickness potential barrier. Beside the emergence of tunneling in deep-submicron electronic devices, some particular electronic elements exploit tunneling principles for their operation [95]. Figure 3.1 shows a finite potential barrier, which does not force the boundary conditions of Ψ (wave function) to be zero at the barrier. W, and V_0 are the thickness and the potential of the barrier, respectively. The continuity of Ψ and its slope $d\Psi/dx$, at the border of barriers are the boundary conditions. Solving the Schrodinger equation, Ψ results in a nonzero value within the barrier and on the right side of it. Hence $\Psi^*\Psi$ (probability density function) is positive, representing the probability of particle presence beyond the barrier.

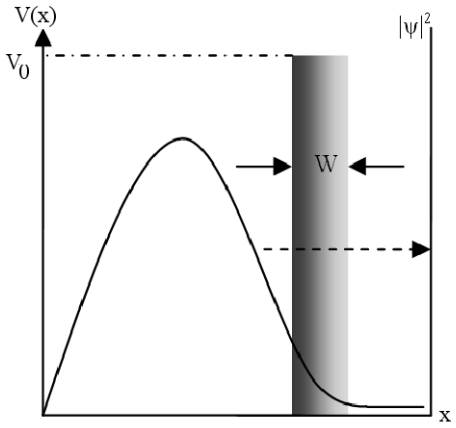

Figure 3.1: *Tunneling probability vs. depth in quantum wells with finite potential barriers.*

Tunneling affects electronic device behaviors in different ways. The penetration of electrons from the valence band through the depletion region and to the conduction band is considered as one of the main tunneling processes affecting SPADs [35]. The probability of this mechanism is highly dependent on the depletion thickness and electric field magnitude. Defining the critical electric field as $6 \times 10^5 V/cm$, some sources suggested considering tunneling, in some particular diode implementations [96]. Reaching the critical electric field is unavoidable in the technology nodes lower than $0.35\mu m$ [35].

Tunneling in SPADs is analyzed and calculated by theoretical analysis and simulations. Moreover, tunneling effects in standard characteristics of SPADs such as I-V curve and DCR are analyzed. It is shown that tunneling can contribute

significantly in some particular SPADs implemented in 0.35 μm standard CMOS technology.

3.2 Theoretical background

Tunneling can contribute to the dark current in photodetectors and DCR in SPADs in two different processes namely *band-to-band (BTBT)*, and *trap assisted tunneling (TAT)* as shown in equation 3.1. Although BTBT is expected in extreme bias conditions, TAT can occur more frequently. The probability of TAT is strongly dependent on the process conditions and steps introducing traps and defects.

$$DCR_{total} = DCR_{BTBT} + DCR_{TAT} \tag{3.1}$$

Figure 3.2 shows the trap role in noise generation of SPADs. A trap being an impurity atom or defect can act as a catalyst for electron and hole generation and recombination. Considering a trap as an atom which lacks one electron, it absorbs electrons from valence band. After the filling process, the charged trap releases its electron to the conduction band, generating dark current. While TAT is considered in this section for n-tub guard ring based SPAD, the trap assisted generation is discussed in the next section for the p-tub guard ring based SPADs.

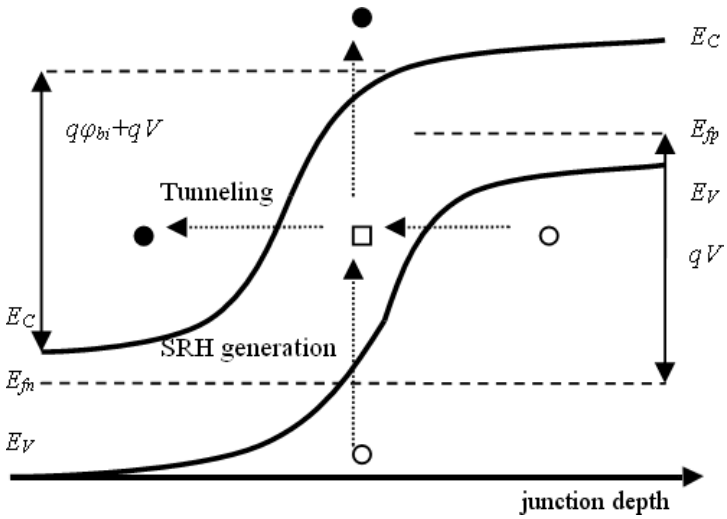

Figure 3.2: *Mechanisms which a trap can affect SPAD behavior (trap assisted tunneling and trap assisted generation).*

TAT can be calculated introducing an integral over the depletion region [34].

The integral consists of the generation rate of electron-hole pairs and the probability that each electron-hole pair may cause an avalanche.

$$DCR_{TAT} = S \cdot \int_{z_0}^{z_w} P_p(z) G_{TAT}(z) dz. \tag{3.2}$$

In this equation, S is the surface between n and p regions, $P_p(z)$ is the probability that an electron-hole pair triggers an avalanche. This probability is useful for both BTBT and TAT calculations and is addressed later in this section. Figure 3.3 shows the cross-section of the junction between n^+ and p-substrate. The elements of S, z_0, z_w and $P_p(z)$ are shown in the figure.

$G_{TAT}(z)$ is the trap assisted tunneling generation rate as calculated in [34]:

$$G_{TAT}(z) = \frac{n_i \cdot (1 + \Gamma(z))}{2\tau \cdot \cosh(\frac{E_t - E_i}{kT})}, \tag{3.3}$$

where Γ is the field enhancement factor of carriers in each depth of the depletion region, n_i is the intrinsic carrier concentration of the material, τ is the minority carrier lifetime, $E_t - E_i$ is the difference between trap energy and the average energy of valence and conduction bands. In this thesis the approximate same value of [34] is used as $E_t - E_i$, k is the Boltzmann's coefficient , and T is the absolute temperature.

The field enhancement factor is calculated using [97]:

$$\Gamma(z) = 2\sqrt{3\pi} \cdot \frac{\xi(z)}{\xi_\Gamma} \cdot \exp((\frac{\xi(z)}{\xi_\Gamma})^2), \tag{3.4}$$

where $\xi(z)$ is electric field at each depth of the substrate, and ξ_Γ is defined as:

$$\xi_\Gamma = \frac{\sqrt{24 m^* (kT)^3}}{q\hbar} \tag{3.5}$$

where q is the charge of an electron, m^* is the effective mass of the electron in silicon, and $\hbar$ is the Planck's constant. The intrinsic concentration n_i of each material depends on temperature T, and band gap E_g [96].

$$n_i \sim T^{\frac{3}{2}} \cdot \exp(\frac{-E_g}{2kT}). \tag{3.6}$$

The bandgap of silicon (eV) is given by [96]:

$$E_g(T) = 1.17 - \frac{4.73 \cdot 10^{-4} \cdot T^2}{(T + 636)} \tag{3.7}$$

Similar to TAT, BTBT is derived from the generation rate and the probability of avalanche creation by each electron-hole pair. The multiplication of the two factors is integrated over the entire depletion region as shown in figure 3.3.

$$DCR_{BTBT} = S \cdot \int_{z_0}^{z_w} P_p(z) G_{BTBT}(z) dz \tag{3.8}$$

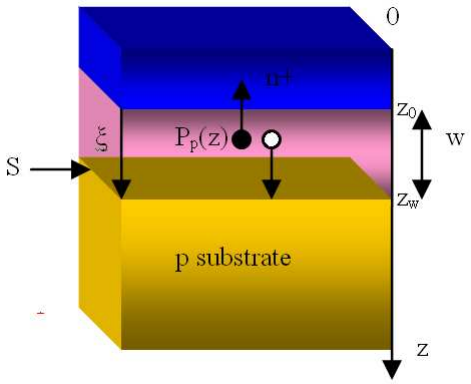

Figure 3.3: *Main elements for tunneling integral calculation in junction cross-section.*

Similarly to DCR generated by TAT in equation 3.2, S is the surface between the n and p regions $P_p(z)$ is the probability that an electron-hole pair triggers an avalanche. The generation rate of BTBT ($G_{BTBT}(z)$) is calculated by [34]:

$$G_{BTBT}(z) = B \cdot |\xi(z)|^{\frac{5}{2}} \cdot D \cdot exp[\frac{-\xi_0}{\xi(z)}] \tag{3.9}$$

In this equation, factor B is $4.10^{14} cm^{-.5} V^{-2.5} s^{-1}$, ξ_0 is a constant which depends on the temperature through the temperature dependence of bandgap and taken to $1.9 \times 10^7 V cm^{-1}$, D is unity except on the edges of depleted region where it vanishes [34].

Finally, for the derivation of the avalanche trigger probability by an electron-hole pair $P_p(z)$ required for BTBT and TAT DCR integrals calculation, the following coupled equations must be solved [98]:

$$\frac{d}{dz} P_e(z) = \alpha_e(z) P_e(z) P_p(z) - \alpha_e(z) P_p(z); \tag{3.10}$$

$$\frac{d}{dz} P_h(z) = \alpha_h(z) P_h(z) P_p(z) - \alpha_h(z) P_p(z); \tag{3.11}$$

where $\alpha_e(z)$ is the electron ionization coefficient and $\alpha_h(z)$, is the hole ionization coefficient. $P_e(z)$, and $P_h(z)$ are calculated using [34]:

$$P_p(z) = P_e(z) + [1 - P_e(z)] \cdot P_h(z) \tag{3.12}$$

3.3 Tunneling in n-tub based guard ring SPADs

In deep-submicron SPADs, tunneling is particularly high due to the high and shallow doping profiles and topologies [99]. To verify our model we built high and low tunneling structures in 0.35 μm CMOS technology.

In the following section these structures are compared: 1) *p-tub guard ring (PTGR)* SPADs, and 2)*n-tub guard ring (NTGR)* SPADs.

Figure 3.4 shows the cross-sections of both structures. Shallow trench isolation is used in combination with low doped p-wells or n-wells to form p-tub and n-tub implants in deep-submicron technologies [71]. In the NTGR SPAD most photons are absorbed in the p-substrate instead of the deep n-well as in the PTGR SPAD. Different sets of circular NTGR SPADs with different types of guard ring layers available in 0.35 μm standard CMOS technology are implemented with different doping concentrations and different thicknesses. Among them one type of SPAD shows the single photon counting behavior being characterized in this thesis. The NTGR SPAD is compared with PTGR realized on the same technology in our group [63].

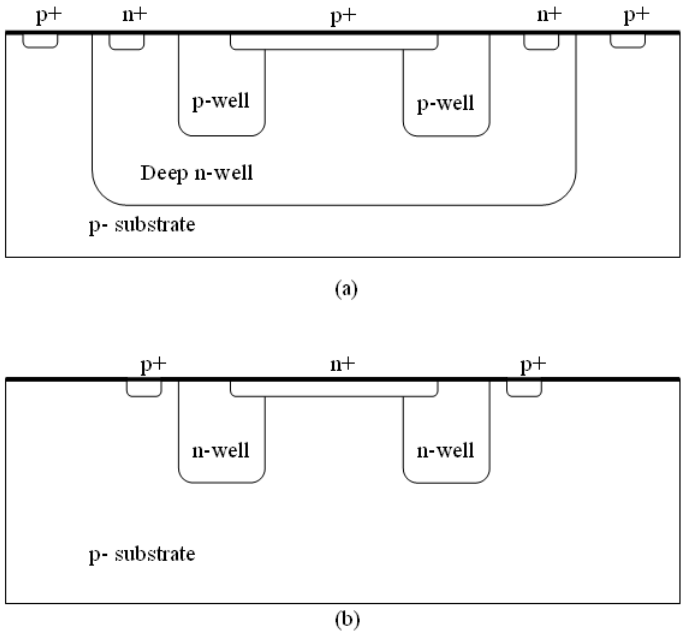

Figure 3.4: *Cross-section of a p-tub guard ring (PTGR) SPAD (a), and a n-tub guard ring (NTGR) SPAD (b).*

3.3.1 SPAD characteristics

Figure 3.5 shows the I-V characteristics in reverse bias region of the NTGR SPAD. The breakdown voltage of this structure at room temperature is 13.6 V, decreased from 17.7 V of the PTGR SPAD integrated with same sizes in this technology. It should be noted that while at relatively low voltages in the I-V curve the current is due to trap-assisted tunneling [100], in middle range voltages, the current is due to band-to-band tunneling [101]. Band-to-band tunneling in the reverse current in CMOS technologies is due to very heavy doping (halo) in the vicinity of the source or drain p/n junctions, supressing short-channel effects by increasing the substrate doping [102].

The breakdown current of the implemented NTGR SPAD is several hundred nanoampere, which is much higher than in previous SPADs [57; 63; 64; 66], and even NTGR structures implemented in 90nm CMOS technology [99]. This high breakdown current shows tunneling current dominance in this topology and it is higher compared to previous SPADs.

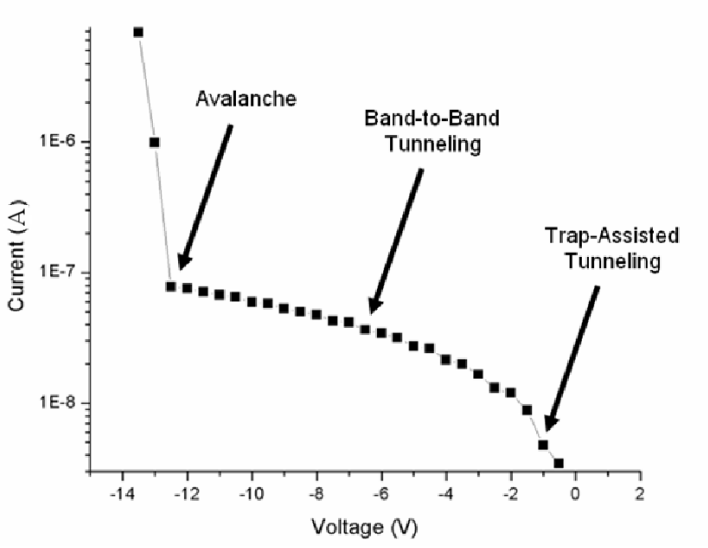

Figure 3.5: *I-V characteristic of a SPAD implemented in 0.35 μm technology.*

Different models are proposed to evaluate generated electrons due to tunneling in reverse biased $p - n$ junction. Among them, Kane's model calculates the generation rate when the local band bending is high enough to make the tunneling probability significant:

$$n = \frac{F^2\sqrt{m_r}}{18\pi\hbar^2\sqrt{E_g}} \exp(\frac{-\pi\sqrt{m_r}E_g^{\frac{3}{2}}}{2\hbar F}), \tag{3.13}$$

Where n is the number of electrons per second per cm^3 leaking from valence band to the conduction band, m_r is the reduced mass considered for combination of electrons and holes, $\hbar$ is the reduced Planck's constant, E_g is the band gap of the material, and F is the magnitude of electric field [103]. The magnitude of the electric field can be calculated both in theory, by considering an abrupt one-sided junction approximation [96], and by simulation, with the exact doping profiles. This equation shows that the electron-hole pair generation in the NTGR structure can reach $3.3 \times 10^{11} pairs/cm^3/s$ and 1.4×10^9 in the PTGR structure.

3.3.2 Tunneling evaluation by device simulations

Both PTGR and NTGR structures are simulated with TCAD^{TM} device simulator to recognize the surfaces contributing in BTBT. Figure 3.6 shows the contours of tunneling generation in the NTGR based SPAD which shows the high contribution of tunneling between the n^+ surface and the p substrate. This tunneling plays a significant role in total tunneling noise since it is integrated over the whole surface as mentioned in Equation (3.8).

Simulation at the same biasing condition is performed for PTGR structures implemented on the same technology. Device simulations show a low contribution of BTBT near the edges of contact between p^+ and guard rings shown in figure 3.6. The inset shows a larger profile of the edge. In addition to the locality of tunneling which does not have a notable effect in Equation (3.8) integral, it does not contribute to total tunneling since the D factor in Equation (3.9) is zero [34].

Extracting simulation predictions, electron-hole pair generation in NTGR structures maximally reaches to 6.1×10^{11} $pairs/cm^3/S$ in the long interface between n^+ and p substrate at normal bias conditions and to 1×10^{12} in PTGR structures only on the edge of p^+ interface with the guard ring.

There are two different approaches to deriving the electric field for Equation (3.8). The first approach is using the single-sided abrupt junction approximation presented in [96], and the second one is the extraction of the electric field values using TCAD^{TM} device simulations [104]. Figure 3.8 shows the electric field distribution of NTGR SPADs at 1.4 V of excess bias voltage extracted from TCAD^{TM} simulations.

Tunneling current and count rate calculations depends on the probability of triggering an avalanche by an electron-hole pair. This probability is calculated at different excess bias voltages as an input to equations 3.2 and 3.8. Figure 3.9 shows P_e, P_h, and P_p as the function of depth at 1.4 volts of excess bias voltage. It should be noted that the boundary conditions force P_e, and P_h to be zero at the start of depletion region and at the end of depletion region, respectively [34].

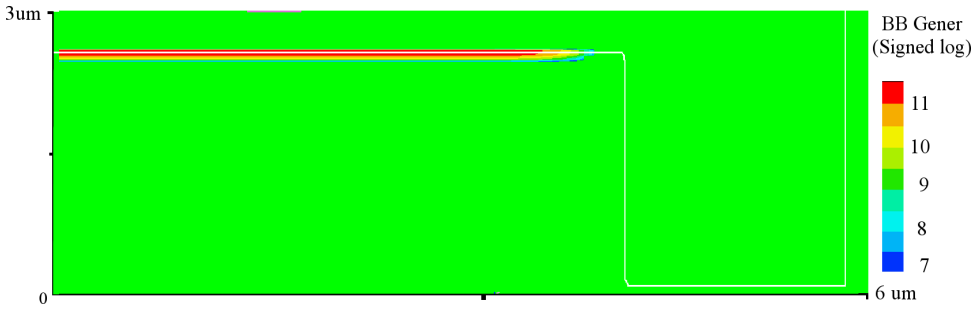

Figure 3.6: *Contours of BTBT generation rate in NTGR SPAD. The picture shows a half of the cross-section of the SPAD.*

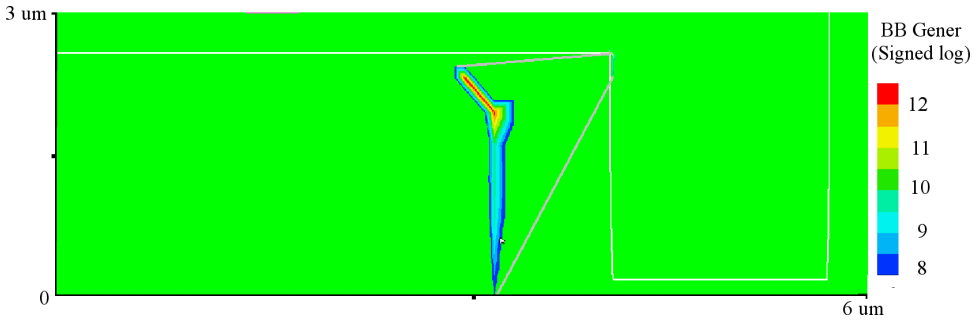

Figure 3.7: *Contours of BTBT generation rate in PTGR SPAD. The picture shows a half of the cross-section of SPAD. The inset shows the tunneling profile on the guard ring edge.*

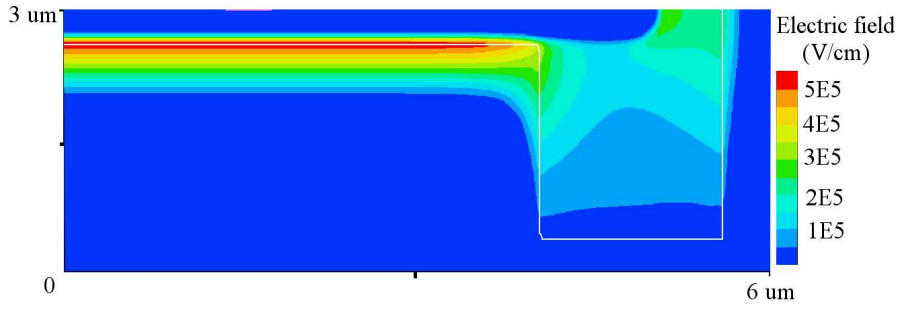

Figure 3.8: *Electric field distribution at 1.4 volts of excess bias voltage in NTGR SPAD.*

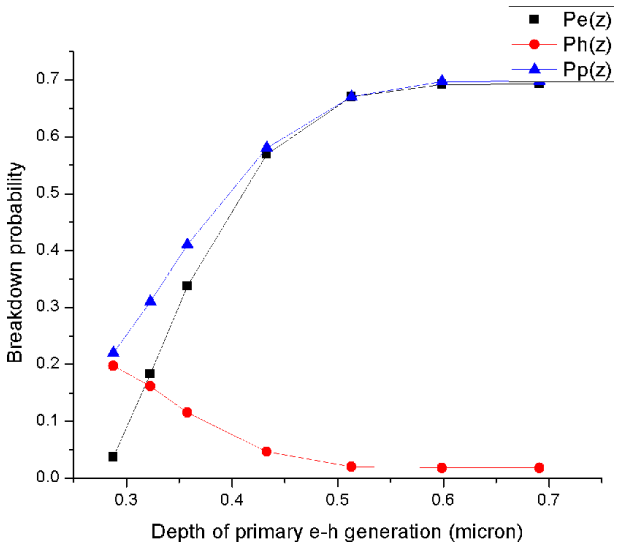

Figure 3.9: *$P_e(z)$, $P_h(z)$, and $P_p(z)$ at different depths. Avalanche trigger probabilities calculated for NTGR SPADs at 1.4 Volts of excess bias voltage. z is the same as Figure 3.3.*

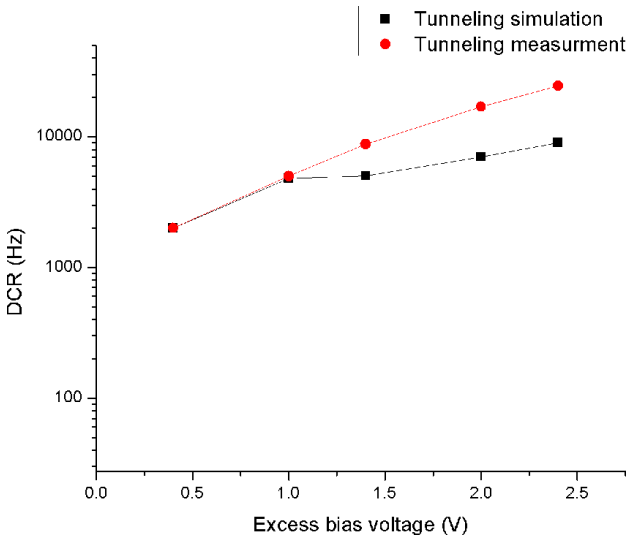

Figure 3.10: *Comparison of theoretical model and measurement of DCR for NTGR SPAD at different excess bias voltages.*

3.3.3 NTGR and PTGR measurement results comparison

Figure 3.11 shows the DCR of a SPAD with different excess bias voltages, measured at different temperatures. As predicted, DCR is higher in NTGR SPADs than PTGR SPADs with the same dimensions [63]. The lower dependency of DCR on temperature in NTGR SPADs suggests dominance of tunneling based dark counts in comparison with PTGR devices.

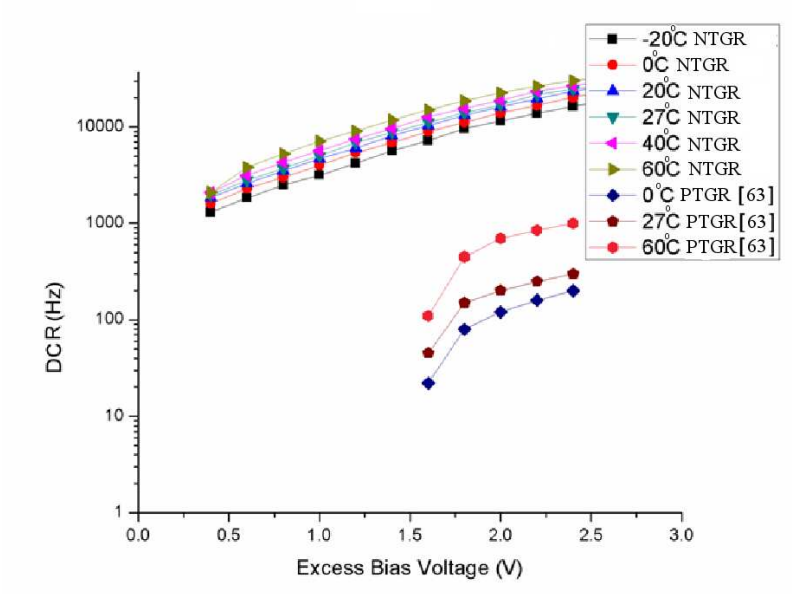

Figure 3.11: *DCR behavior at different excess bias voltages as a function of temperature.*

Figure 3.12 shows the photon detection probability (PDP) measured with a monochromator at room temperature at different wavelengths. While the excess bias voltage is being reduced from 4V in [63], to 1.4 V in this technology the maximum PDP is increased by 3 percent with regard to PTGR SPADs. The wavelength which shows the highest amount of PDP is 420 nm, the same as [63]. The higher PDP can be due to a thicker depletion region on the p-substrate side, which is less doped than deep n-well in PTGRs.

Timing jitter was measured for the NTGR SPAD. Two fast laser sources with a pulse width of 40 ps and repetition rate of 40 MHz emitting at of 637 nm and 405 nm respectively, were used for the jitter measurement. The time interval between the laser output trigger and the leading edge of the SPAD signal was measured via a high performance oscilloscope operating as a TDC.

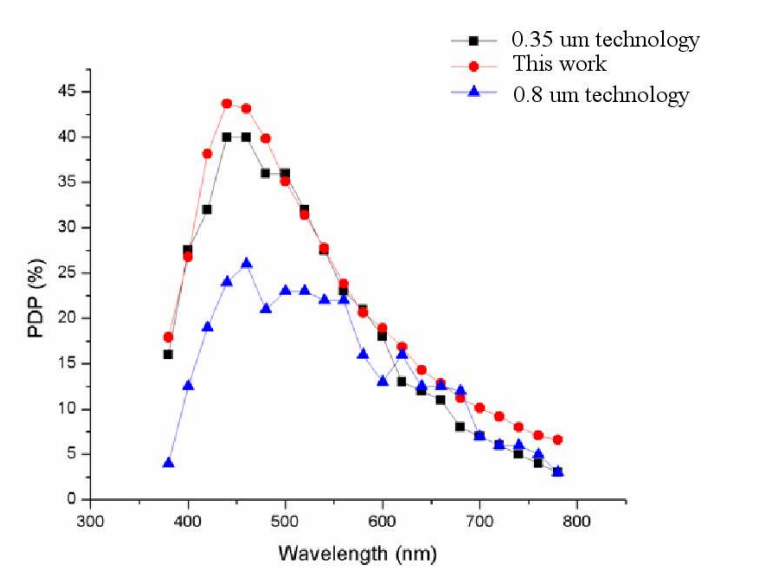

Figure 3.12: *Photon detection probability of NTGR SPAD at different wavelengths.*

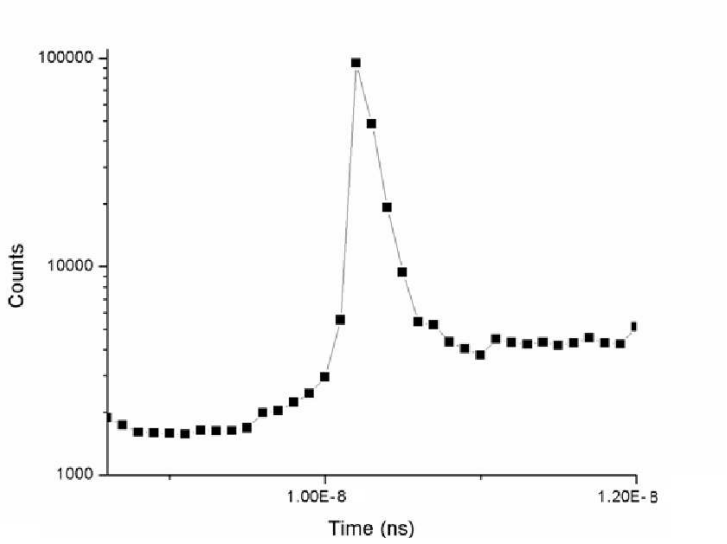

Figure 3.13: *Output signal of a SPAD exposed to 40 MHz pulsed laser for jitter measurements at 405 nm wavelength at room temperature.*

A histogram was constructed from the time difference of the edges of the laser light and sensing signal. Figure 3.13 shows the histogram of the time difference at wavelength of 405 nm. The Full Width at Half Maximum (FWHM) of the jitter is measured as 157 ps and 152 ps at wavelengths of 637 nm and 405 nm, respectively. Higher jitter is expected due to higher thickness of the depletion region and higher tunneling counts.

Figure 3.14 shows the afterpulsing measurement of the NTGR SPAD in time domain. Afterpulsing increases due to the higher number of carriers involved in an avalanche and higher defect density. For the same reasons, crosstalk also increases in deep-submicron SPADs (both optical and electrical).

In addition to the presence of separate deep n-wells in [63] isolating SPADs from each other, integration of an inverter on chip reduces the parasitic capacitance and crosstalk simultaneously. Afterpulsing degradation by 7 percent is observed in this technology consistently with [99], where the inverter was not integrated on chip.

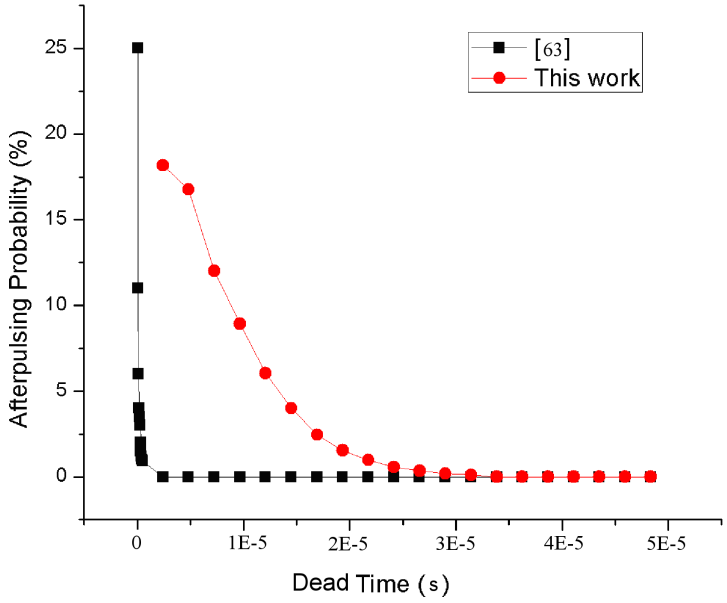

Figure 3.14: *Afterpulsing probability of NTGR SPAD vs dead time.*

To evaluate the guard ring in SPADs, photoemission was measured. This method shows the guard ring effectiveness and homogeneity of the electric field across the active region [35]. Since emission of photons occurs during an avalanche as a byproduct of impact ionization, that leads to enable a quantitative assessment

of the effectiveness of the guard ring [35]. An Avalanche current of 100 μA was running through SPAD to capture photoemission using a microscope and a standard CCD camera with the collection time of 1 s.

Figure 3.15 shows the photoemission experiment of a non-functional guard ring (NTGR SPAD). The figure shows most of the photoemission occurring on the guard ring which is an indication of PEB. Figure 3.16 shows photoemission from the light sensitive region of a functional guard ring (NTGR SPAD) characterized in this section. The non-uniformity of light emission which is observable in the low excess bias volatges, is due to non uniformity of doping in the guard ring [35]. The complexity of these lateral non-uniformities was not included in the TCAD^{TM} simulations.

Performance	This thesis (NTGR)	[63] (PTGR)	Comments
DCR	10 kHz	15 Hz	At 1.5 V excess bias voltage at room temperature
PDP	43%	40%	Maximum PDP
Timing jitter (FWHM)	152 ps	80 ps	At 405 nm illumination
Afterpulse probability	25 %	18 %	At nominal dead time
V_{br}	13.6 V	17.7 V	Breakdown voltage at room temperature

Table 3.1: *Comparison of NTGR and PTGR main characteristics, implemented in 0.35 μm standard CMOS technology.*

3.3.4 Conclusion

In this section the tunneling effects of two main families of SPADs were analyzed. The fundamental equations for tunneling count rate in SPADs are presented. Moreover, the evaluation is performed by implementing a new n-tub guard ring SPAD in 0.35 μm standard CMOS technology. The NTGR SPAD is characterized theoretically by simulations, and experiments. Table 3.1 shows the main characteristics of NTGR and PTGR SPADs designed in the same technology with identical sizes. While DCR is increased due to the tunneling, PDP is improved by 3 percents due to depletion region expansion. Figure 3.17 shows the photomicrograph of the SPAD implemented in this work.

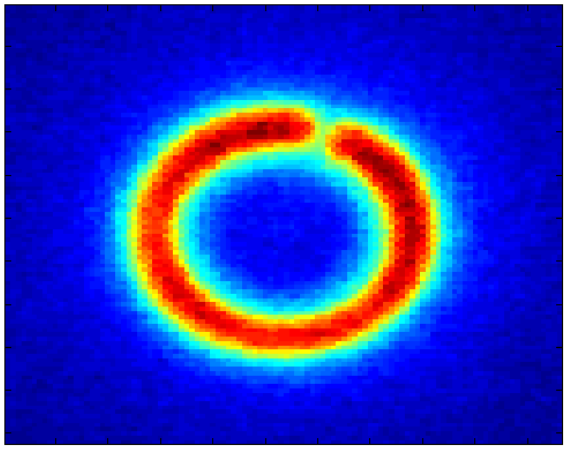

Figure 3.15: *Photoemission of a non-functional NTGR SPAD. The figure shows the light emission from the guard ring due to PEB.*

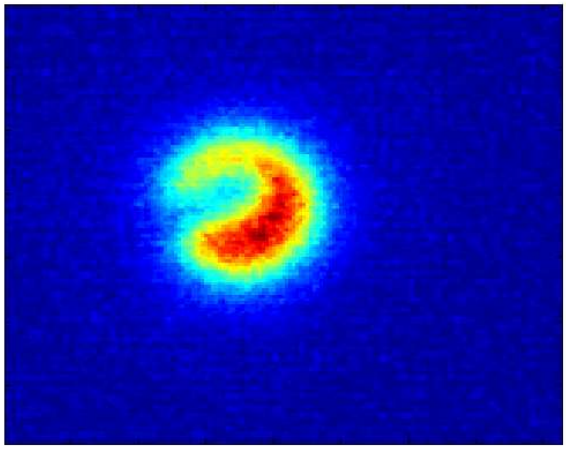

Figure 3.16: *Photoemission from the light sensitive area of a working NTGR SPAD. The figure shows virtually no PEB on the guard ring.*

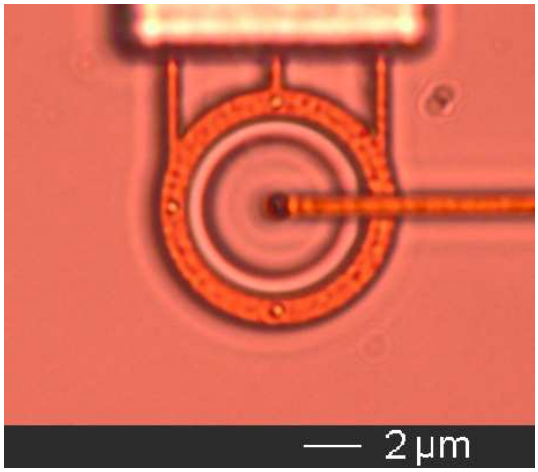

Figure 3.17: *Photomicrograph of the circular NTGR SPAD characterized in this section.*

3.4 Random Telegraph Signal (RTS) in SPADs

In addition to the mentioned noise generating processes, the direct excitonic recombination, Auger recombination, and trap assisted generation can be regarded as sources of carrier generation and recombination in solid state electronic devices [105].

The trap assisted generation and recombination is considered among indirect processes of generation or recombination in solid state electronic devices [94]. The indirect processes of recombination occurs in indirect bandgap semiconductors (Si, Ge). Figure 3.18 shows the energy-wavevector diagram of three main semiconductors. Traps have a strong role in recombination in Si, and Ge due to the indirect bandgap of the materials. Although a photon having enough energy is sufficient to bring an electron from valence band to conduction band, in direct semiconductors, a wavevector match is required between an electron going from the top of valence to bottom of conduction band in indirect semiconductors. The role of traps or defects is to take an electron from valence band with the same wavevector of the valence band top, and then exchange with the electrons in the conduction band with the help of phonons, giving enough wavevector.

It is shown that trap levels close to the mid-gap can help in a trap assisted generation process. In contrast, the trap levels which are close to the conduction band like a donor or the trap levels which are close to the valence band like an acceptor, are not contributing significantly to trap assisted generation. Hence most effective traps are materials which band diagrams lie in the middle of silicon

bandgap (such as Cu, Au) [105].

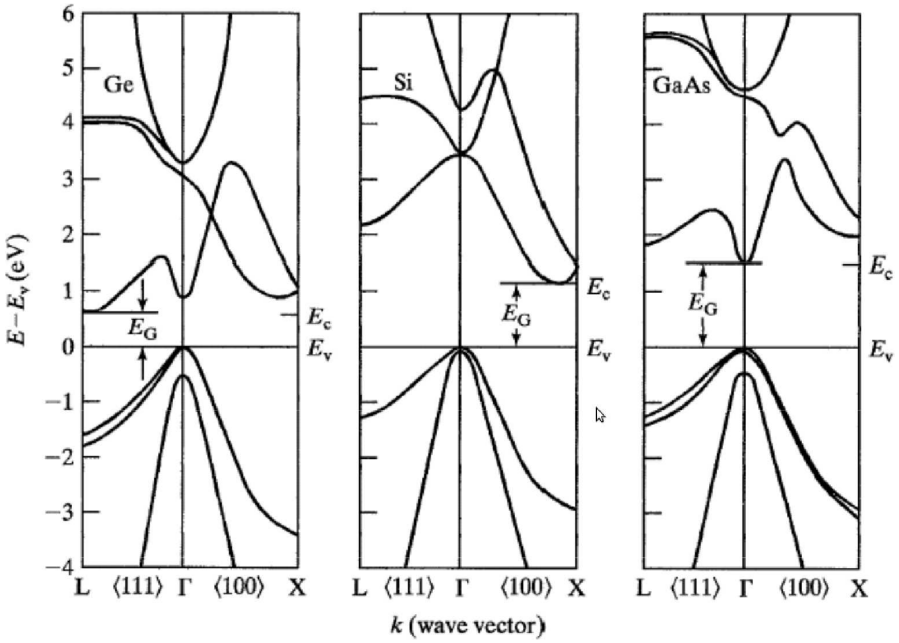

Figure 3.18: *The energy vs wavevector of indirect (Si, Ge), and direct (GaAs) semiconductors [94].*

The trap assisted generation or trap assisted tunneling (TAT) can produce Random Telegraph Signals (RTS) in the main characteristics of SPADs. RTS is observed in the count rate of a SPAD fabricated in 0.8 μm CMOS technology and in four proton irradiated SPADs designed and fabricated in 0.35 μm CMOS technology. In this section RTS characteristics are evaluated experimentally and verified theoretically with respect to bias and temperature.

3.4.1 RTS fundamentals

When the charge transport through a solid-state device is controlled by changing a trap state, defect configuration, or a cluster of defects, it gives rise to a discrete switching of the current, called burst or popcorn noise, or more often random telegraph signal (RTS) [106]. Although RTS is important for small-area scaled devices, researchers have shown that the superposition of RTSs will result in flicker (1/f) noise in large devices [107].

Understanding and modeling DCR instability may provide a powerful tool to predict and analyze defects in standard CMOS processes while suggesting new

techniques to overcome their effects by design. We describe temporal fluctuations of DCR between different states in SPADs. The statistics of such fluctuation follows a RTS behavior.

RTS phenomena have been observed in currents of MOSFETs, JFETs, DRAMS, and in proton-irradiated CCDs, neutron irradiated avalanche photodiodes, proton and heavy-ion irradiated active pixel sensors [108],[109]. Moreover, the presence of individual bistable defects in APDs, and neutron irradiated APDs was suggested when the APDs were biased 10 V above their 250 V breakdown voltage [110]. The different configuration of a trap, defect, or cluster of defects causes the superposition of several levels of fluctuation (multistable) or a two level fluctuation [111]. The bias and temperature dependence of a bistable RTS is analyzed in a SPAD design fabricated in 0.8 μm CMOS technology [112]. Moreover the main characteristics of RTS is analyzed in a proton irradiated SPAD array fabricated in 0.35 μm technology [113]. It should be noted that the RTS behavior could not be observed among the non-irradiated SPAD arrays fabricated in 0.35 μm technology.

Figure 3.19 shows the cross-section of a particular implementation of a SPAD in the 0.8 μm process which shows RTS behavior in some pixels without any irradiation. Figure 3.20 shows the SEM and photomicrograph of the SPAD implemented in 0.8 and 0.35 μm technology respectively.

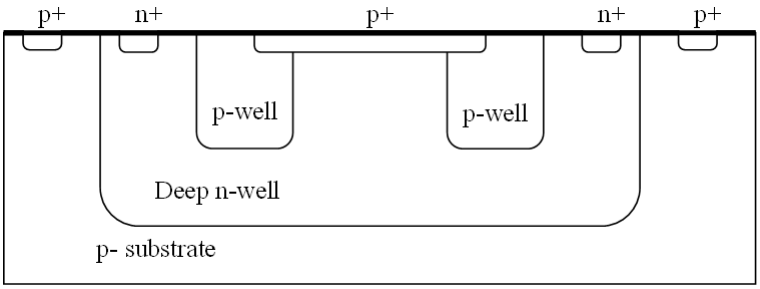

Figure 3.19: *Cross-section of a SPAD implemented in 0.8 μm CMOS process.*

A 32 × 32 SPAD imager designed and implemented in 0.8 μm CMOS technology is evaluated for observing the RTS behavior [58]. Figure 3.21 shows the DCR distribution in each pixel of the chip. Figure 3.22 shows the RTS behavior of DCR observed in one of the pixels of the imager. Each frame in this figure consists of the summation of counts in a constant time span.

The figure shows two frequencies between which DCR is oscillating. The transition from one to the other frequency is abrupt and has a statistical behavior consistent with RTS.

Figure 3.23 shows DCR for three different proton irradiated pixels of a SPAD fabricated in 0.35 μm CMOS technology. While the third pixel shows the standard

bistable behavior, the other two show multistable and multi-bistable behaviors. The autocorrelation and temperature behavior of DCR fluctuations enabled us to show that the RTS is free from afterpulsing [112]. Moreover, the power spectral density of DCR shows a Lorentzian spectrum which is a necessary but not sufficient condition of RTS signals [112].

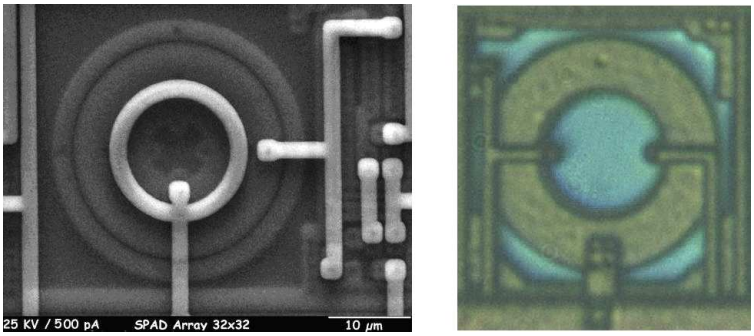

Figure 3.20: *Scanning Electron Microscope (SEM) image of a SPAD implemented in 0.8 μm CMOS process, and photomicrograph of a SPAD implemented in 0.35 μm CMOS process which shows the RTS behavior after irradiation.*

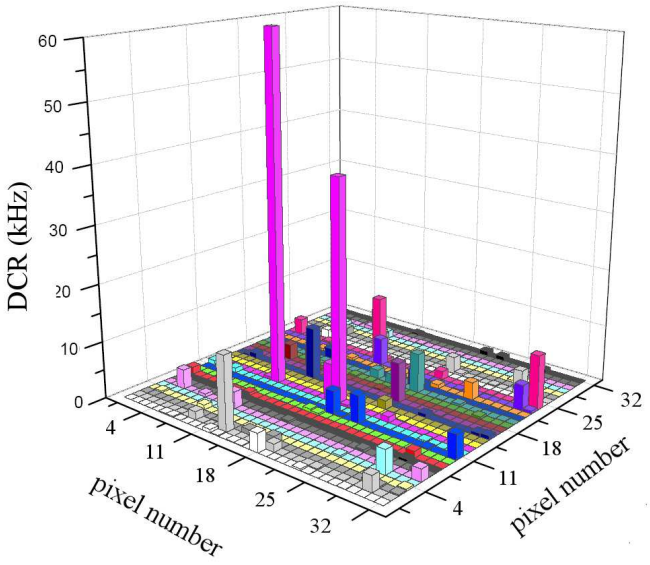

Figure 3.21: *DCR distribution among the pixels of the 32×32 SPAD array.*

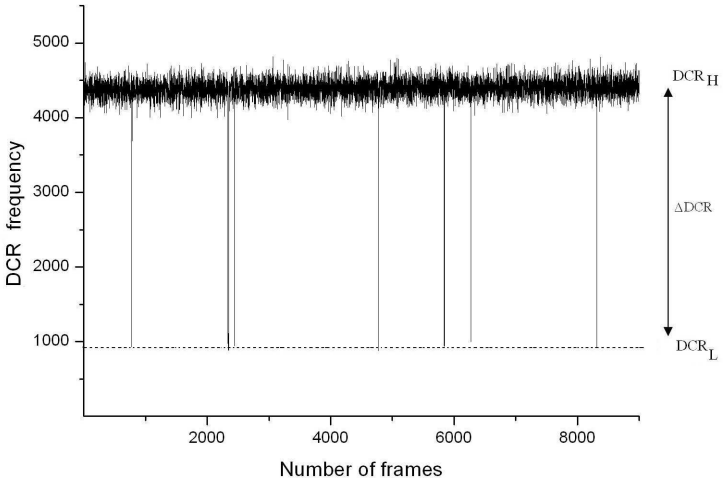

Figure 3.22: *Bistable DCR levels in* $0.8\mu m$ *CMOS non-irradiated SPAD.*

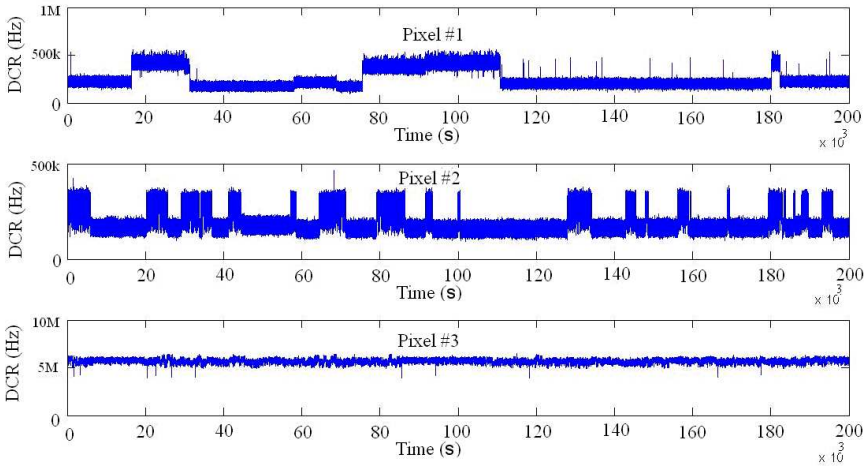

Figure 3.23: *Different forms of RTS in 3 pixels of a proton irradiated SPAD.*

A simple two-level RTS is defined by three parameters: the time spent in the up (or high) state t_u, the time spent in the down (or low) state t_d, and the amplitude [106]. Considering the RTS behavior in SPADs, the up state occurs when the trap in the bandgap is not filled, and the down state is when the trap is filled with an electron. When the trap is not filled it contributes to the electron

exchange in trap assisted generation and trap assisted tunneling mechanisms.

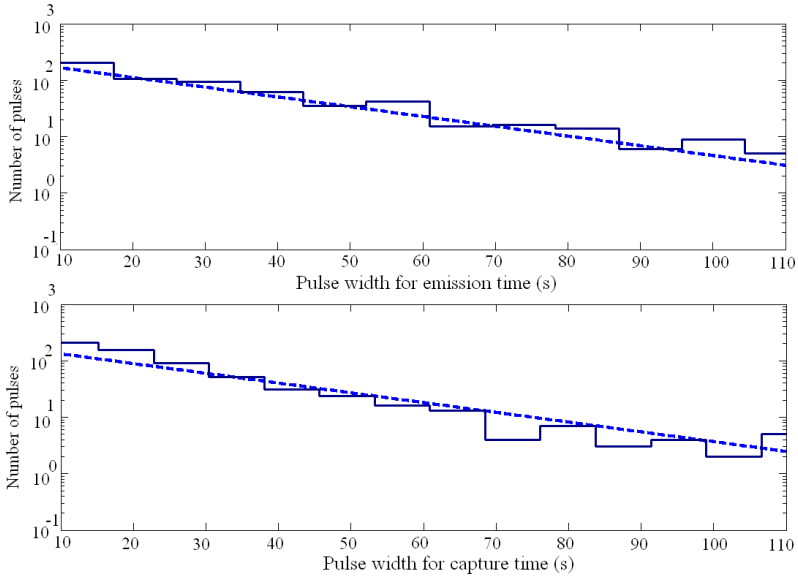

Figure 3.24: *Pulse width histograms obtained by sorting and counting the capture and emission times of a bistable pixel at room temperature for 8hrs. Dashed lines are the expected number of pulses estimated by equation (3.14).*

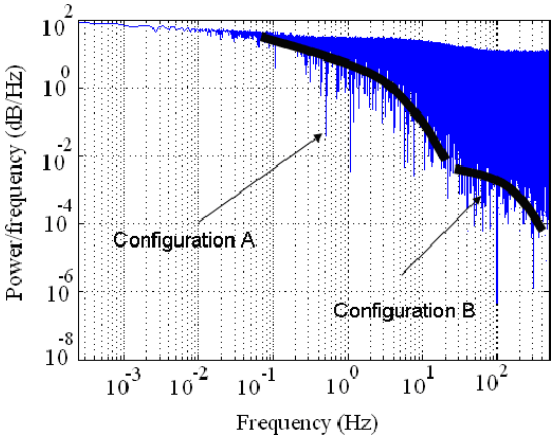

Figure 3.25: *Different configurations of the traps or defects effects in the power spectral density of the DCR.*

$$P_{up,down}(t) = \frac{1}{\tau_{up,down}} \cdot \exp(\frac{-t}{\tau_{up,down}}) \tag{3.14}$$

The exponential behavior in the time distributions of both up (capture) and down (emission) states being observed in the timing characteristics of SPADs is in accordance with equation (3.14) [106]. In this equation $P_{up,down}$ is the probability of up state or down state in a typical RTS. Figure 3.24 shows the matching of the pulse width histogram of the RTS behavior with the exponential characteristic considering the emission time as 25 seconds, and capture time as 18 seconds.

The power spectral density is another indication of RTS noise. When the DCR is due to flicker noise, or to RTS noise, the slope of $1/f$ or $1/f^2$ can be recognized in the power spectral density. Again, this observation is not in itself sufficient to prove the nature of noise. This change in power spectral density of the DCR is observed in SPADs [112]. Moreover, the power spectral density of DCR in some pixels shows different configurations of traps or defects, causing the multi-stable behavior as shown in figure 3.25 [114].

3.4.2 RTS effect on main characteristics of SPADs

The RTS behavior is observed in a pixel of a 32 × 32 SPAD imager designed and implemented in $0.8\mu m$ CMOS technology, and in four pixels of a 32 × 32 SPAD imager designed and implemented in $0.35\mu m$ CMOS technology. The latter is irradiated with a 11MeV proton source [58], and [115]. Normally the irradiated pixels show some temporal behaviors after the irradiation process. The temporal behavior vanishes after certain annealing processes. The RTS behavior being reported in this work is among stable behaviors which remains in the SPADs after several steps of annealing and long disposal time in different temperatures.

Traps and defects are identified by activation energy introducing in the host material. The activation energy suggests about the kind of trap material which is introduced in the fabrication process [28]. The activation energy for the traps and defects observed by analysing the RTS in SPADS are shown in Table 3.2. The activation energies are calculated by considering the Arrhenius equation [111].

$$\frac{1}{\tau_{up,down}} = R \cdot e^{\frac{E_{act}}{kT}} \tag{3.15}$$

In this equation E_{act} is the activation energy of the trap, R is the rate constant τ_{up} is the time constant for the up state, and τ_{down} is the time constant for the down state. In other words τ_{down} is the average time constant taken to fill the trap, and τ_{up} is the average time constant which takes from the trap to release it's electron.

$E_{act}(eV)$	up state	down state
$0.8\mu m$ chip	0.6108	0.7233
pixel No.1 $0.35\mu m$ chip	0.4592	0.4286
pixel No.3 $0.35\mu m$ chip	0.5243	0.4572

Table 3.2: *The activation energy calculated by equation 3.15 for the pixels which show RTS behavior.*

The activation energy of the traps and defects in these cases are close to the midgap of the silicon. This closeness of the activation energy and midgap is the main reason for the contribution of the traps in the avalanche, and impact ionization process [94]. Moreover the difference between the activation energy of 0.8 μm chip and 0.35 μm irradiated chips is a sign of different origins of traps and defects.

Analyzing the RTS behavior of the pixels by increasing the reverse bias voltage shows the appearance of new DCR levels in the DCR histogram (Figure 3.26). This behavior is being observed in some of the pixels of proton irradiated chip. For generating the histogram, the DCR of the pixels are observed in continuous time at room temperature.

By considering the low probability of multiple electron trapping at one defect site in conventional sensors [116], the appearance of these new levels is due to the existence of a second trap with a slightly different energy level. This trap causes the emergence of a new DCR level when the excess bias voltage reaches a certain threshold. Moreover the shift in the frequency of the DCR in the figure with regard to excess bias voltage, is due to the increase in the breakdown probability.

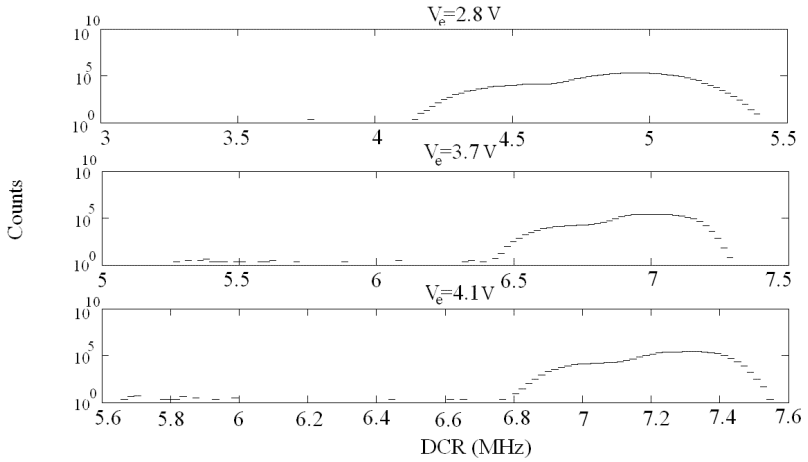

Figure 3.26: *DCR histogram of a pixel: DCR levels number increasing vs V_{excess}.*

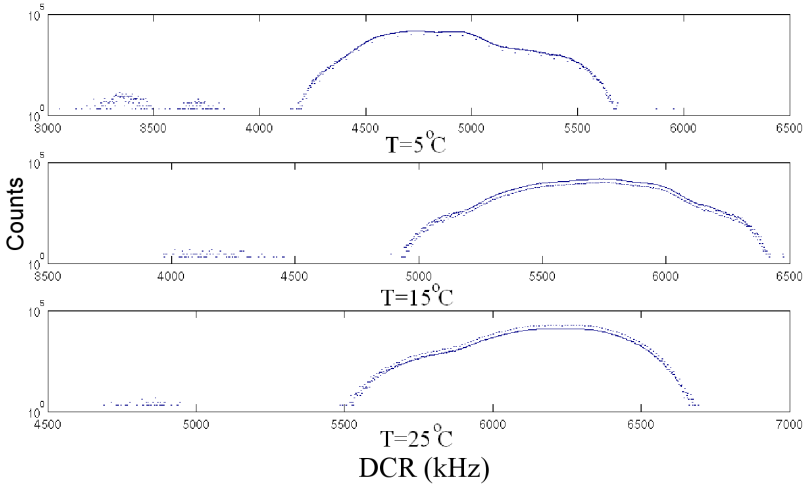

Figure 3.27: *DCR histogram of a pixel: DCR levels number decreasing vs temperature.*

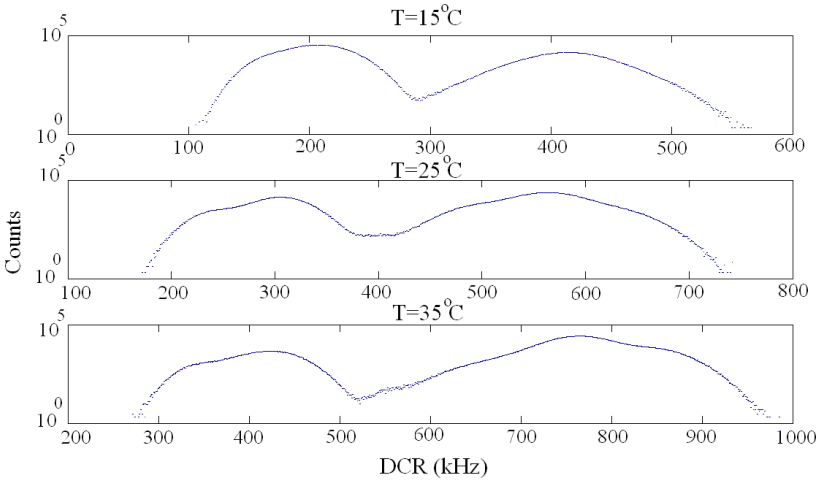

Figure 3.28: *DCR histogram of another pixel: An increase in DCR levels number in temperature.*

Figure 3.27 and 3.28 show a histogram of DCR at different temperatures for two different pixels. Figure 3.27 shows that while there are five distinct levels

for DCR at 5^oC and their superposition in histogram, the number of DCR levels decreases to two levels in 35^oC. In contrast, Figure 3.28 shows that the number of different DCR levels increase to five by increasing the temperature from 15^oC, to 45^oC.

The trap location in the band diagram defines the number of DCR levels in the histogram. The location of the trap which causes reducing of DCR levels by increasing temperature is labeled as 1 in figure 3.29 , band diagram. By considering the location as the temperature increases the probability of occupation of the trap in this location by an electron increases and correspondingly the SRH center is more frequently filled. By filling the SRH center, some DCR levels would decrease. The location of the trap which results in appearing of new DCR levels by increasing the temperature is labeled as 2 in figure 3.29. It is obvious from the Fermi-Dirac function at the location of the trap that by increasing the temperature, the probability of trap occupation is reduced and thus a new DCR level will appear.

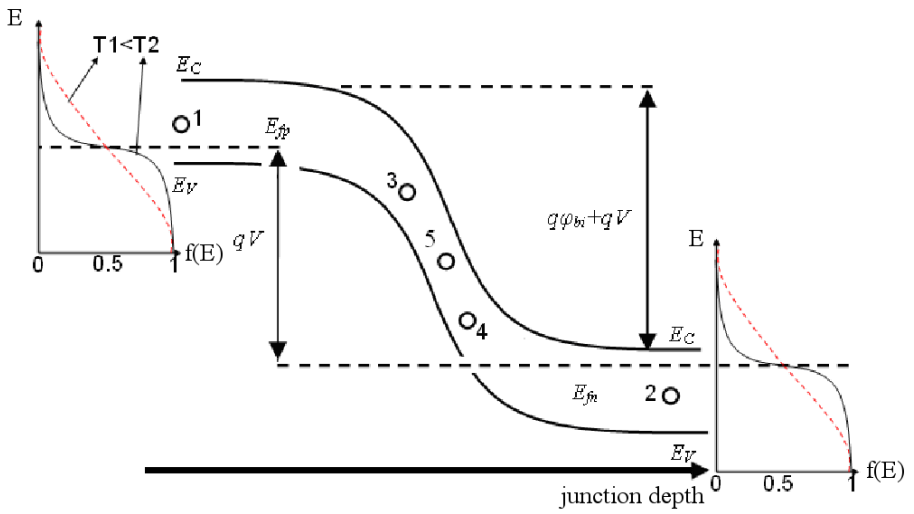

Figure 3.29: *Changing in Fermi-Dirac distribution of electrons in the band diagram of the SPAD. $f(E)$ is the Fermi-Dirac distribution function.*

Differet pixels show different behaviors by changing the temperature. When no new DCR levels appear in the histogram as a function of temperature, it implies that the traps are located in the center which changing temperature does not affect them in the normal bias conditions.

Figure 3.30 shows the temperature dependency of the RTS amplitude. The one-dimensional Poole Frenkel theory describes RTS variation in the electrical

current of devices by introducing a factor χ as described in equation 3.16 [117]. In this equation k is the Boltzmann's constant, ε_s is the dielectric constant of the silicon, q is the charge of an electron, and $|E|$ is the strength of the electric field.

$$\chi = \exp(\beta), \beta = \frac{\sqrt{q^3|E|/\pi\varepsilon_s}}{kT} \tag{3.16}$$

The same behavior is being observed in DCR instead of dark current which can be explained using the same theory. The clear difference between the two different technologies can be explained by the difference in electric field. For the pixels from the same technology the difference in slope can also be due to the same reason or measurement inaccuracies.

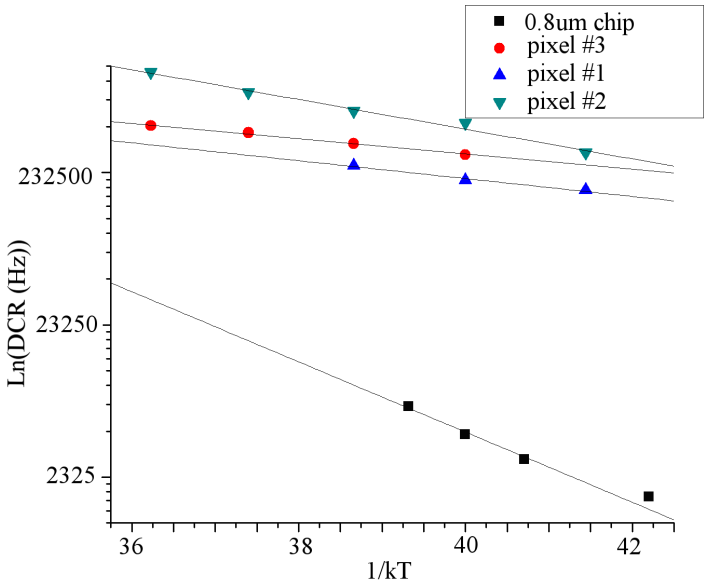

Figure 3.30: *DCR RTS amplitude vs. temperature in four different pixels at constant excess bias voltage. The fitting function is* $y = ae^bx$, *where a, and b are equal to 3.3e12, and -0.53 for* $0.8\mu m$ *chip, 3.75e5, and -0.022 for pixel No.3, 2.04e7, and -0.1352 for pixel No.1, and 1.92e6, and -0.05 for pixel No.2.*

Figure 3.31 shows the increase in RTS amplitude as a function of excess bias voltage. In the conventional trap-assisted tunneling model, the increase of electric field is being described by considering the time constants for tunneling between the valence or conduction band ($\tau(t)$) and the trap described by equation 3.17 [118]. In this equation q is the magnitude of electronic charge, F is the magnitude of electric field at the defect site, m^* is the effective mass for the electron or hole

in th silicon, ΔE is the ionization energy or trap depth, and $\hbar$ is the Plancks constant [112].

The DCR increases rapidly in accordance with decreasing the tunneling time constant between the valence and conduction band. This variation in the tunneling time constant describes the increase in electrical current of the device. We have observed the same behavior in DCR which can be explained by the same theory.

$$\tau_t^{-1} = \frac{qF}{2\sqrt{2m^*\Delta E}} \times \exp(-\frac{8\pi(2m^*)^{1/2}\Delta E^{3/2}}{3qhF}) \tag{3.17}$$

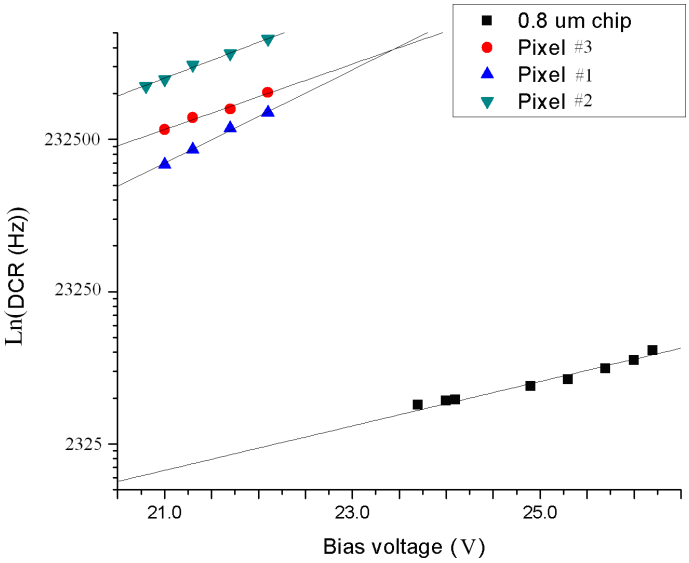

Figure 3.31: *DCR RTS amplitude vs bias voltage at room temperature. The fitting function is* $y = ae^bx$*, where* a*, and* b *are equal to 0.5783, and 0.336 for* $0.8\mu m$ *chip, 3.52, and 0.4954 for pixel No.3, 0.026, and 0.704 for pixel No.1, and 2.89, and 0.541 for pixel No.2.*

3.4.3 Conclusion

Random Telegraph Signal generated by trap assisted generation and trap assisted tunneling is observed in both non-irradiated and irradiated SPADs. The RTS behavior is recognized as the DCR bistability and multistability. The RTS hypothesis of the fluctuations is verified by measurements that are in excellent agreement with the theory. RTS is characterized at different bias voltages and

different temperatures. The RTS behavior is expected to affect large array of SPAD detectors being used for 3D imaging applications.

Chapter 4

Deep-Submicron Single-Photon Avalanche Diodes

The implementation and characterization of SPADs in a 90nm standard CMOS technology is described in this chapter. To the best of our knowledge, this is the first time SPADs are implemented in a feature size smaller than 130nm [119]. At the time of the writing of this thesis, a second successful attempt has been reported of 90nm CMOS SPADs [120].

By implementing different SPAD structures, we could study the geometric trade-offs involved in the design of deep-submicron SPADs. 162 SPADs with different arrangements of doping layers and different guard ring sizes were implemented. Among them, 45 structures are functional with a range of well-defined, reproducible breakdown voltages. The implemented structures were also simulated and characterized.

The next sections describe the design process, modeling, and characterization of the proposed SPAD structures. Simulation considerations of SPADs designed in 65nm CMOS technology are described in the last section. The design of SPADs in 65nm technology is based on the characterization results of functional SPADs in 90nm technology.

4.1 Design process

4.1.1 Design, simulation, and implementation of the SPAD farm

Three Different families of SPADs were implemented in 90nm standard CMOS technology. (1) *p-well based guard ring structures*, (2) *substrate based guard ring structures* (3) *n-well based guard ring structures.*

1) p-well based guard ring structures

Different p-well guard rings used for different breakdown voltages are used in combination with STI. Figure 4.1 shows the structure proposed in [35]. A polysilicon layer is used to prevent STI creation near the active region. Different n-wells are used as the absorption region.

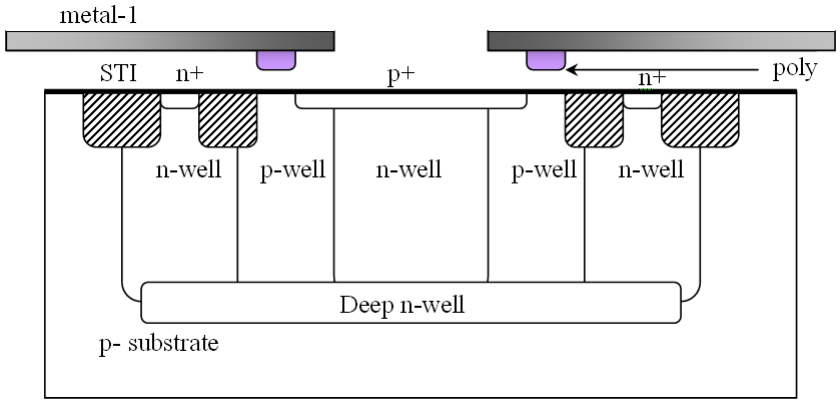

Figure 4.1: *Cross-section of the p-well based guard ring SPAD in 90nm CMOS technology.*

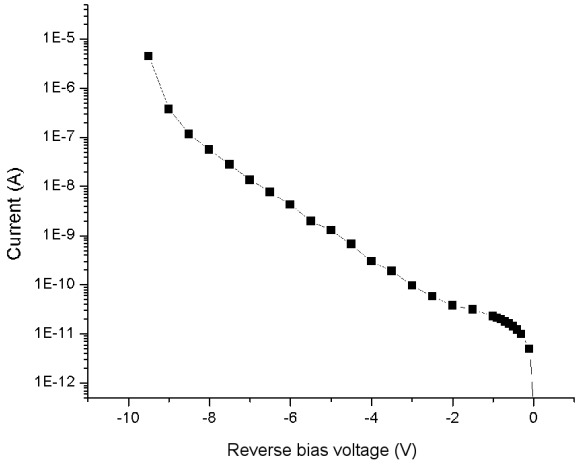

Figure 4.2: *I-V characteristic of the p-well guard ring based SPAD.*

Spectra imaging device simulator [121] is used for simulating the SPAD. The device simulations are used to find the best dimensional values minimizing PEB. Figure 4.4 shows the electric field distribution around the breakdown voltage as computed by Spectra. The figure shows the maximum electric field in the active region. Some hot spots are located on guard ring edges which may cause PEB.

Figure 4.2 shows the measured I-V characteristic of one such structure in reverse bias mode. The current is originated from trap assisted tunneling, band-to-band tunneling and avalanche impact ionization process respectively, with increasing reverse bias [100]. These regions of operation were discussed in earlier sections. The transition between band-to-band tunneling and the avalanche is not sharp enough for SPAD to operate in Geiger mode.

2) substrate based guard ring SPADs

The second family of SPADs is based on defining no additional guard ring explicitly. In other words, spacing between the active region and STI is achieved by substrate intrinsic doping. This structure is based on connecting n-wells surrounding an island of substrate as guard ring. Deep n-wells can connect two n-wells and isolate the guard ring between the n-wells. Two separate n-wells can also be connected directly without using deep n-wells by careful positioning of doping profiles.

Figure 4.3 shows the cross-section of a p-substate based guard ring structure. Different n-wells and deep n-well arrangements are used to form the guard ring.

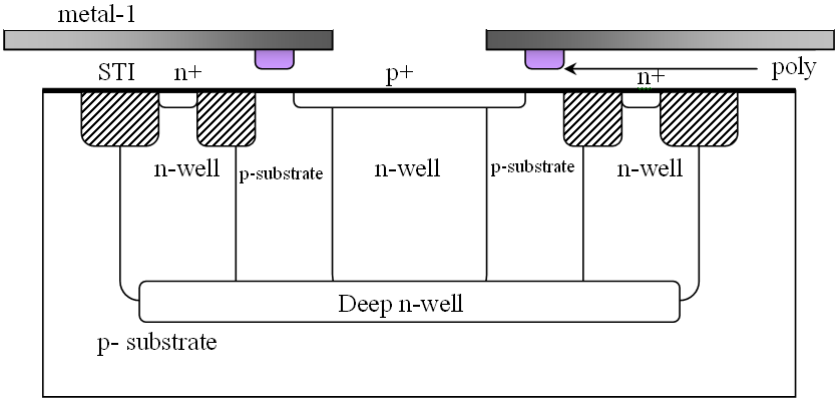

Figure 4.3: *Cross-section of p-substrate based guard ring SPAD in 90nm technology.*

Figure 4.5 shows a Spectra simulation of the electric field distribution around the breakdown voltage. It shows that the maximum electric field is occurring in the active region which is a necessity for operation. The connection between the n-wells also show a high electric field region which can cause the edge breakdown.

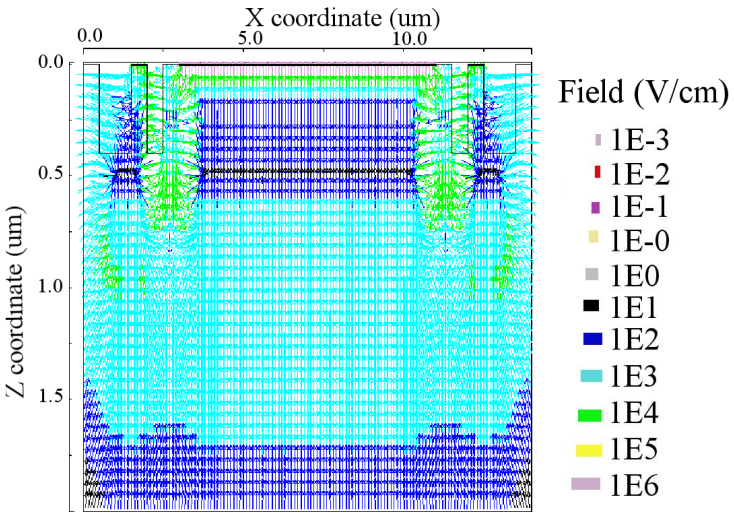

Figure 4.4: *Electric field distribution on the p-well guard ring based SPAD.*

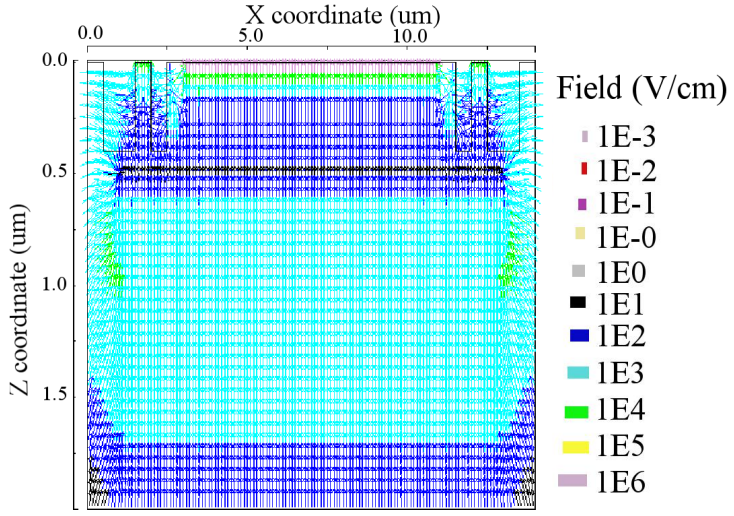

Figure 4.5: *Spectra simulation of the electric field distribution of the p-substrate guard ring based SPAD in 90nm technology.*

Figure 4.6 shows the measured I-V characteristic of one of the structures in reverse bias. This structure also could not show the sharp difference between the

tunneling region and the avalanche region. Smoothness of the transition makes Geiger mode operation impossible using different ballast resistances as quenching element. The main reasons for these types of diodes not to work as single-photon detector are analyzed later in this section.

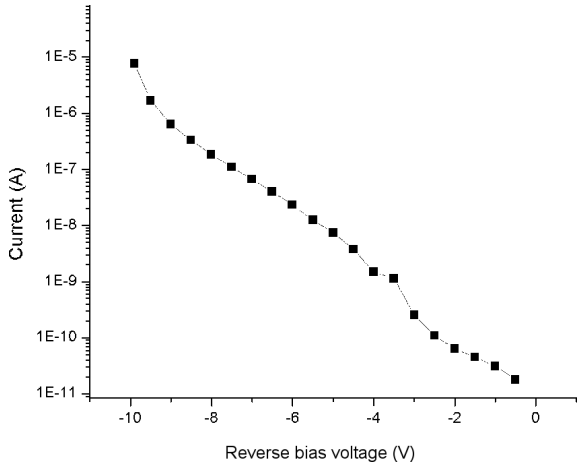

Figure 4.6: *I-V characteristic of the p-substrate guard ring based SPAD in 90nm technology.*

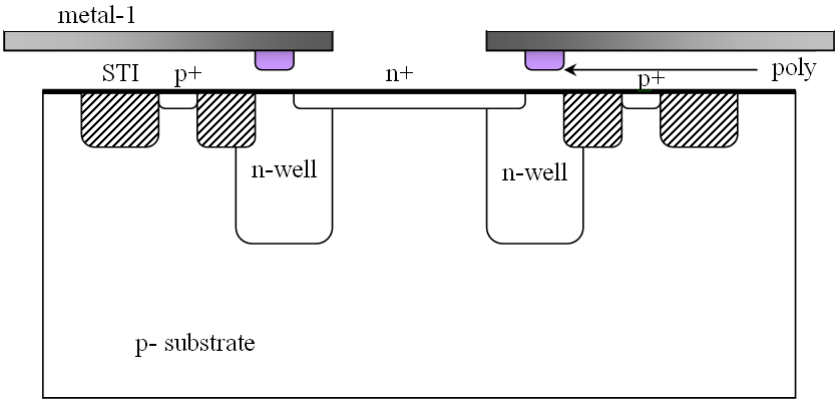

Figure 4.7: *Cross-section of n-well based guard ring SPADs in 90nm technology.*

3) n-well based guard ring structures

The n-well guard ring SPAD uses the connection between the highly doped n

region and the substrate as the active region. Figure 4.7 shows the cross-section of n-well guard ring SPADs which is discussed in more detail.

Numerous structures were implemented, and a systematic search of optimal geometric parameters was conducted. Three different parameters were investigated for each family of SPADs. Figure 4.8 shows the parameters. *(1) active diffusion in the guard ring* for edge breakdown prevention *(2) polysilicon width* to control the separation between the active region and the STI traps, and *(3) STI overlap with the guard ring* to provide maximum separation and edge breakdown prevention. The third parameter can also be used to modulate the surface voltage with poly bias.

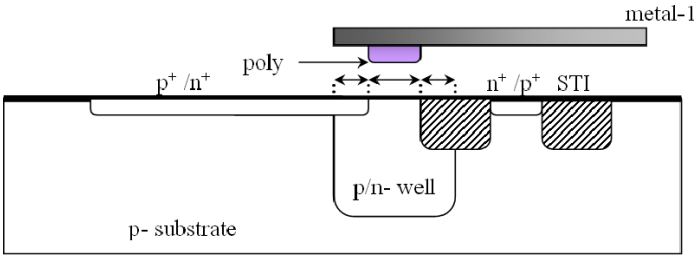

Figure 4.8: *The three main parameters investigated in each family namely: poly width, extention of p/n+ in guard ring and extension of STI in guard ring.*

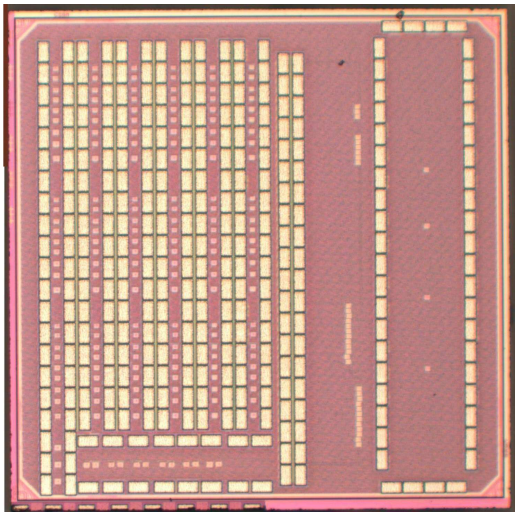

Figure 4.9: *Photomicrograph of the SPAD farm.*

A SPAD farm was implemented to test the most promising SPADs in a 1.8 × 1.8 mm chip. Figure 4.9 shows the photomicrograph of the chip. Each SPAD has three different connections, the anode, the cathode and the polysilicon gate for the guard ring. Most of the SPADs were integrated with ballast resistance as quenching element and an inverter for impedance matching.

Among the three families of structures (p-well, p-substrate, and n-well guard ring) only the latter could be biased in Geiger mode. We concluded that the other two families of structures suffer from the following shortcomings:

Punch-through : the doping profiles and distances between junctions cause the fact that the depletion regions of two superimposed layers touch, thus causing an ohmic resistance that effectively shortcircuits cathode and anode. In the cases where the n-wells could not join each other in the second family of structures (p-substrate guard ring based structures) this effect was observed frequently. Figure 4.10 shows how punch-through can occur between depletion layers of n-wells.

High tunneling : Since the doping profiles at some locations were extremely high in order to decrease the resistance of channels in the MOSFETs, tunneling noise may be elevated in these structures. The existence of Lightly Doped Drain (LDD) between drain and source in MOSFETs also causes some imperfections in the electric field simulation predictions. Due to elevated doping profiles, tunneling dominates thus preventing the control over DCR and causing the diode to operating as a Zener diode.

Premature edge breakdown : the guard ring does not reduce the effective electric field, thus not preventing PEB.

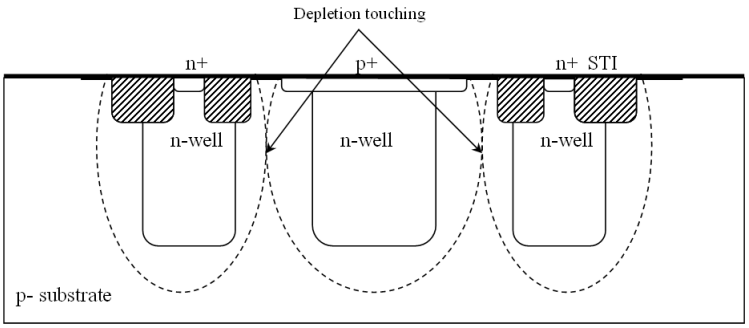

Figure 4.10: *Punch-through between the depletion regions of n-wells, resulting in short circuit.*

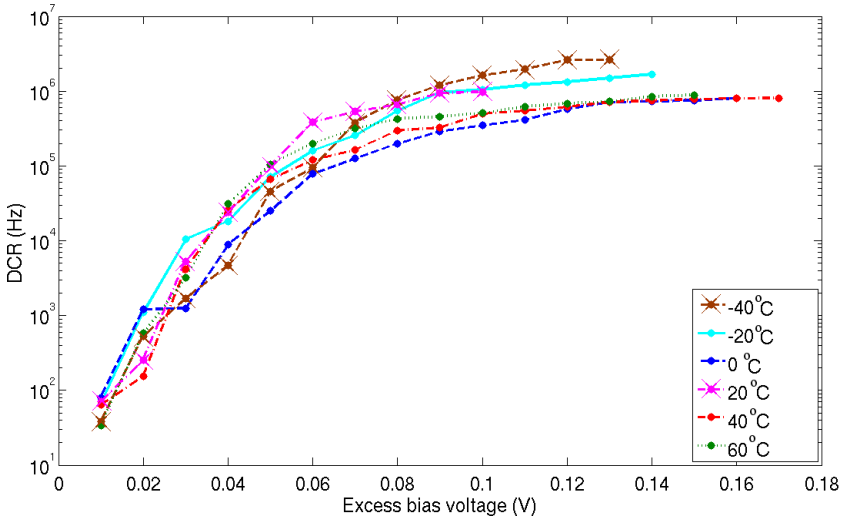

Figure 4.11: *DCR of the first sub-family of SPADs in 90nm technology at different excess bias voltages and different temperatures. (All the SPADs are based on n-well guard rings).*

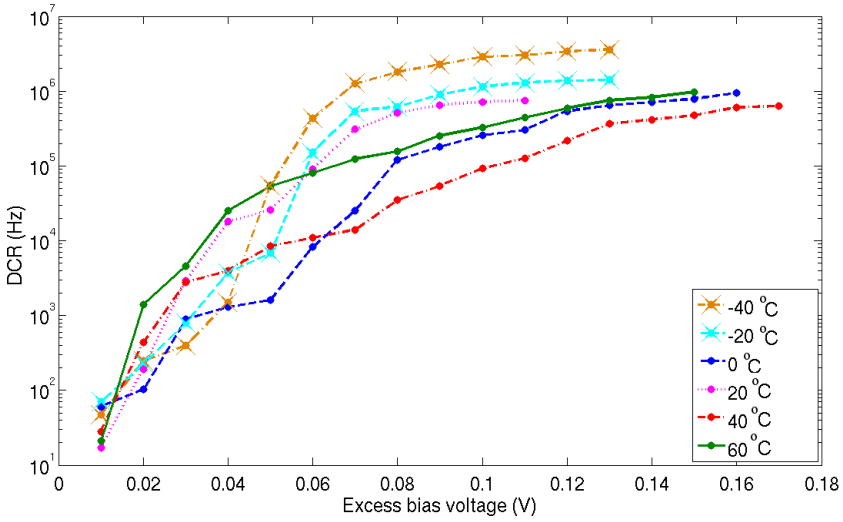

Figure 4.12: *DCR of the second sub-family of SPADs in 90nm technology at different excess bias voltages and different temperatures. (All the SPADs are based on n-well guard rings).*

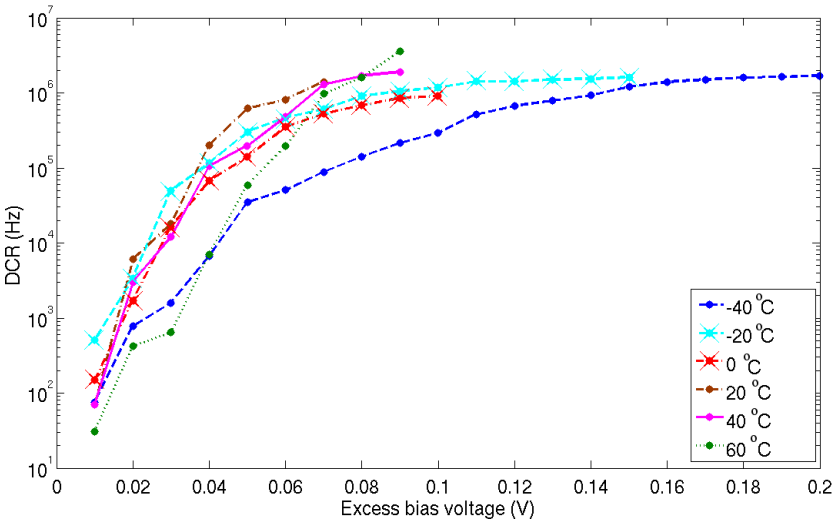

Figure 4.13: *DCR of the third group of SPADs in 90nm technology at different excess bias voltages and different temperatures. (All the SPADs are based on n-well guard rings).*

SPADs of the third family (n-well guard ring based) could operate in Geiger mode. The family of structures consists of four different sub-families. In each family, the guard ring is implemented using different layer contributions. DCR results of three of the four sub-families are presented. Figures 4.11, 4.12, and 4.13 show the DCR of each sub-family of SPADs at different excess bias voltages and different temperatures. At some points the DCR at lower tempertaure is higher than the DCR in the higher temperature which can be due to temperature instability of the testing device.

4.2 SPAD characterization

Figure 4.7 shows the cross-section of the SPAD family being characterized in this section. Figure 4.14 shows the SPAD photomicrograph. The photon sensitive area is in the center of the octagonal shape and other parts are covered with metal to reduce probability of photon absorption in the guard ring area. A diameter of 8 μm is used for SPADs. Circular shaped SPADs have better characteristics, due to sharp edge prevention which reduces hot spots of electric field in corners. However design rules prevented us from using circular geometries.

Most of the free carriers generated by photons in the active region will trigger the avalanche breakdown while free carriers in the guard ring do not gain enough energy to initiate the avalanche. Figure 4.15 shows the electric field distribution

of the SPAD at 0.2 V of excess bias voltage. The combination of different doping layers with different sizes is arranged in such a way to confine the highest electric field to the planar active area. The simulation of the designed SPAD is carried out by importing the doping profiles from the 90nm technology using the Spectra imaging device simulator.

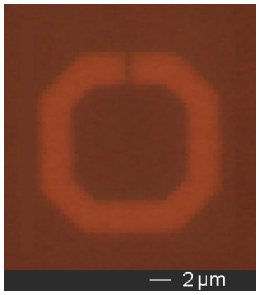

Figure 4.14: *Photomicrograph of an octagonal SPAD. The guard ring is covered with metal to prevent light absorption in the guard ring region.*

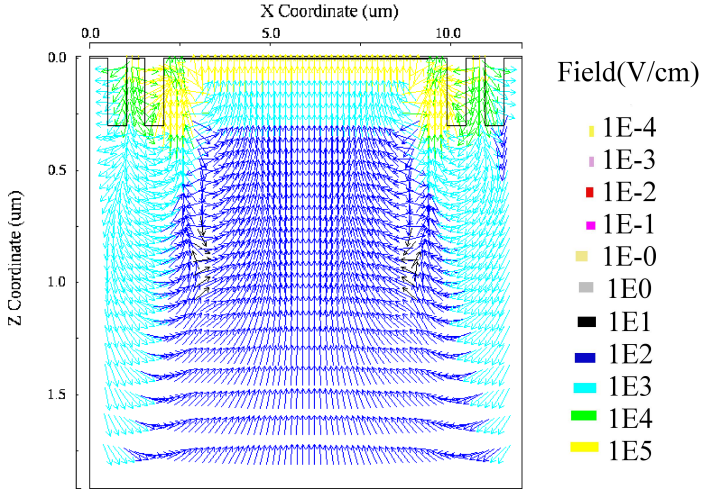

Figure 4.15: *Electric field distribution of the n-well guard ring based SPAD in 90nm technology, biased above breakdown.*

Figure 4.16 shows the strength of the electric field at different depths in the absence of any radiation. The picture shows that the guard ring is mostly effective near the surface of the SPAD, where most of the photon absorption takes place.

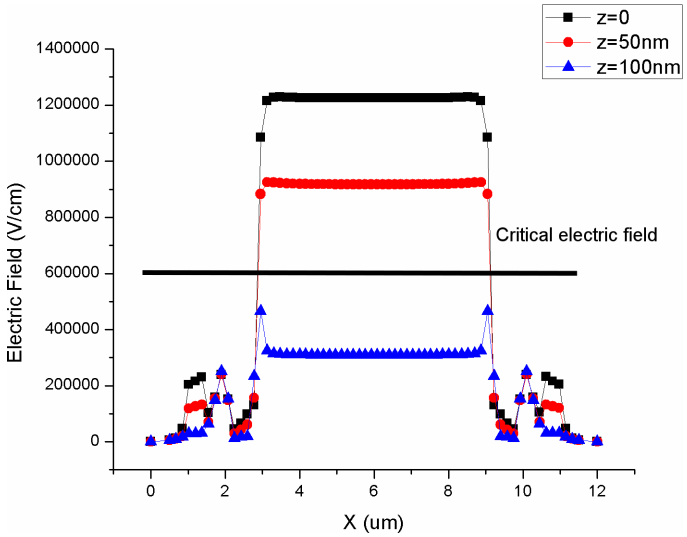

Figure 4.16: *Electric field strength in the cross-section of the SPAD at different depths.*

While the diameter of the anode was kept 8 μm, different combinations of n^+ and n-wells were examined. The thickness of the n-wells is constrained by design and rule check (DRC) from 500 nm to 2 μm. Simulations were conducted using STI overlap with n-well guard ring (100 $\sim$ 500 nm), and Anode in the guard ring (200 $\sim$ 500 nm). Among the simulated and implemented structures, thicker and deeper guard rings show better performance. It should be noted that the deep n-well with low doping concentration played an important role in DCR reduction of the reported SPAD. The reported device has no isolation to separate the potential of the anode from the ground.

With this design indeed there is a need for introducing decoupling between the cathode and the input to the next stage if this is to be used in an array of detectors. The obvious techniques to solve this problem are:

1. A capacitive decoupler or resistive partitor that require more area and possibly introduce noise, especially in time-resolved imaging. Note that high voltage capacitances and resistors need to be used in this case; Figure 4.40 shows the capacitive decoupling from cathode providing the desired isolation.

2. Placing the electronics in a deep n-well with a fully embedded p-well and bias the entire substrate at a high negative voltage. This is a non-standard approach and may interfere with a conventional semi-custom design flow. In addition, a p-well/deep n-well option may not always be available in standard deep-submicron CMOS technologies. If that were the case, NMOS

digital circuitry may be used as a workaround, with consequences in terms of speed and power consumption. Figure 4.17 shows the cross-section of a SPAD built on separate deep n-well and p-well eliminating the coupling issues.

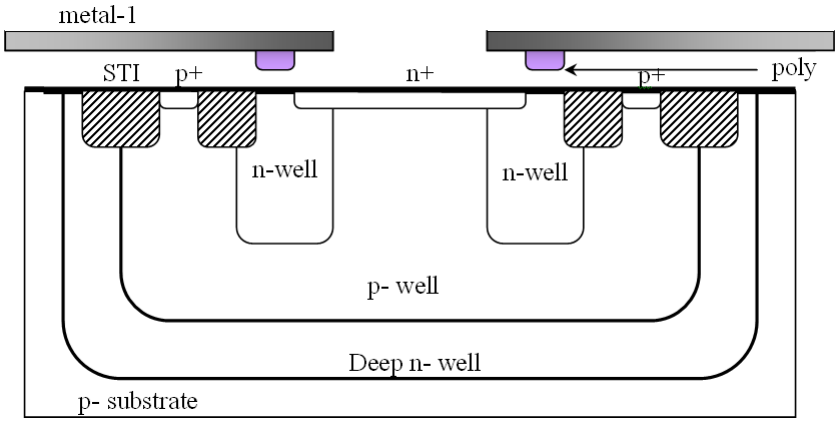

Figure 4.17: *Eliminating coupling between SPAD and substrate using deep n-well and p-well.*

3. Use of SOI process. This solution is a valuable one and could actually offer other advantages in terms of noise isolation and PDP boosting. Figure 4.18 shows how buried oxide layers (BOX) of SOI technologies eliminate the problem of SPAD coupling with substrate.

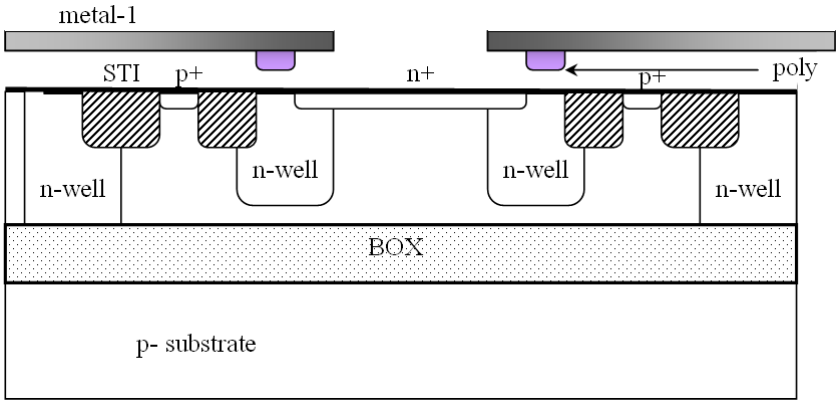

Figure 4.18: *Buried oxide layers for eliminating coupling between SOI and substrate.*

4.2.1 Experimental results

Figure 4.19 shows the I-V characteristics of the SPAD in reverse bias mode of operation. Although in some of the implemented SPADs in this technology series resistance affects the I-V curve, the observed breakdown behavior of the diode shown in this figure is proper for the photon counting applications. The sharp transition between the band-to-band tunneling region in I-V curve and the avalanche caused by impact ionization is sharp enough for the device to work in Geiger mode.

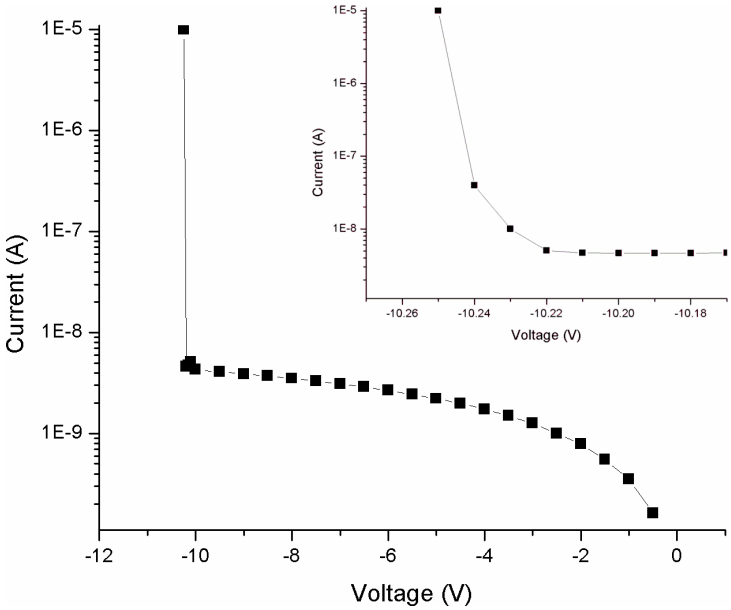

Figure 4.19: *The I-V characteristic of the diode implemented in 90nm technology which shows photon counting capability. The sharp transition between the band-to-band tunneling region and the avalanche region makes the SPAD suitable for the Geiger mode operation.*

The breakdown behavior indicates that tunneling is high in these devices. This is to be expected due to the relatively high doping levels of the substrate. That is why the excess bias has to be kept low to avoid high DCR while PDP still has acceptable levels. The pre-breakdown current is in the nA range which is indeed a large dark current. However, this does not present a problem since it is not related to the DCR mechanisms. It only implies a slightly larger power dissipation especially in regimes of high illumination.

In order to characterize the number of spurious pulses generated by tunneling and SRH processes, DCR is measured. Designing low DCR SPADs is important

for photon-starved applications. However, medium or high range noise SPADs can be used in commercial imaging systems, e.g. 3D vision systems. These systems can sustain relatively high DCRs (in the order of a few tens of kilohertz) due to the higher level of background noise in typical systems [58].

Figure 4.20 shows the DCR of the fabricated SPAD at different temperatures and excess bias voltages. Passive quenching is used in the first prototype of the implemented SPAD.

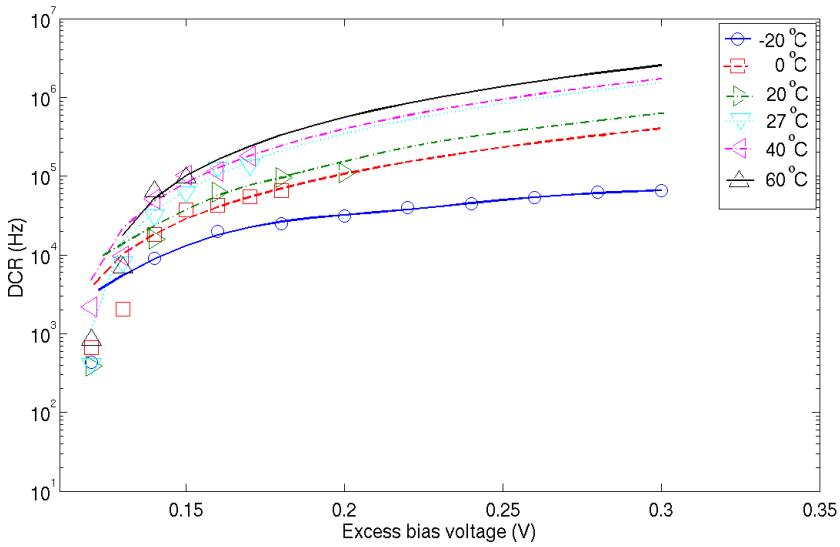

Figure 4.20: *DCR vs. excess bias voltage for different temperatures. The curves were fitted to show the trend of DCR in different biasing conditions.*

Although tunneling effects increase DCR with temperature, the major temperature-dependent DCR contribution is SRH. The SPADs described here exhibit 8.1 kHz of DCR at room temperature with 0.13 Volts of excess bias voltage.

The DCR can be decreased by decreasing the temperature and/or excess bias voltage. Due to relatively large dead time, the SPAD enters saturation relatively early. The DCR in the plot of Figure 4.20 is not shown after saturation is reached. Higher DCR is to be expected due to higher tunneling.

Tunneling and SRH dark counts can be isolated and independently measured by varying the temperature: below a certain temperature, tunneling DCR dominates and above that temperature SRH dominates.

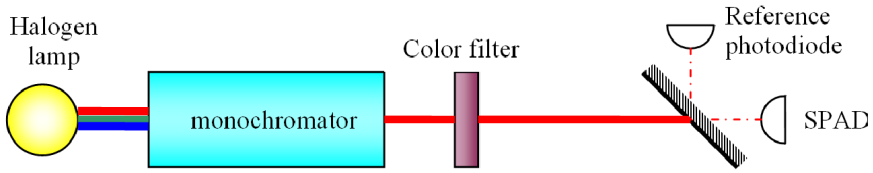

Figure 4.21: *Schematic configuration of PDP measurement setup.*

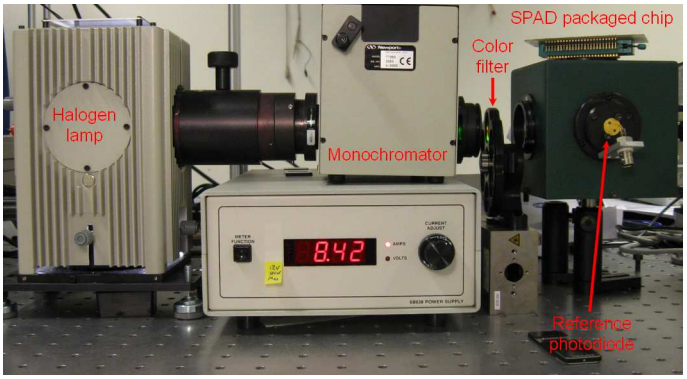

Figure 4.22: *Optical setup for PDP measurement.*

The PDP is measured for the entire spectrum of interest (360-800 nm). Figure 4.23 shows the PDP at room temperature at different bias voltages. The figure shows that the detection probability can be as high as 15 % at 520 nm. The shallow doping profile of the n^+ layer in the 90nm CMOS technology results in higher PDP in the ultraviolet region. The overall lower PDP if compared with other technologies [63] is predictable due to the shallower multiplication region. In addition, a thicker optical stack with worse refraction index matching, causes more attenuation and a higher ripple in the PDP profile.

Figures 4.21, 4.22 show the schematic and real images of the PDP measurement setup. Wide spectrum light is generated by a halogen lamp and is filtered by a monochromator. Extra filters are located after the monochromator, filtering the unexpected spectrum of light. Light is divided in two identical paths, comparing SPAD counts with the reference photon counts.

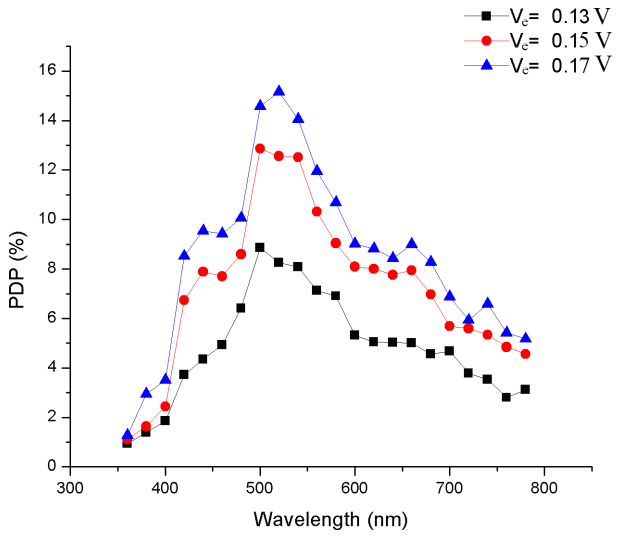

Figure 4.23: *Photon Detection Probability (PDP) of a SPAD implemented in 90nm technology. The PDP was measured with a monochromator and reference photodetector at room temperature.*

Timing jitter is characterized using two fast laser sources with a pulse width of 40 ps and repetition rate of 40 MHz emitting a beam of light with the wavelength of 637 nm and 405 nm, respectively. The time interval between the laser output trigger and the leading edge of the SPAD signal is measured via a high performance oscilloscope. A histogram is constructed from the time difference of the edges of the laser light and sensing signal. Figure 4.24 shows the resulting histogram at 405 nm with an excess bias voltage of 0.13 V.

The FWHM of the time difference histogram was measured to 398ps for 637 nm and 435 ps for 405 nm. The jitter in this technology is higher than older technologies [120]. Since the doping concentration of the p-substrate in this technology is lower than the previous deep n-wells, the depletion region extends deeper into the substrate. As a result, photocarriers generated relatively deep into the substrate may be captured, after diffusing, through the substrate, in the multiplication region, thus increasing the uncertainty of the timing of electron-hole pair generation [120].

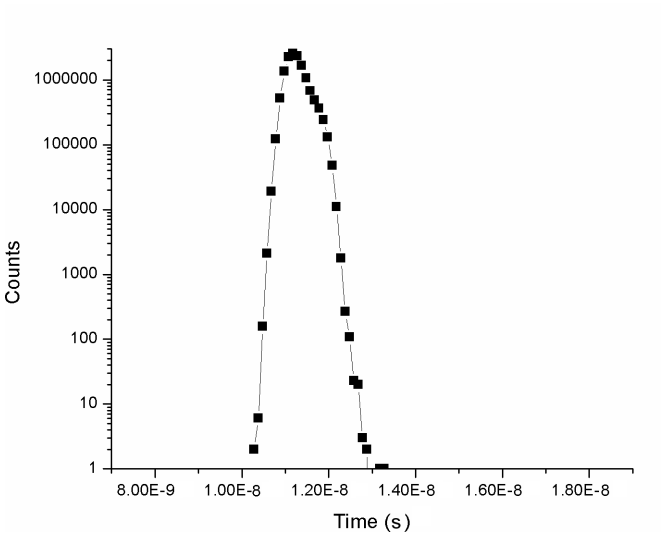

Figure 4.24: *Histogram of the time between the laser pulse and the SPAD receiving digital pulse at 405nm wavelength.*

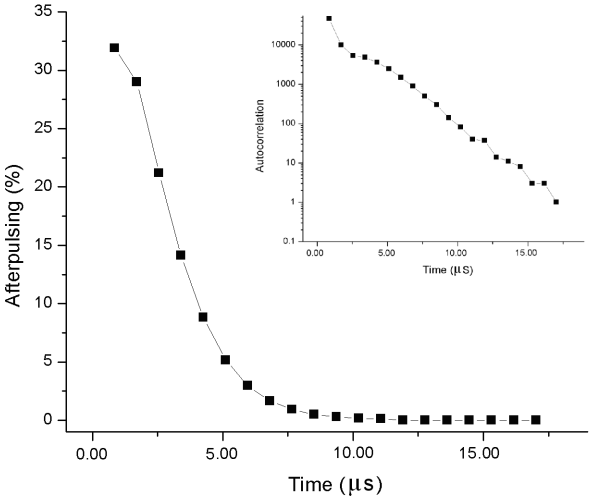

Figure 4.25: *Afterpulse probability as a function of dead time. In the inset the corresponding autocorrelation function.*

Figure 4.25 shows the result of the afterpulse probability measurement as a function of dead time. The afterpulsing measurement was carried out by the discrete autocorrelation measurement of the signal. The main reason for the afterpulsing generation is the traps in the depletion region that add a history effect by initiating the avalanche independent of photon absorption. Since the depletion region in the substrate is thicker, the probability of having traps in the depletion region is higher and the afterpulsing probability increases correspondingly. The dead time of the detector is 1.2 μs. A large parasitic capacitance in parallel to the anode of the device is largely responsible for a long recharge time that, in turn, dominates the dead time. High afterpulsing probability is due to a relatively large charge of approximately 17.85 pC involved in each avalanche. Large numbers of carriers are undesirable in SPADs as they result in higher probability of trapped carriers that are prone to the generation of secondary spurious avalanches.

A photoluminescence test was conducted to verify the effectiveness of the guard ring. When edge breakdown occurs in SPADs the avalanche generates photoluminescence as by-product to impact ionization in the guard ring. In contrast, in the SPADs where the premature edge breakdown is prevented, the avalanche occurs in the active region of the SPAD. Figure 4.26 shows the 90nm technology SPAD photoluminescence. The left panel shows operation in Geiger mode, while the right panel shows operation in thermal equilibrium. In both cases no photoluminescence outside the active area can be obsreved, thus implying that no or negligible avalanche activity is occurring in the outer rim of the detector and the premature edge breakdown prevention is effective.

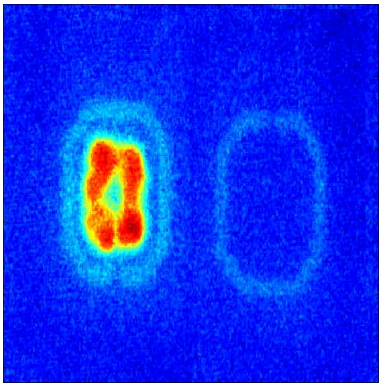

Figure 4.26: *Photoluminescence of a SPAD operating in Geiger and sub-Geiger mode.*

4.2.2 Conclusion

We reported on the first implementation of a single-photon avalanche diode in 90nm CMOS technology. The detector features an octagonal multiplication region

and a guard ring to prevent premature edge breakdown using a standard mask set, exclusively.

Performance	Min	Typ	Max	Units	Comments
SPAD diameter		8		μm	
DCR		8.1		kHz	V_e=0.13V, T=293K
Timing jitter		398		ps	FWHM at 637nm wavelength
Timing jitter		435		ps	FWHM at 405nm wavelength
PDP			9	%	V_e=0.13V, T=293K
Afterpulse probability		32		%	At nominal dead time
Breakdown voltage	10.28	10.4	10.43	V	
Wavelength range	360		800	nm	

Table 4.1: *Summary of experimental results. All measurements were conducted at room temperature.*

The proposed structure emerged from a systematic study aimed at miniaturization, while optimizing the overall performance. The guard ring design is the result of an extensive modeling effort aimed at constraining the multiplication region within a well-defined area where the electric field exceeds the critical strength for impact ionization. The device exhibits a dark count rate of 8.1 kHz, a maximum photon detection probability of 14 % at maximum excess bias. At 0.13 V of excess bias, a PDP of 9 % and a jitter of 398ps at a wavelength of 637nm, were measured at room temperature. An afterpulsing probability of 32 % was measured at the nominal dead time.

The performance of the SPAD implemented in the 90nm standard CMOS technology is summarized in Table 4.1. While the high DCR is expected because of the tunneling and high doping profiles, the PDP is lower due to high, unoptimized optical stack and thinner multiplication regions.

4.3 Outlook: 65nm CMOS process

4.3.1 Technology migration

The methodology used in the design of SPADs in 65nm technology was the same as in the 90nm technology, but with more freedom in choosing distances. Different wells for 1.5,and 2.5 V devices are available in this technology. Interconnect metal spacing and thicknesses were reduced comparing with 90nm technology, helping to thin optical stacks. The scaling table shows the technology differences between 90nm and 65nm technologies.

Parameter	Scaling factor
metal to metal distance	0.96 S
guard ring	0.3 S
SPAD diameter	1
Transistor channel length	S
Transistor voltages	S, 1.5 S
Number of passivation layers	2

Table 4.2: *Scaling table comparison between 90nm and 65nm technology S=90/65.*

4.3.2 New implemented SPAD design and simulations

In addition to the main three families of structures described in 90nm technology (p-well, p-substrate, n-well guard ring based structures), five new families are investigated for this technology. The remainder of this chapter describes these five families whereas simulation results are presented.

Deep n-well guard ring structures

Figure 4.27 shows a deep n-well (DNW) guard ring structure. When a complete deep n-well is formed, high tunneling and PEB are expected. Thus, we eliminate the DNW below the multiplication region expecting the DNW to act as a guard ring by reversing the surrounding region onto a lightly n-doped region. The vertical effects of deep n-wells are reduced due to fewer annealing steps in deep-submicron technologies.

Figure 4.29 shows the electric field strength obtained by Spectra simulations at 11V of bias voltage. Simulation results show the maximum electric field in the active region which is appropriate for PEB prevention.

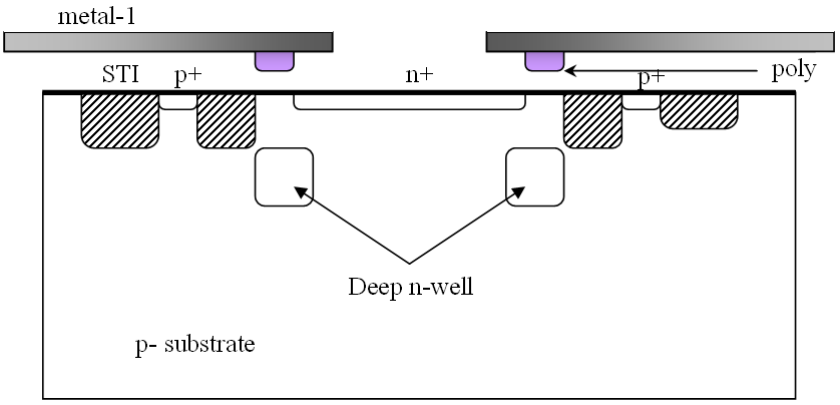

Figure 4.27: *Cross-section of deep n-well guard ring SPAD integrated in 65nm technology.*

Low voltage p-well guard ring structures

Shallower p-wells, in comparison with thick wells used in previuos technologies, are used to evaluate the guard ring. The difference between this type of structures and similar structures in 90nm technology is the separation between guard ring and deep n-wells. This separation is performed by a special layer present in 65nm technology. Lower PDP is expected because of thinner depletion region but this SPAD can work independently from the substrate voltage. Figure 4.28, shows the low voltage p-well guard ring SPAD structure designed in 65nm technology.

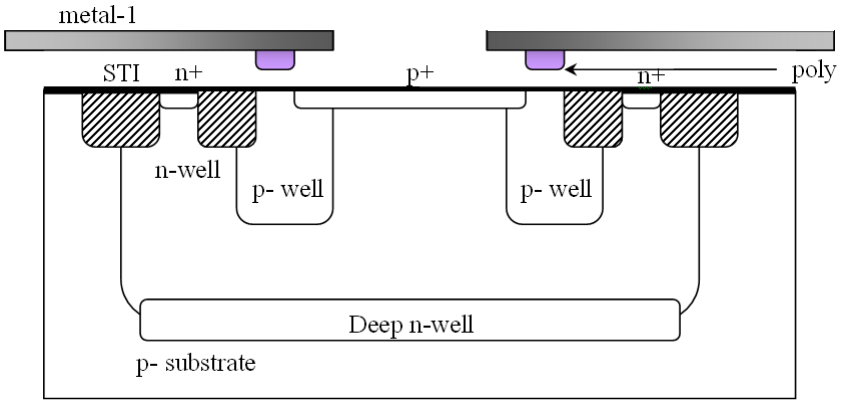

Figure 4.28: *Cross-section of low voltage p-well guard ring SPADs.*

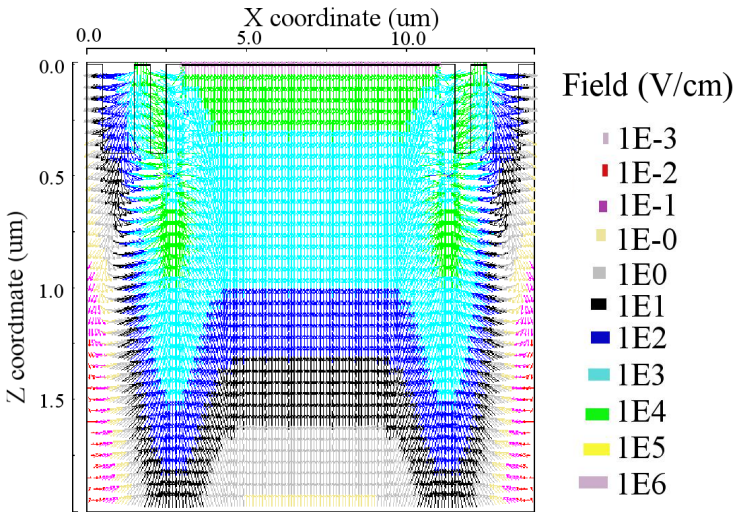

Figure 4.29: *Electric field simulation result of deep n-well guard ring based SPAD at room temperature. The critical electric field for impact ionization is* $6 \times 10^5 V/cm$ *in silicon.*

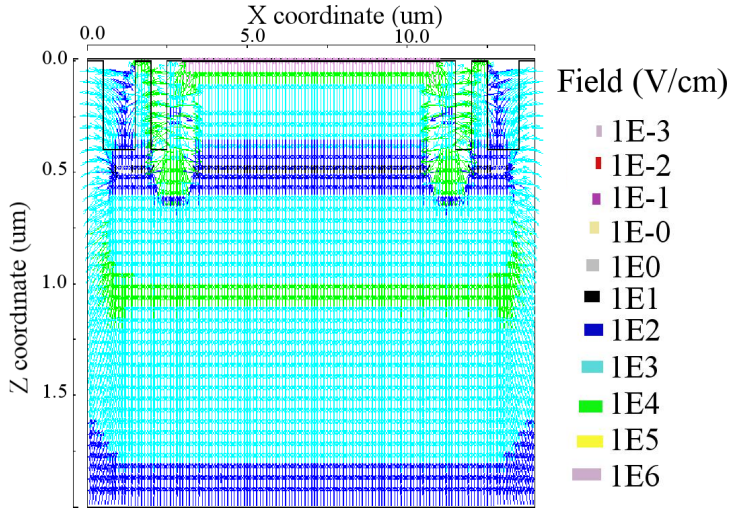

Figure 4.30: *Electric field simulation result of low voltage p-well guard ring SPAD at room temperature.*

Electric field distribution is shown in figure 4.30 using Spectra at 11V of bias voltage. While the narrow maximum electric field is located in the active region some high electric field spots are inside the guard rings. Due to shallow p-wells PEB is predicted, not providing enough isolation with n^+.

STI-bound P-substrate guard ring based structures

Shallower n-wells with minimum separation are used in these structures. Different separate n-wells are used reaching the minimum distance between n-wells connecting the two separate n-wells. Minimum separation results in a guard ring region doped as intrinsic substrate. Figure 4.31, shows the STI-bound p-substrate guard ring SPAD structure using a special layer in this technology.

Figure 4.33 shows the electric field strength resulting from by Spectra simulations at 11 V of bias voltage. A uniform high electric field region can be seen in the guard rings which is highly dependent. Some high electric spots are observed close to the guard ring causing band-to-band tunneling (BTBT). If hot spots are connected, band-to band tunneling should be considered, otherwise local points high electric field do not inject BTBT noise.

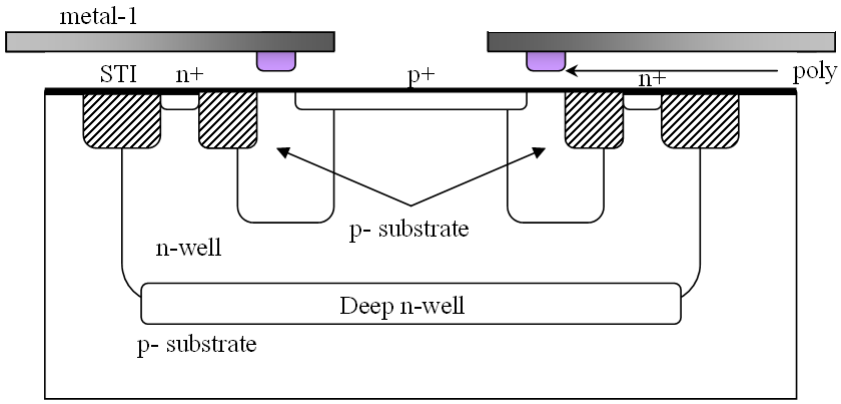

Figure 4.31: *Cross-section of STI-bound p-substrate guard ring SPAD.*

LDD guard ring based structures

Shallow LDD layers are used in these structures instead of separate guard rings. The SPAD is placed on top of deep n-well and surrounded by n-wells. The STI regions are placed at different distances from the active region which is predicted to have more impact on the SRH noise.

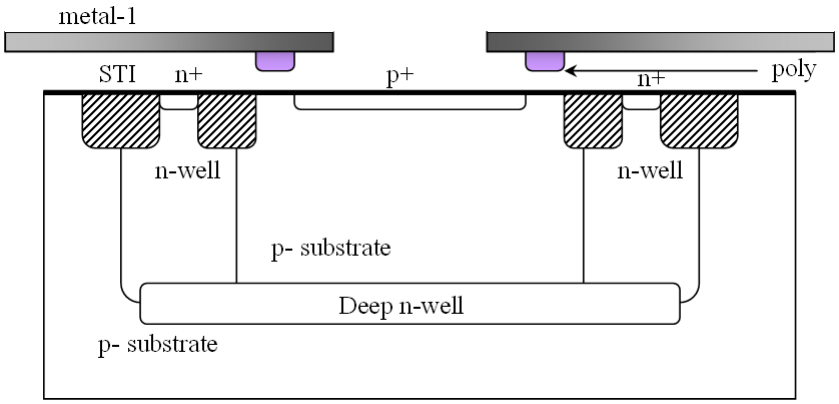

Figure 4.32: *Cross-section of LDD guard ring based structures integrated in 65nm CMOS technology.*

Figure 4.32, shows the cross-section of LDD guard ring SPAD structure. The lowest doping concentration is expected for the substrate making it effective for light absorption and tunneling prevention.

Figure 4.34 shows Spectra simulation results at 11 V of bias voltage. The connection between n-wells and deep n-well which is indicated in figure 4.32 also shows a high electric field region which can be troublesome for device in breakdown.

New structures designed based on double epitaxial technologies

The guard ring used in [22] shows effectiveness in PEB prevention. This device is based on an in-house process of double epitaxial silicon substrate. The devices exploit the difference between the doping depth of highly doped p-regions and highly doped n-regions. The deep p^+ region in the center of the device is called enrichment region and confines the electric field to the center of the device and prevents high electric fields on the edges between n^+, and the substrate.

Figures 4.35, shows the proposed cross-section. This structure is similar to [23] without the use of epitaxy layers. The p^+ enrichment layer is placed with different radiuses to achieve maximum isolation from the edges.

Figure 4.37 shows the electric field distribution simulated using Spectra at 11 V of bias voltage. The electric field shows the highest values in the active region, while some high electric field spots are distinguishable on the edges of n^+. This device is expected to suffer from coupling with the substrate, since the double epitaxy was used to prevent this to occur.

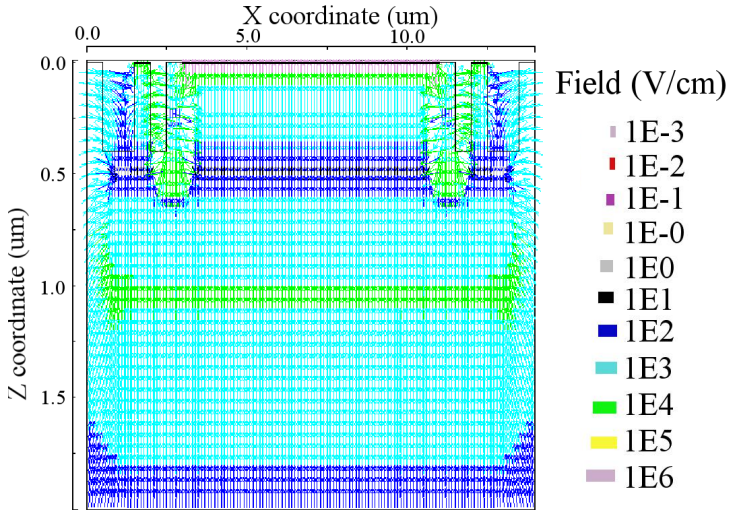

Figure 4.33: *Electric field simulation result of the STI-bound p-substrate guard ring SPAD at room temperature.*

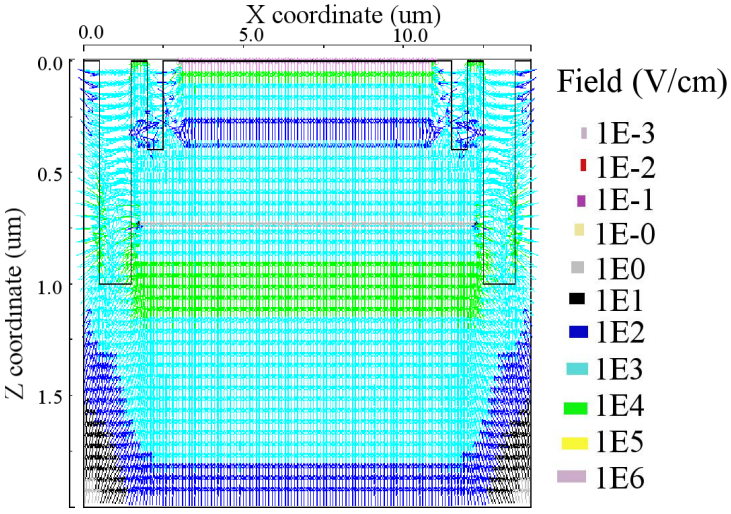

Figure 4.34: *Electric field simulation of LDD guard ring based structure at room temperature.*

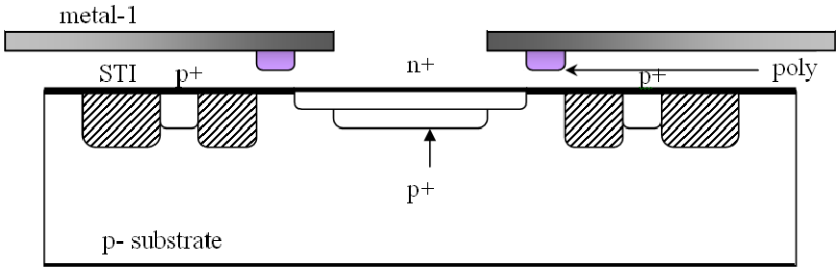

Figure 4.35: *Cross-section of the proposed structure based on an enrichment p^+ region.*

Varying the parameters in SPAD families

Five different parameters are varied to find the most efficient guard ring in the main families of SPADs. Figure 4.36 shows the three parameters on top of the silicon substrate. *(1) active diffusion in the guard ring* which is changed mainly for the edge breakdown prevention *(2) poly width* which implies the separation between the active region and the STI traps, and *(3) STI overlap with the guard ring* which is mainly used to give the maximum separation from active region and edge breakdown prevention *(4) STI overlap with p/n-wells* which gives the separation between the STI and the guard ring *(5)Deep n-well width* which defines the guard ring width and separation for the deep-n-well guard rings. Figure 4.38 shows the photomicrograph of the octagonal SPAD designed and implemented in 65nm standard CMOS technology.

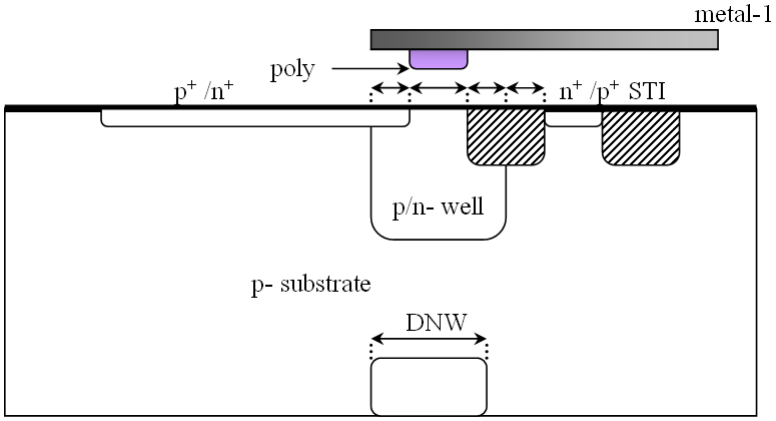

Figure 4.36: *The Five main parameters varied in each family of SPAD: poly width, extention of p/n+ in guard ring and extension of STI in the guard ring, STI overlap with p/n-wells, and the Deep n-well width.*

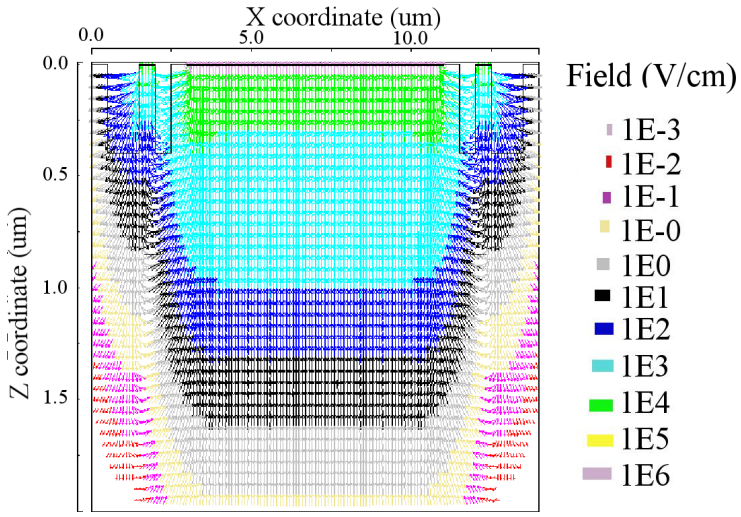

Figure 4.37: *Electric field simulation of the SPAD structure at room temperature.*

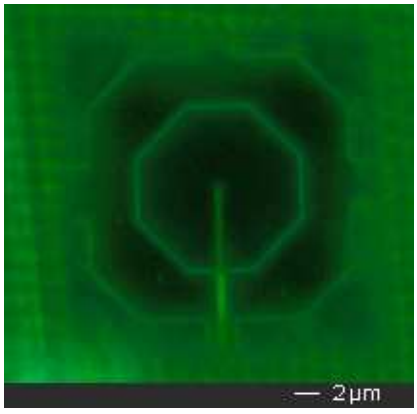

Figure 4.38: *Photomicrograph of an octagonal SPAD implemented in 65nm technology. The guard ring is covered with metal to prevent light absorption in the guard ring region.*

4.3.3 SPAD biasing

SPADs are biased in two ways, applying negative potential on the anode, or positive potential on the anode. Minimizing the detector dead time and charge flow quantity during avalanche should be considered in any biasing. For SPADs

that are implemented within deep nwells two passive quenching circuits can be considered as shown in figure 4.39.

Figure 4.39a shows the complementary voltages of high negative breakdown and very low positive excess bias voltage, where the passive quenching is performed by a PMOS connected to the cathode. Figure 4.39 b shows a large bias voltage applied to the cathode while a NMOS quenches the avalanche. In configuration a the cathode is the moving node and additional capacitance of n-wellL to p-substrate parasitic diode must be charged and discharged during the detector operating cycle. This adds to the charge volume flow through the detector, and increases the probability of charge trapping and afterpulsing probability. In the case of b structure, the charging of SPAD junction capacitance should only be considered.

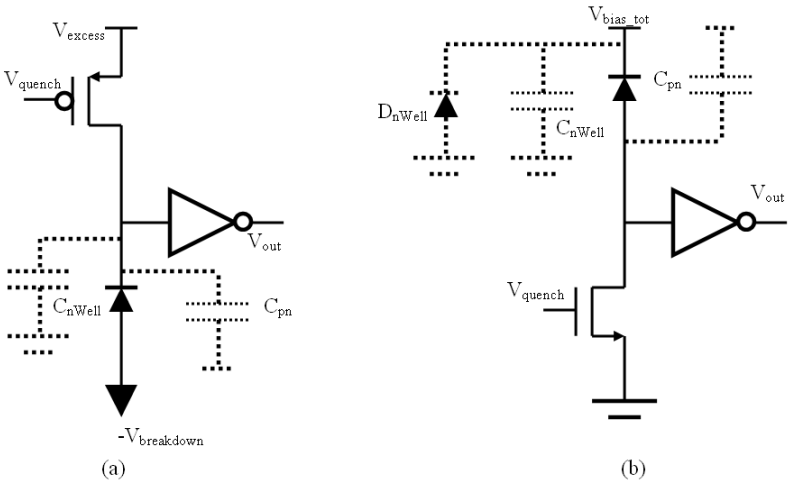

Figure 4.39: *SPAD biasing options (a) cathode (b) anode, as the reference for output signal.*

4.3.4 Implemented circuits

Quenching is needed to limit the current passing through the SPAD. Different resistors located outside the chip are used to find the appropriate quenching resistance. One of the simple ways of quenching is using a resistor to limit the current passing through SPAD after avalanche. The resistors can be used as an external element or be implemented by a PMOS transistor. This resistance shown in Figures 4.40, and 4.41 is modulated with v_b.

The most promising SPAD which had worked in 90nm technology is integrated with inverter in the 65nm technology. The functional SPAD being shown in Figure 4.7 is directly connected to the substrate. A special circuit is needed to prevent

interaction between the SPAD and the logic parts: a capacitor has been inserted between the high voltage SPAD and the gate of the transistor of the next stage. The circuit designed and implemented is shown in figure 4.40.

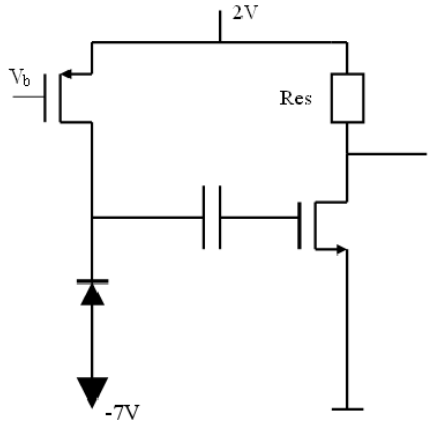

Figure 4.40: *Implemented circuit for digitalizing the SPAD output.*

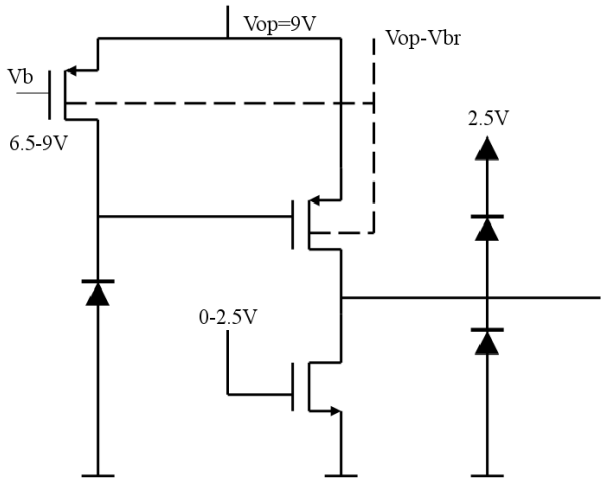

Figure 4.41: *Second circuit for digitalizing the SPAD output.*

Other types of circuits can also be used for n-well guard ring SPAD as shown

in figure 4.41. This configuration of digitalizing SPAD output is problematic due to direct connectivity of SPAD output and a transistor gate, which makes it vulnerable to high voltages.

Figure 4.42 shows simulation results of next stage transistor gate and SPAD output. The top graph shows the SPAD output during photon detection and the bottom shows the input gate of the transistor both simulated at the room temperature.

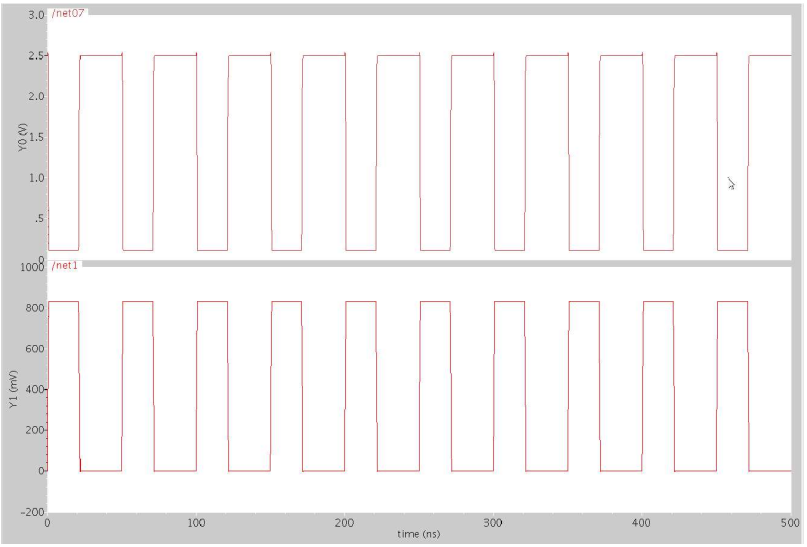

Figure 4.42: *The transient time simulation of the capacitance part voltage convertor.*

4.4 Conclusion

This chapter describes the implementation and characterization of SPADs in a 90nm standard CMOS technology. By implementing different SPAD structures, we could study the geometric trade-offs involved in the design of deep-submicron SPADs. The implemented structures were also simulated and characterized. Moreover, the simulation considerations of SPADs designed in 65nm CMOS technology were described.

Chapter 5

Summary and Future Work

5.1 Main achievements

Quantum parasitic effects and miniaturization of Single Photon Avalanche Diodes in deep-submicron technologies have been studied in this thesis in detail. Tunneling noise and Random Telegraph Signal (RTS) noise have been the main two parasitic effects addressed comprehensively.

While the fundamental equations for tunneling count rate in SPADs have been presented, the tunneling study has been performed by comparing two different topologies of SPADs on the same technology.

The tunneling evaluation has been performed by implementing a new n-tub guard ring SPAD in 0.35 μm standard CMOS technology. The NTGR SPAD has been characterized theoretically by simulations, and experiments. The tunneling noise effects have been verified in the I-V characteristics and Dark Count Rate responses of the new NTGR SPAD.

While the NTGR SPAD has shown 3 % improvement in the maximum Photon Detection Probability (PDP), the timing jitter Full Width Half Maximum (FWHM) has been increased from 80 ps in p-tub guard ring SPAD to 152 ps in the NTGR SPAD. The breakdown voltage has been decreased 4.1 V due to using different doping profiles for NTGR SPAD implementation. Avalanche photoemission of the active area of NTGR SPAD demonstrated the ability for single photon detection in comparison with the other diodes, which show photoemssion in guard rings as a sign of Premature Edge Breakdown (PEB).

The second parasitic effect which has been studied in this thesis was the RTS behavior of dark count rate (DCR); this is characterized as a bistability and multistability of DCR. RTS behavior of DCR has been observed in a SPAD fabricated in 0.8 μm CMOS technology and in four proton-irradiated SPADs designed and fabricated in 0.35 μm CMOS technology. To the best of our knowledge, this

was the first time RTS behavior of DCR has been reported in SPADs in any CMOS technology.

RTS characteristics have been evaluated experimentally and verified theoretically with respect to the bias and temperature. The RTS hypothesis of the fluctuations has been verified by measurements that have been in excellent agreement with the theory. The RTS behavior is expected to affect large array of SPAD detectors being used for 3D imaging applications.

Demonstrating the functionality of SPADs in deep-submicron technologies has been the second main challenge of this thesis. We have demonstrated that single-photon detectors can be fabricated in commercial deep-submicron CMOS processes. Miniaturization has been explored with the first SPAD designed and successfully tested in technologies smaller than 130nm.

The proposed structures, implemented in 90nm standard CMOS technology, emerged from a systematic study aimed at miniaturization, while optimizing overall performance. The guard ring design has been the result of an extensive modeling effort aimed at constraining the multiplication region within a well-defined area where the electric field exceeds the critical value for impact ionization.

By implementing different SPAD structures, we have studied the geometric trade-offs involved in the design of deep-submicron SPADs. Numerous SPADs with different arrangements of doping layers and different guard ring sizes have been implemented. Among them, as many as 45 structures were functional with a range of well-defined, reproducible breakdown voltages. The implemented structures have also been simulated and characterized.

The detectors feature an octagonal multiplication region and a guard ring to prevent PEB using a standard mask set, exclusively. The proposed structure emerged from a systematic study aimed at miniaturization, while optimizing the overall performance.

The devices exhibit a DCR of 8.1 kHz, a maximum PDP of 14 % at maximum excess bias. At 0.13 V of excess bias, a PDP of 9 % and a jitter of 398ps at a wavelength of 637 nm, were measured at room temperature. An afterpulsing probability of 32 % was measured at the nominal dead time.

The main weak point of SPADs which have been demonstrated in this thesis in 90nm technology was the lack of isolation with the substrate, while different solutions have been proposed confining the NTGR SPAD, not to communicate with the substrate.

Simulation based analysis of SPADs designed in 65nm CMOS technology has been described in the last section. The design of SPADs in 65nm technology was based on the characterization results of functional SPADs in 90nm technology.

While implementation of SPADs in imaging CMOS technologies can lead to better performance due to the special layers, one of the main achievements of this thesis has been the demonstration of SPAD in a standard CMOS technology without employment of imaging doping layers. Because of the isolation issues described above, novel circuits have been implemented for the SPAD quenching and pulse detection.

5.2 Outlook and future work

While the preliminary results show a successful implementation of SPADs in 65nm standard CMOS technology for the first time, different challenges should be addressed in the future for SPAD implementations in deep-submicron technologies namely:

5.2.1 PDP improvement

One of the main areas of concern for deep-submicron SPADs is the PDP response. Low PDP may limit the field of applicability. The doping profile of the chosen CMOS process prevents a high PDP in the red/IR wavelength range. Special efforts need to concentrate upon the increase of PDP. Different solutions can be applied to increase the PDP, such as modifying the optical stack, applying post processing steps or increasing the multiplication depths, and changing one or more process parameters at a particular layer.

Widening the multiplication region is achievable by introducing specific doping layers or applying post processing doping profiles for doping reduction and compensation. Special effort should be employed extending the multiplication region in order to make SPAD sensible to Infra Red and near IR wavelengths.

5.2.2 Tunneling and DCR reduction

Tunneling is the main noise source in deep-submicron SPADs. The tunneling noise can be reduced by using p-well guard rings or by making the active area smaller. Increasing the active area of the diodes results in higher DCR due to the introduction of more defects and traps.

5.2.3 Jitter reduction

Special effort should be devoted to confining the depletion region and decreasing tunneling noise in order to reduce jitter performance. These efforts may be somewhat in conflict to the increase of PDP, and thus the proper trade-offs should be explored. Special studies should be conducted to analyze the roots of tunneling noise surge in NTGR SPADs.

5.2.4 Dead time and afterpulsing reduction

The dead time and afterpulsing of SPADs implemented in 90nm CMOS technology were high due to lack of integration of standard inverters after at SPAD output thus resulting in high capacitive loads and the floating node of the SPAD. Capacitive reduction at this node should be implemented to decrease the carriers

involved in each avalanche. The dead time can be further reduced by an optimization of the value of the quenching resistor. Active quenching methods can be employed to reduce the afterpulsing probability and the dead time.

5.2.5 RTS noise reduction

RTS noise was observed in two different technologies. The trend in technological development is the reduction of the defects and traps thus decreasing SRH noise as well. Special design and biasing methods should be employed to suppress RTS effects in SPADs, as this could be beneficial in future, more advanced deep-submicron CMOS technologies and in SPADs that must operate in particularly hostile environments, where high energy radiation and cold particles are present.

5.2.6 SPAD isolation

The NTGR SPAD proposed in thesis for 0.35 μm, 90nm and 65nm technologies is not isolated from the substrate. Different methods were proposed for isolating the NTGR SPAD from the substrate. All these methods should be evaluated and further post processing steps should be employed to remove this main weakness of NTGR SPADs.

5.2.7 Lifetime enhancement

Experimental results obtained by testing the SPAD farm in 90nm and 65nm technologies, show lower lifetime of SPADs in deep-submicron technologies. Smart biasing of the polysilicon layers on top of the SPAD can increase the lifetime of the SPAD. Furthermore, additional studies should be performed increasing the SPAD lifetime in deep-submicron technologies.

5.2.8 Employing new standard technologies

The investigation of other industrial CMOS processes is one possible way to fabricate larger SPADs with enhanced detection probability in the red/ IR spectral range. Different new technologies should be studied such as back-side illumination, SOI, and SiGe technologies in order to improve the SPAD performance.

Bibliography

[1] Hamamatsu, Japan, *Photomultiplier tubes: basics and applications*, 2007.

[2] T. Ohnuki, X. Michalet, A. Tripathi, S. Weiss, and K. Arisaka, "Development of an ultrafast single photon counting imager for single molecule imaging," in *Ultrasensitive and Single-Molecule Detection Technologies*, (Boston), 2006.

[3] H. Kume, K. Koyama, K. Nakatsugawa, S. Suzuki, and D. Fatlowitz, "Ultrafast microchannel plate photomultipliers," *Applied optics*, vol. 27, pp. 1170–1178, 1988.

[4] X. Michalet, O. H. W. Siegmund, J. V. Vallerga, P. Jelinsky, J. E. Millaud, and S. Weiss, "Ultrafast microchannel plate photomultipliers," *Nuclear Instruments Methods in Physics Research Section A- Accelerators Spectrometers Detectors and Associated Equipment*, vol. 567, pp. 133–136, 2006.

[5] H. Finkelstein, "Shallow-trench-isolation bounded single-photon avalanche diodes in commercial deep submicron CMOS technologies." PhD Thesis, University of California, San Diego, 2007.

[6] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*. Hoboken, NJ: Wiley, 2006.

[7] K. Nishida, K. Taguchi, and Y. Matsumoto, "Ingaasp heterostructure avalanche photodiodes with high avalanche gain," *Applied Physics Letters*, vol. 35, pp. 251–252, 1979.

[8] H. W. Li, B. E. Kardynal, P. See1, A. J. Shields, P. Simmonds, H. E. Beere, and D. A. Ritchie, "Quantum dot resonant tunneling diode for telecommunication wavelength single photon detection," *Applied Physics letters*, vol. 91, no. 7, pp. 1–3, 2007.

[9] B. E. Kardynal, S. S. Hees, A. J. Shields, C. Nicoll, I. Farrer, and D. A. Ritchie, "Photon number resolving detector based on a quantum dot field effect transistor," *Applied Physics letters*, vol. 90, no. 181114, pp. 1–3, 2007.

[10] M. A. Rowe, E. J. Gansen, M. Greene, R. H. Hadfield, T. E. Harvey, M. Y. Su, S. W. Nam, and R. P. Mirin, "Single-photon detection using a quantum dot optically gated field-effect transistor with high internal quantum efficiency," *Applied Physics Letters*, vol. 89, no. 253503, pp. 1–3, 2006.

[11] J. Kitaygorsky, I. Komissarov, A. Jukna, O. Minaeva, N. Kaurova, A. Divochiy, A. Korneev, M. Tarkhov, B. Voronov, I. Milostnaya, G. Goltsman, and R. Sobolewski, "Dark counts in nanostructured NbN superconducting single photon detectors and bridges," *IEEE Trans. Applied Superconductivity*, vol. 17, pp. 275–278, 2007.

[12] G. Goltsman, O. Minaeva, A. Korneev, M. Tarkhov, I. Rubtsova, A. Divochiy, I. Milostnaya, G. Chulkova, N. Kaurova, B. Voronov, D. Pan, J. Kitaygorsky, A. Cross, A. Pearlman, I. Komissarov, W. Slysz, M. Wegrzecki, P. Grabiec, and R. Sobolewski, "Middle-infrared to visible-light ultrafast superconducting single-photon detectors," *IEEE Trans. Applied Superconductivity*, vol. 17, pp. 246–251, 2007.

[13] A. J. Kerman, E. A. Daulerand, B. S. Robinson, R. Barron, D. O. Caplan, M. L. Stevens, J. J. Carney, S. A. Hamilton, W. E. Keicher, J. K. W. Yang, K. Rosfjord, V. Anant, and K. Bergren, "Superconducting nanowire photon- counting detectors for optical communications," *Lincoln Laboratory Journal*, vol. 16, pp. 217–224, 2006.

[14] E. Charbon and S. Donati, "SPAD sensors come of age," *Optics and photonics news (OPN)*, vol. 21, pp. 35–41, 2010.

[15] F. Zappa, S. Tisa, S. Cova, P. Maccagnani, D. B. Calia, G. Bonanno, M. Belluso, R. Saletti, and R. Roncella, "Pushing technologies: single-photon avalanche diode arrays," in *Advancement in adaptive optics*, 2004.

[16] H. Dautet, P. Deschamps, B. Dion, A. D. Macgregor, D. Macsween, R. J. Mcintyre, C. Trottier, and P. P. Webb, "Photon-counting techniques with silicon avalanche photodiodes," *Applied optics*, vol. 32, pp. 3894–3900, 1993.

[17] B. F. Aull, A. H. Loomis, D. J. Young, R. M. Heinrichs, B. J. Felton, P. J. Daniels, and D. J. Landers, "Geiger-mode avalanche photodiodes for three-dimensional imaging," *Lincoln Laboratory Journal*, vol. 13, pp. 335–350, 2002.

[18] B. Aull, "3D imaging with Geiger-mode avalanche photodiodes," *Optics and Photonics News*, vol. 16, pp. 42–46, 2005.

[19] B. F. Aull, A. H. Loomis, D. J. Young, A. Stern, B. J. Felton, P. J. Daniels, D. J. Landers, L. Retherford, D. D. Rathman, R. M. Heinrichs, R. M. Marino, D. G. Fouche, M. A. Albota, R. E. Hatch, G. S. Rowe, D. G. Kocher, J. G. Mooney, M. E. O'Brien, B. E. Player, B. C. Willard, Z. L. Liau, and J. J. Zayhowski, "Three-dimensional imaging with arrays of Geiger-mode avalanche photodiodes," in *Proceedings of the SPIE*, (US), 2004.

[20] A. Goetzberger, R. M. Scarlett, R. H. Haitz, and B. Mcdonald, "Avalanche effects in silicon p-n junctions .II. structurally perfect junctions," *Journal of Applied Physics*, vol. 34, pp. 1591–1600, 1963.

[21] S. Cova, M. Ghioni, A. Lotito, I. Rech, and F. Zappa, "Evolution and prospects for single-photon avalanche diodes and quenching circuits," *Journal of Modern Optics*, vol. 51, p. 2004, 1267-1288.

[22] A. Lacaita, M. Ghioni, and S. Cova, "Double epitaxy improves single-photon avalanche-diode performance," *Electronics letters*, vol. 25, p. 1989, 841-843.

[23] I. Rech, A. Ingaragiola, R. Spinelli, I. Labanca, S. Marangoni, M. Ghioni, and S. Cova, "A new approach to optical crosstalk modeling in single-photon avalanche diodes," *IEEE Photonics Technology Letters*, vol. 20, no. 5, pp. 330–332, 2008.

[24] A. Spinelli, M. A. Ghioni, S. D. Cova, and L. M. Davis, "Avalanche detector with ultraclean response for time-resolved photon counting," *IEEE Journal of Quantum Electronics*, vol. 34, pp. 817–821, 1998.

[25] E. Sciacca, S. Lombardo, M. Mazzillo, G. Condorelli, D. Sanfilippo, A.Contissa, M. Belluso, F. Torrisi, S. Billotta, A. Campisi, L. Cosentino, A.Piazza, G. Fallica, P. Finocchiaro, F. Musumeci, S. Privitera, S. Tudisco, G.Bonanno, and E. Rimini, "Arrays of Geiger mode avalanche photodiodes," *IEEE Photonics Technology Letters*, vol. 18, pp. 1633–1635, 2006.

[26] A. Rochas, A. R. Pauchard, P. A. Besse, D. Pantic, Z. Prijic, and R. S. Popovic, "Low-noise silicon avalanche photodiodes fabricated in conventional CMOS technologies," *IEEE Transaction on Electron Devices*, vol. 49, pp. 387–394, 2002.

[27] A. Rochas, M. Gani, B. Furrer, P. A. Besse, R. S. Popovic, G. Ribordy, and N.Gisin, "Single photon detector fabricated in a complementary metal-oxide-semiconductor high-voltage technology," *Review of scientific instruments*, vol. 74, pp. 3263–3270, 2003.

[28] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*. Hoboken, NJ: Wiley-Interscience, 2007.

[29] R. J. Locker and G. C. Huth, "A new ionizing radiation detection concept which employs semiconductor avalanche amplification and tunnel diode element," *Applied Physics Letters*, vol. 9, pp. 227–230, 1966.

[30] E. Gramsch, M. Szawlowski, S. Zhang, and M. Madden, "Fast, high-density avalanche photodiode-array," *IEEE Transactions on Nuclear Science*, vol. 41, pp. 762–766, 1994.

[31] J. P. Donnelly, K. A. McIntosh, D. C. Oakley, A. Napoleone, S. H. Groves, S. Vernon, L. J. Mahoney, K. Molvar, J. Mahan, J. C. Aversa, E. K. Duerr, Z. L. Liau, B. F. Aull, and D. C. Shaver, "1 um Geiger-mode detector development," *Proceedings of the SPIE*, vol. 5791, pp. 286–292, 2005.

[32] H. Yilmaz, "Optimization and surface-charge sensitivity of high-voltage blocking structures with shallow junctions," *IEEE Transactions on Electron Devices*, vol. 38, pp. 1666–1675, 1991.

[33] R. H. Haitz, A. Goetzberger, R. M. Scarlett, and W. Shockley, "Avalanche effects in silicon p-n junctions .I. localized photomultiplication studies on microplasmas," *Journal of Applied Physics*, vol. 34, pp. 1581–1589, 1963.

[34] A. Rochas, "Single photon avalanche diodes in CMOS technology." PhD Thesis, Ecole Polytechnique Federale de Lausanne (EPFL), November 2003.

[35] M. Gersbach, "Single-photon detector arrays for time-resolved fluorescence imaging." PhD Thesis, Ecole Polytechnique Federale de Lausanne (EPFL), November 2009.

[36] W. Shockley and W. T. Read, "Statistics of the recombination holes and electrons," *Physics review*, vol. 87, pp. 835–842, 1952.

[37] M. C. Beard, K. P. Knutsen, P. Yu, J. M. Luther, Q. Song, W. K. Metzger, R. J. Ellingson, and A. J. Nozik, "Multiple exciton generation in colloidal silicon nanocrystals," *Nano Letters*, vol. 7, no. 8, pp. 2506–2512, 2007.

[38] A. Franceschetti, J. M. An, and A. Zunger, "Impact ionization can explain carrier multiplication in PbSe quantum dots," *Nano letters*, vol. 6, no. 10, pp. 2191–2195, 2006.

[39] R. D. Schaller, M. A. Petruska, and V. I. Klimov, "Effect of electronic structure on carrier multiplication efficiency: Comparative study of PbSe and CdSe nanocrystals," *Applied Physics letters*, vol. 87, p. 253102, 2005.

[40] M. Belluso, M. C. Mazzillo, G. Bonanno, S. Billotta, S. Scuderi, and A. A. Cali, "SPAD array detectors for astrophysical applications," *Memorie della Societa Astronomica Italiana Supplement*, vol. 9, pp. 430–433, 2006.

[41] F. Zappa, A. Gulinatti, P. Maccagnani, S. Tisa, and S. Cova, "SPADA: Single-photon avalanche diode arrays," *IEEE photonics technology letters*, vol. 17, pp. 657–659, 2005.

[42] C. Y. Chang and S. M. Sze, *ULSI technology*. New York: McGraw-Hill, 1996.

[43] IBM Corporation, Mixed Signal Technology Development Department, IBM Microelectronics Division, *CMOS7RF design manual*, 2005.

[44] S. Cova, A. Longoni, A. Andreoni, and R. Cubeddu, "A semiconductor detector for measuring ultraweak fluorescence decays with 70 ps FWHM resolution," *IEEE journal of quantum electronics*, vol. 19, pp. 630–634, 1983.

[45] C. L. Niclass, "Single photon image sensor in CMOS: Picosecond resolution for three dimensional imaging." PhD Thesis, Ecole Polytechnique Federale de Lausanne, 2008.

[46] A. Lacaita, M. Mastrapasqua, M. Ghioni, and S. Vanoli, "Observation of avalanche propagation by multiplication assisted diffusion in p-n junctions," *Applied Physics Letters*, vol. 57, pp. 489–491, 1990.

[47] A. Spinelli and A. L. Lacaita, "Physics and numerical simulation of single photon avalanche diodes," *IEEE Transaction on Electron Devices*, vol. 44, pp. 1931–1943, 1997.

[48] A. Lacaita, S. Cova, A. Spinelli, and F. Zappa, "Photon-assisted avalanche spreading in reach-through photodiodes," *Applied Physics Letters*, vol. 62, pp. 606–608, 1993.

[49] A. Lacaita and M. Mastrapasqua, "Strong dependence of time resolution on detector diameter in single-photon avalanche diodes," *Electronic letters*, vol. 26, pp. 2053–2054, 1990.

[50] J. A. Richarson, L. A. Grant, and R. K. Henderson, "Low dark count single-photon avalanche diode structure compatible with standard nanometer scale CMOS technology," *IEEE Photonics Technology Letters*, vol. 21, no. 14, pp. 1020–1022, 2009.

[51] I. Rech, I. Labanca, M. Ghioni, and S. Cova, "Modified single photon counting modules for optimal timing performance," *Review of Scientific instruments*, vol. 77, pp. 4–5, 2006.

[52] M. Gersbach, Y. Maruyama, E. Labonne, J. Richardson, R. Walker, L. Grant, R. Henderson, F. Borghetti, D. Stoppa, and E. Charbon, "A parallel 32x32 time-to-digital converter array fabricated in a 130nm imaging CMOS technology," in *European Solid-State Circuits Conference (ESSCIRC 09)*, (Athens, Greece), 2009.

[53] S. Cova, A. Lacaita, and G. Ripamonti, "Trapping phenomena in avalanche photodiodes on nanosecond scale," *IEEE Electron device letters*, vol. 12, pp. 685–687, 1991.

[54] S. Donati, E. Randone, M. Fathi, J.-H. Lee, E. Charbon, and G. Martini, "Uniformity of concentration factor and back focal length in molded polymer microlens arrays," in *Conference on Lasers and Electro-Optics (CLEO)*, (OSA), 2010.

[55] M. Gersbach, D. L. Boiko, C. Niclass, C. C. H. Petersen, and E. Charbon, "Fast-fluorescence dynamics in nonratiometric calcium indicators," *Optics Letters*, vol. 34, no. 3, pp. 362–364, 2009.

[56] R. Newman, "Visible light from a silicon p-n junction," *Physical review*, vol. 100, pp. 700–703, 1955.

[57] A. Rochas, M. Gosch, A. Serov, P. A. Besse, R. S. Popovic, T. Lasser, and R. Rigler, "First fully integrated 2D-array of single photon detectors in standard CMOS technology," *IEEE Photonics technology letters*, vol. 15, pp. 963–965, 2003.

[58] C. Niclass, A. Rochas, P. A. Besse, and E. Charbon, "Design and characterization of a CMOS 3-D image sensor based on single photon avalanche diodes," *Journal of Solid-state Circuits*, vol. 40, pp. 1847–1854, 2005.

[59] R. H. Haitz, "Studies on optical coupling between silicon p-n junctions," *Solid state electronics*, vol. 8, pp. 417–425, 1965.

[60] I. Prochazka, K. Hamal, L. Kral, and J. Blazej, "Silicon photon counting detector optical cross-talk effect," in *Photonics devices and systems III*, 2007.

[61] W. J. Kindt, H. W. van Zeijl, and S. Middlehoek, "Optical crosstalk in Geiger mode avalanche photodiode arrays: modeling, prevention and measurement," in *28th European solid state*.

[62] I. Rech, A. Ingargiola, R. Spinelli, I. Labanca, S. Marangoni, M. Ghioni, and S. Cova, "Optical crosstalk in single photon avalanche diode arrays: a new complete model," *Optics Express*, vol. 16, no. 12, pp. 8381–8394, 2008.

[63] C. Niclass, M. Sergio, and E. Charbon, "A single photon avalanche diode array fabricated in 0.35um CMOS and based on an event-driven readout for TCSPC experiments," in *Proc. SPIE optics*, (Southwest), 2006.

[64] N. Faramarzpour, M. J. Deen, S. Shirani, and Q. Fang, "Fully integrated single photon avalanche diode detector in standard CMOS 0.18-um technology," *IEEE Transaction on Electronic Devices*, vol. 55(3), pp. 760–767, 2008.

[65] C. Niclass and M. Soga, "A miniature actively recharged single-photon detector free of afterpulsing effects with 6ns dead time in a 0.18um CMOS technology," in *International Electron Device Meeting*, (San Francisco), 2010.

[66] M. Gersbach, J. Richardson, E. Mazaleyrat, S. Hardilier, C. Niclass, R. Henderson, L. Grant, and E. Charbon, "A low noise single photon detector implemented in a 130 nm CMOS imaging process," *Solid state Electronics*, vol. 53, no. 7, pp. 803–808, 2009.

[67] J. Richardson, L. Grant, and R. Henderson, "A low-dark count single-photon avalanche diode structure compatible with standard nanometer scale CMOS technology," in *International Image Sensor Workshop*, (Bergen, Norway), 2009.

[68] J. B. Kuo and S. C. Lin, *Low-voltage SOI CMOS VLSI Devices and Circuits*. Newyork: Wiley and Sons, 2001.

[69] T. Hamamoto, "Sidewall damage in a silicon substrate caused by trench etching," *Applied Physics Letters*, vol. 58, pp. 2942–2944, 1991.

[70] H. Kwon, M. Kang, B. Park, J. Lee, and S. Park, "The analysis of dark signals in CMOS APS imagers from the characterization of test structures," *IEEE Transaction on Electronic Devices*, vol. 51, no. 2, pp. 178–184, 2004.

[71] H. Finkelstein, M. J. Hsu, and S. C. Esener, "STI- bounded single-photon avalanche diode in a deep-submicrometer CMOS technology," *IEEE Electron Device Letters*, vol. 27, pp. 887–890, 2006.

[72] R. H. Hadfield, "Single-photon detectors for optical quantum information applications," *Nature Photonics*, vol. 3, pp. 696–705, 2009.

[73] W. Becker, *Advanced Time-Correlated Single Photon Counting Techniques*. Germany: Springer, 2005.

[74] A. Migdall, "Introduction to journal of modern optics special issue on single-photon: detectors, applications, and measurement methods," *Journal of Modern Optics*, vol. 51, pp. 1265–1266, 2004.

[75] X. Michalet, O. H. W. Siegmund, J. V. Vallerga, P. Jelinsky, J. E. Millaud, and S. Weiss, "Detectors for single-molecule fluorescence imaging and spectroscopy," *Journal of Modern Optics*, vol. 54, pp. 239–281, 2007.

[76] R. T. Thew, S. Tanzilli, L. Krainer, S. C. Zeller, A. Rochas, I. Rech, S. Cova, H. Zbinden, and N. Gisin, "Low jitter up-conversion detectors for telecom wavelength GHz QKD," *New Journal of Physics*, vol. 8, pp. 32–33, 2006.

[77] D. Rosenberg, S. W. Nam, P. A. Hiskett, C. G. Peterson, R. J. Hughes, J. E. Nordholt, A. E. Lita, and A. J. Miller, "Quantum key distribution at telecom wavelengths with noise-free detectors," *Applied Physics Letters*, vol. 88, p. 021108, 2006.

[78] D. K. Joseph, T. J. Huppert, M. A. Franceschini, and D. A. Boas, "Diffuse optical tomography system to image brain activation with improved spatial resolution and validation with functional magnetic resonance imaging," *Applied Optics*, vol. 45, pp. 8142–8151, 2006.

[79] P. Andersson, "Long-range three-dimensional imaging using range-gated laser radar images," *Optical Engineering*, vol. 45, p. 034301, 2006.

[80] M. A. Albota, R. M. Heinrichs, D. G. Kocher, D. G. Fouche, B. E. Player, M. E. OBrien, B. F. Aull, J. J. Zayhowski, J. Mooney, B. C. Willard, and R. R. Carlson, "Three-dimensional imaging laser radar with a photon-counting avalanche photodiode array and microchip laser," *Applied Optics*, vol. 41, pp. 7671–7678, 2002.

[81] R. M. Marino and W. R. J. Davis, "Jigsaw: a foliage-penetrating 3D imaging laser radar system," *Lincoln Laboratory Journal*, vol. 15, pp. 23–36, 2005.

[82] G. S. Buller, R. D. Harkins, A. McCarthy, P. A. Hiskett, G. R. MacKinnon, G. Smith, R. Sung, A. M. Wallace, R. A. Lamb, K. D. Ridley, and J. G. Rarity, "Multiple wavelength time-of-flight sensor based on time-correlated single-photon counting," *Review of Scientific Instruments*, vol. 76, p. 083112, 2005.

[83] W. E. Moerner and D. P. Fromm, "Methods of single-molecule fluorescence spectroscopy and microscopy," *Review of Scientific Instruments*, vol. 74, pp. 3597–3619, 2003.

[84] M. Gersbach, R. Trimananda, Y. Maruyama, M. Fishburn, D. Stoppa, J. Richardson, R. Walker, R. K. Henderson, and E. Charbon, "High frame-rate TCSPC-FLIM using a novel spad-based image sensor," in *SPIE Optics and Photonics*, (San Diego), 2010.

[85] V. Ntziachristos, A. G. Yodh, M. Schnall, and B. Chance, "Concurrent mri and diffuse optical tomography of breast after indocyanine green enhancement," in *Proceedings of the National Academy of Sciences of the United States of America*, (vol.97), 2000.

[86] W. Becker, *The bh TCSPC handbook*. Berlin GmbH: Becker and Hickl, 2006.

[87] A. P. Gibson, J. C. Hebden, and S. R. Arridge, "Recent advances in diffuse optical imaging," *Physics in Medicine and Biology*, vol. 50, pp. 1–43, 2005.

[88] P. Mengel, L. Listl, B. Konig, C. Topfer, M. Pellkofer, W. Brockherde, B. Hosticka, O. Elkhalili, O. Schrey, and W. Ulfig, "Three-dimensional CMOS image sensor for pedestrian protection and collision mitigation.," in *The international forum on advanced microsystems for automotive applications*, (AMAA), 2006.

[89] S. Hsu, S. Acharya, A. Rafii, and R. New, "Performance of a time-of-flight range camera for intelligent vehicle safety applications.," in *The international forum on advanced microsystems for automotive applications*, (AMAA), 2006.

[90] H. Du, "Application of 3D range camera in virtual human-computer interfaces." PhD Thesis, Ecole Polytechnique Federale de Lausanne (EPFL), 2008.

[91] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, "Quantum cryptography," *Review of modern physics*, vol. 74, pp. 145–195, 2002.

[92] D. Achilles, C. Silberhorn, and I. A. Walmsley, "Direct, loss-tolerant characterization of nonclassical photon statistics," *Physics Review Letters*, vol. 97, p. 043602, 2006.

[93] X. Li, X. Ma, L. Quan, L. Yang, L. Cui, and X. Guo, "Quantum efficiency measurement of single-photon detectors using photon pairs generated in optical fibers," *Journal of Optical Society of America B*, vol. 27, no. 9, pp. 1857–1865, 2010.

[94] M. A. Alam. Lecture Notes on Solid State Devices, 2009.

[95] C. G. Streetman and S. Banerjee, *Solid State Electronic Devices.* Upper Saddle River, NJ: Prentice Hall, 2000.

[96] S. M. Sze, *Semiconductor devices: Physics and technology.* Hoboken, NJ: Wiley, 1985.

[97] G. A. M. Hurkx, D. B. M. Klassen, and M. P. G. Knuvers, "A new recombination model for device simulation including tunneling," *IEEE Transaction on Electron Devices*, vol. 39, no. 2, pp. 2090–2098, 1992.

[98] W. G. Oldham, R. S. Samuelson, and P. Antognetti, "Triggering phenomena in avalanche diodes," *IEEE Transaction on Electron Devices*, vol. 19, no. 9, pp. 1056–1060, 1972.

[99] M. A. Karami, M. Gersbach, H. J. Yoon, and E. Charbon, "A new single-photon avalanche diode in 90nm standard CMOS technology," *Optics Express*, vol. 18, no. 21, pp. 803–808, 2010.

[100] V. Milovanovic, "Advanced breakdown modeling for solid-state circuit design." PhD Thesis, Technical University of Delft), July 2010.

[101] P. M. Solomon, D. J. Frank, J. Jopling, C. D. Emic, O. Dokumaci, and P. R. W. E. Haensch, "Tunnel current measurement on p/n junction diodes and implication for future device design," in *Internation Electron Device Meeting (IEDM)*, (San Francisco), 2003.

[102] P. M. Solomon, J. Jopling, D. Frank, C. D. Emic, O. Dokumaci, P. Ronsheim, and W. E. Haensch, "Universal tunneling behavior in technologically relevant p-n junction diodes," *Applied Physics Letters*, vol. 95, no. 1, pp. 5800–5812, 2004.

[103] E. O. Kane, "Zener tunneling in semiconductors," *Journal of Physics and Chemistry Solids*, vol. 12, pp. 181–188, 1959.

[104] SYNOPSYS, US, *Taurus Medici device simulator Manual*, 2007. March 2007.

[105] R. F. Pierret, *Advanced Semiconductor Fundamentals*. Upper Saddle River, NJ: Prentice Hall, 2002.

[106] E. Simoen and C. Claeys, "Random telegraph signal: a local probe for single point defect studies in solid-state devices," *Materials Science and Engineering B*, vol. 91, pp. 136–143, 2002.

[107] M. J. Kirton and M. J. Uren, "Noise in solid-state microstructures: A new perspective on individual defects, interface states and low frequency (1/f) noise," *Advances in Physics*, vol. 38, pp. 367–468, 1989.

[108] I. H. Hopkins and G. R. Hopkinson, "Further measurements of random telegraph signals in proton irradiated CCDs," *Transactions on nuclear science*, vol. 42, pp. 2074–2081, 1995.

[109] J. Bogarets, B. Dierickx, and R. Mertens, "Random telegraph signals in radiation-hardened CMOS active pixel sensors," *Transactions on nuclear science*, vol. 49, pp. 249–257, 2002.

[110] F. Buchinger, A. Kyle, J. K. P. Lee, and C. Webb, "Identification of individual bistable defects in avalanche photodiodes," *Applied Physics Letters*, vol. 66, no. 18, pp. 2367–2369, 1995.

[111] G. R. Hopkinsons, V. Gouffon, and A. Mohammadzadeh, "Random telegraph signals in proton irradiated CCDs and APS," *Transaction on nuclear science*, vol. 55, no. 4, pp. 2197–2204, 2008.

[112] M. A. Karami, C. Niclass, and E. Charbon, "Random telegraph signal in single-photon avalanche diodes," in *International Image Sensor Workshop (IISW)*, (Bergen), 2009.

[113] M. A. Karami, L. Carrara, C. Niclass, M. Fishburn, and E. Charbon, "RTS noise characterization in single photon avalanche diodes," *IEEE Electron Device Letters*, vol. 31, no. 7, pp. 692–694, 2010.

[114] S. R. Li, W. McMahon, Y. Lu, and Y. Lee, "RTS noise characterization in flash cells," *IEEE Electron Device letters*, vol. 29, pp. 106–108, 2008.

[115] L. Carrara, C. Niclass, N. Scheidegger, H. Shea, and E. Charbon, "A gamma, X-ray and high energy proton radiation tolerant CIS for space applications," in *International Solid-solid state circuits ISSCC*, (San Francisco), 2009.

[116] G. R. Hopkinsons, V. Gouffon, and A. Mohammadzadeh, "Random telegraph signals in proton irradiated CCDs and APS," *Transaction on nuclear science*, vol. 55, no. 4, pp. 2197–2204, 2008.

[117] I. H. Hopkins and G. R. Hopkinsons, "Random telegraph signals from proton-irradiated CCDs," *Transaction on nuclear science*, vol. 40, no. 6, pp. 1567–1574, 1993.

[118] G. I. Andersson, M. O. Andersson, and O. Engstrom, "Discrete conductance fluctuations in silicon emitter junctions due to defect clustering and evidence for structural changes by high-energy electron irradiation and annealing," *Journal of Applied Physics*, vol. 72, no. 7, pp. 2680–2691, 1992.

[119] C. Niclass, M. Gersbach, R. Henderson, L. Grant, and E. Charbon, "Single photon avalanche diode implemented in 130-nm CMOS technology," *Selected Topics on Quantum Electronics*, vol. 13, no. 4, pp. 863–869, 2007.

[120] R. Henderson, E. Webster, R. Walker, J. Richardson, and L. Grant, "A 3x3, 5um pitch, 3-transistor single photon avalanche diode array with integrated 11v bias generation in 90nm CMOS technology," in *Proceedings of International Electron Device Meeting (IEDM)*, (San Francisco), 2010.

[121] Link Research Corporation, Japan, *SPECTRA 3-dimensional device simulator Manual*, 2009. Release 5.4.

Samenvatting

Quantum parasitaire effecten en miniaturisatie van Single Photon Avalanche Diodes in diep-submicron technologieën worden in dit proefschrift in detail bestudeerd. Tunneling ruis en Random Telegraph Signal (RTS) ruis zijn de twee belangrijkste parasitaire effecten die ruimschoots worden aangepakt. Terwijl de fundamentele vergelijkingen voor de tunnelingstelsnelheid in SPADs worden ingediend, heeft de tunneling studie uitgevoerd door het vergelijken van twee verschillende topologieën van SPADs op dezelfde technologie. De tunneling evaluatie is uitgevoerd door het implementeren van een nieuwe N-tub borgring SPAD in 0.35 μm standaard CMOS technologie. De NTGR SPAD is theoretisch gekenmerkt door simulaties en experimenten. De tunneling ruis effecten zijn in de IV karakteristieken geverifieerd en Dark Count Rate reacties van de nieuwe NTGR SPAD.

Terwijl de NTGR SPAD heeft 3% verbetering in de maximale Photon Detection Probability (PDP) aangetoond, de timing jitter Full Half Maximum (FWHM) wordt van 80 ps in p-tub borgring SPAD tot 152 ps naar de NTGR SPAD verhoogd. De doorslagspanning is verlaagd met 4.1 V vanwege het gebruik van de verschillende dopingprofielen voor NTGR SPAD implementatie. Lawine fotoemissie van het actieve deel van de NTGR SPAD aangetoond het vermogen voor enkele foton detectie in vergelijking met de andere diodes, die aantonen photoemssion in guard ringen als een teken van Premature Edge Breakdown (PEB).

Het RTS-gedrag van donkere telsnelheid (DCR) was het tweede parasitaire effect die is onderzocht in dit proefschrift. Dit is gekarakteriseerd als een bistabiliteit en multistabiliteit van DCR. RTS gedrag van DCR is waargenomen in een SPAD gefabriceerd in 0.8 μm CMOS-technologie in vier proton-bestraalde SPADs ontworpen en gefabriceerd in 0.35 μm CMOS-technologie. Om het beste van onze kennis, dit was de eerste keer dat RTS gedrag van DCR is gemeld bij SPADs in een CMOS-technologie. RTS kenmerken zijn experimenteel geëvalueerd en theoretisch geverifieerd met betrekking tot de vertekening en de temperatuur.

De RTS-hypothese van de fluctuaties is geverifieerd door metingen die zijn in uitstekende overeenstemming met de theorie. De RTS gedrag verwacht wordt dat om te invloeden op groot schaal van SPAD detectoren die worden gebruikt voor 3D-imaging-toepassingen. Het aantonen van de functionaliteit van SPADs in diepsubmicron technologieën is de tweede belangrijkste uitdaging van dit proefschrift. We hebben aangetoond dat singlephoton detectoren kunnen worden vervaardigd in commercieel deep-submicron CMOS-processen. Miniaturisering is onderzocht met de eerste SPAD ontworpen en wordt succesvol getest in technologieën kleiner dan 130nm.

De voorgestelde constructies, uitgevoerd in 90nm standaard CMOS technologie, voortgekomen uit een systematische studie gericht op miniaturisatie, terwijl het optimaliseren van de algemene prestaties. Het bewaker ring ontwerp is het resultaat van een uitgebreide modellering inspanning gericht op beperken van de vermenigvuldiging regio binnen een welomschreven gebied waar het elektrische veld groter is dan de kritische waarde voor de impact ionisatie. Door het implementeren van verschillende SPAD structuren, hebben we onderzocht de geometrische trade-offs die betrokken zijn bij het ontwerp van deep-submicron SPADs. Talrijke SPADs met verschillende arrangementen van doping lagen en verschillende maten van guard ring zijn uitgevoerd. Onder hen, zo veel als 45 structuren waren functioneel met een reeks van goed gedefinieerde, reproduceerbare doorslagspanningen. De gemplementeerde structuren zijn ook gesimuleerd en gekarakteriseerd. De detectors zijn voorzien van een achthoekige vermenigvuldiging regio en een bewaker ring om te voorkomen dat BEP een standaard masker set gebruikt, exclusief. De voorgestelde structuur blijkt uit een systematische studie gericht op miniaturisatie, terwijl het optimaliseert de algemene prestaties.

De apparaten vertonen een DCR van 8,1 kHz, een maximum PDP van 14% bij maximale overschrijding bias. Bij 0,13 V van overmaat bias, een PDP van 9% en een jitter van 398ps bij een golflengte van 637 nm, werden gemeten in kamertemperatuur. Een afterpulsing waarschijnlijkheid van 32% werd gemeten bij de nominale dode tijd.Het belangrijkste zwakke punt van SPADs dat in dit proefschrift is aangetoond, in 90nm-technologie, was het ontbreken van isolatie met de ondergrond, die verschillende oplossingen zijn voorgesteld beperken van de NTGR SPAD, niet te communiceren met het substraat. Simulatie op basis analyse van SPADs ontworpen 65nm CMOS-technologie, is beschreven in het laatste deel. Het ontwerp van SPADs in 65nm-technologie is gebaseerd op de karakterisering resultaten van functionele SPADs in 90nm-technologie.

Terwijl de uitvoering van de SPADs in beeldvorming CMOS-technologieën kan tot betere prestaties leiden te wijten aan de speciale lagen. De demonstratie van SPAD in een standaard CMOS technologie zonder gebruik van de beeldvorming van doping lagen is een van de belangrijkste resultaten van dit proefschrift geweest. Vanwege het isolement problemen zoals hierboven beschreven, zijn nieuwe circuits uitgevoerd voor de SPAD afschrikken en de pols detectie.

Publications

Journal papers:

M.A. Karami, M. Gersbach, H.J. Yoon, and E. Charbon, "A new single-photon avalanche diode in 90nm standard CMOS technology ", Optics Express, Vol. 18, Issue 21, pp. 22158-22166 (2010).

M.A. Karami, L. Carrara, C. Niclass, M. Fishburn, E. Charbon, "RTS noise characterization in Single Photon Avalanche Diodes ", Electron Device Letters, Vol. 31, Issue 7, (2010).

M.A. Karami, and E. Charbon, "Tunneling in Submicron CMOS Single-Photon Avalanche Diodes", submitted to Electron Device Letters

Conference papers:

M.A. Karami, H.J. Yoon, and E. Charbon, "Single-photon avalanche diodes in sub- 100nm CMOS technologies", to be appeared in the proceedings of International Image Sensor Workshop, IISW , 2011.

M.A. Karami, M. Gersbach, and E. Charbon, "A new Single photon Avalanche-diode fabricated in 90nm CMOS standard technology ", SPIE conference in single-photon imaging 2010.

M.A. Karami, C. Niclass, and E. Charbon, "RTS noise in Single-photon Avalnche Diodes ", International Image Sensor Workshop, IISW, Bergen, Norway, 2009.

Patents:

E. Charbon, **M.A. Karami**, "Plasmonic Integrated Circuits ", Patent pending, European patent no.EP10157793, Delft, Netehrlands, March 2010.

About the Author

Mohammad Azim Karami was born in Tehran, Iran on September 12, 1983. He got Bachlor degree in Electronics from Shahid Beheshti University in 2005, and Master degree in Electronic devices from University of Tehran in 2007. His master thesis was focused on nanowire lasers modeling and analysis. From January 2008, He joined the group of Prof. Charbon in EPFL, Switzerland, and TU Delft, designing Single-Photon Avalanche Diodes, testing, and post processing them. During his Ph.D he worked on the idea of plasmonic ICs implementation, replacing electrons with surface plasmon polaritons.